

**TC
AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
TEZLİ YÜKSEK LİSANS PROGRAMI**

**PISA 2015 MATEMATİK OKUR-YAZARLIĞI TESTİNİN FARKLI
DİLLERİ KONUŞAN ÜLKELER ARASINDA ÖLÇME DEĞİŞMEZLİĞİNİN
İNCELENMESİ**

YÜKSEK LİSANS TEZİ

ESMA ASAR

Antalya, 2019

**TC
AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
TEZLİ YÜKSEK LİSANS PROGRAMI**

**PISA 2015 MATEMATİK OKUR-YAZARLIĞI TESTİNİN FARKLI
DİLLERİ KONUŞAN ÜLKELER ARASINDA ÖLÇME DEĞİŞMEZLİĞİNİN
İNCELENMESİ**

YÜKSEK LİSANS TEZİ

ESMA ASAR

Danışman:

Dr. HAKAN KOĞAR

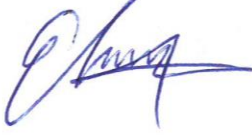
Antalya, 2019

DOĐRULUK BEYANI

Yüksek lisans tezi olarak sunduĐum bu alıřmayı, bilimsel ahlak ve geleneklere aykırı dűşecek bir yol ve yardıma bařvurmaksızın yazdıĐımı, yararlandıĐım eserlerin kaynakalardan gösterilenlerden oluřtuĐunu ve bu eserleri her kullanımında alıntı yaparak yararlandıĐımı belirtir; bunu onurumla doĐrularım. Enstitű tarafından belli bir zamana baĐlı olmaksızın, tezimle ilgili yaptıĐım bu beyana aykırı bir durumun saptanması durumunda, ortaya ıkacak tüm ahlaki ve hukuki sonulara katlanacaĐımı bildiririm.

29/05/2019

Esmas ASAR



İMZA SAYFASI

T.C.

AKDENİZ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Esme..... Aso.c. 'nın bu çalışması 03.05.2019... tarihinde jürimiz tarafından Eğitim Bilimleri Anabilim Dalı Ölme ve Değerlendirme Tezli Yüksek Lisans Programında Yüksek Lisans Tezi olarak oy birliği/oy çokluğu ile kabul edilmiştir

İMZA

Başkan : (Unvan)Adı Soyadı Dr. Öğr. Ü. Seyma UYAR
(Çalıştığı Kurum, Fakülte, Bölüm)
Mehmet Atif Ersoy Üniversitesi, Eğitim Fakültesi,
Eğitim Bilimleri Bölümü

Üye : (Unvan)Adı Soyadı Doç. Dr. Bayram BIGAÇ
(Çalıştığı Kurum, Fakülte, Bölüm)
Akdeniz Üniversitesi Eğitim Fakültesi, Eğitim
Bilimleri Bölümü

Üye (Danışman) : (Unvan)Adı Soyadı Dr. Öğr. Ü. Hakan KOBAK
(Çalıştığı Kurum, Fakülte, Bölüm)
Akdeniz Üniversitesi, Eğitim Fakültesi, Eğitim
Bilimleri Bölümü

YÜKSEK LİSANS TEZİNİN ADI:

ONAY: Bu tez, Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından uygun görülmüş ve Enstitü Yönetim Kurulunun tarihli ve sayılı kararıyla kabul edilmiştir.

Doç. Dr. Ramazan KARATAŞ

Enstitü Müdürü

TEŐEKKÜR

Beni bugünlere getiren canım aileme, tezimi dört gözle bitirmemi beklerken benden maddi manevi desteęini esirgemeyen canım babama, tez sürecinde hep yanımda olan canım eőime ve danıőman hocam Hakan KOęAR a sonsuz teőekkürler.

ÖZET

PISA 2015 MATEMATİK OKUR-YAZARLIĞI TESTİNİN FARKLI DİLLERİ KONUŞAN ÜLKELER ARASINDA ÖLÇME DEĞİŞMEZLİĞİNİN İNCELENMESİ

ASAR, Esmâ

Yüksek Lisans, Eğitim Bilimleri Ana Bilim Dalı

Tez Yöneticisi: Dr. Öğr. Üyesi Hakan KOĞAR

Mayıs, 2019 67 sayfa

Bu araştırmanın amacı, PISA 2015 Matematik okuryazarlığı testinin farklı dilleri konuşan ülkeler arasında ölçme değişmezliğinin incelenmesidir. İlişkisel tarama modeli ile yürütülen bu araştırmada 54 numaralı kitapçık kullanılmış ve araştırmaya, farklı dilleri konuşan 12 ülke ve 2816 öğrenci dahil edilmiştir. İlk olarak bütün gruplar için ayrı ayrı doğrulayıcı faktör analizi uygulanarak, ölçme modelinin bütün gruplarda doğrulandığı belirlenmiştir. Ardından kovaryans matrislerinin eşitliğini test etmek için ölçme değişmezliğinin aşamaları sırası ile uygulanmıştır. Modelin yapısal ve metrik değişmezlik aşamasını sağladığı ancak ölçek değişmezliği aşamasını sağlamadığı görülmüştür. Böylelikle analiz sonuçlandırılmış ve araştırma sonuçları, farklı dilleri konuşan 12 ülke arasında ölçme değişmezliğinin sağlanmadığını göstermiştir.

Anahtar Sözcükler: Matematik okuryazarlığı, Ölçme değişmezliği, Çoklu gruplu doğrulayıcı faktör analizi, PISA 2015

ABSTRACT

AN INVESTIGATION OF THE MEASUREMENT OF THE MEASUREMENT BETWEEN SPEAKING COUNTRIES OF DIFFERENT LANGUAGES OF THE PISA 2015 MATHEMATICS LITERACY TEST

ASAR, Esma

Master of Arts, Department of Educational Sciences

Supervisor: Dr. Öğr. Üyesi Hakan KOĞAR

May, 2019 67 page

The aim of this study is to examine the measurement invariance of the PISA 2015 Mathematics literacy test among the countries that speak different languages. In this study conducted with the relational survey model, booklet 54 was used and 12 countries and 2816 students who speak different languages were included in the study. First, confirmatory factor analysis was performed for all groups and it was determined that the measurement model was confirmed in all groups. Then, in order to test the equality of the covariance matrices, the stages of measurement invariance were applied respectively. It was found that the model provided the stage of structural and metric invariance but did not provide the stage invariance. As a result, the analysis was finalized and the results of the study showed that there was no measurement invariance between the 12 countries speaking different languages.

Keywords: Mathematics literacy, Measurement invariance, Multiple group confirmatory factor analysis, PISA 2015

İÇİNDEKİLER

TEŞEKKÜR.....	i
ÖZET.....	ii
ABSTRACT	iii
İÇİNDEKİLER.....	iv
TABLO, ŞEKİL VE GRAFİKLER LİSTESİ.....	vi
KISALTMALAR LİSTESİ.....	vii

BÖLÜM I

GİRİŞ

1.1. Problem Durumu	1
1.1.1. Problem Cümlesi	5
1.2. Araştırmanın Amacı	5
1.3. Araştırmanın Önemi	6
1.4. Araştırmanın Varsayımları	6
1.5. Araştırmanın Sınırlılıkları	6
1.6. Tanımlar	7

BÖLÜM II

KURAMSAL ÇERÇEVE İLE İLGİLİ ARAŞTIRMALAR

2.1. Ölçme Değişmezliği	8
2.2. Doğrulayıcı Faktör Analizi	10
2.3. Çoklu Grup Doğrulayıcı Faktör Analizi	11
2.4. Yapısal / Şekilsel / Biçimsel Değişmezlik (Configural invariance)	12
2.5. Metrik/ Zayıf / Faktöriyel Değişmezlik (Weak invariance / Metric invariance)	13
2.6. Ölçek / Güçlü / Skalar Değişmezlik (Scalar / Strong / invariance)	13
2.7. Katı / Tam Değişmezlik (Strict invariance)	14
2.8. Kısmi Ölçme Değişmezliği (Partial Measurement Invariance)	14
2.9. Yapısal Eşitlik Modeli.....	15
2.9.1 Yapısal Eşitlik Modeline İlişkin Temel Kavramlar.....	16
2.9.2 Yapısal Eşitlik Modellerinin Betimlenmesi ve Model Uyumunun Değerlendirilmesi	18

2.9.2.1. Modelin Betimlenmesi ve Tanımlanması.....	18
2.9.2.2. Model Uyumunun Değerlendirilmesi.....	19
2.10. İlgili Araştırmalar.....	22
2.10.1. Türkiye’de ölçme değişmezliği ile ilgili araştırmalar.....	22
2.10.2. Dünya’da ölçme değişmezliği ile ilgili araştırmalar.....	24

BÖLÜM III

YÖNTEM

3.1. Araştırmanın Türü.....	26
3.2. Çalışma Grubu	26
3.3. Veri Toplama Araçları	27
3.4. Verilerin Analizi.....	28

BÖLÜM IV

BULGULAR

4.1. Araştırmaya dahil olan Ülkeler İçin Ölçme Değişmezliğine İlişkin Bulgular...31	
4.1.1. Şekilsel Değişmezlik.....	33
4.1.2. Metrik Değişmezlik.....	33
4.1.3. Ölçek Değişmezliği.....	34

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

5.1. Sonuç ve Tartışma	35
5.2. Öneriler	36
KAYNAKÇA	39
EKLER	45
ÖZGEÇMİŞ	59
İNTİHAL RAPORU	60

TABLO, ŐEKİL VE GRAFİKLER LİSTESİ

Tablo1. Model Uyum İndekslerinin Kriterleri Ve Kabulu	22
Tablo 2. Ülkelere Göre Örnekleme Dağılımı	27
Tablo 3. Analizde Kullanılan Matematik Okuryazarlığı Maddeleri	28
Tablo 4. Analize Dahil Edilen Ülkeler İçin Normallik Testi Sonuçları	31
Tablo 5. Analize Dahil Edilen Ülkeler İçin DFA Testi Sonuçları	32
Tablo 6. Ölçme Değişmezliği Aşamalarına İlişkin Uyum İndeksileri	33
Őekil 1. Matematik Okuryazarlığı Deęerlendirme Aşamaları	4
Őekil 2. Ölçme Modeli	17
Őekil 3. Yapısal Model	17
Őekil 4. Őekil Deęişmezliği Path Diagramı Sonuçları	34
Őekil 5. Metrik Deęişmezliği Path Diagramı Sonuçları	35
Grafik 1	2
Grafik 2	2

KISALTMALAR LİSTESİ

DFA: Doğrulayıcı Faktör Analizi

ÇGDFA: Çok Gruplu Doğrulayıcı Faktör Analizi

YEM: Yapısal Eşitlik Modeli

PISA: Uluslararası Öğrenci Değerlendirme Çalışması (Program for International Student Assessment)

OECD: Ekonomik İşbirliği ve Kalkınma Örgütü (Organisation for Economic Co-operation and Development)

BÖLÜM I

GİRİŞ

1.1. Problem Durumu

Uluslararası düzeyde 15 yaş grubu öğrencilerin, Matematik, Fen Bilimleri okuryazarlığı ve Okuma Becerilerini, öğrencilerin motivasyonlarını, kendileri hakkındaki görüşlerini araştırmak ve öğrenme biçimlerini, okul ortamlarını ve aileleri ile ilgili tutumlarının ülkeden ülkeye farklılıklarını incelemek için kurulmuş olan ve merkezi Fransa'nın başkenti Paris' de bulunan OECD Eğitim Direktörlüğü'ne bağlı PISA projesi, 2000 yılından bu yana uluslararası alanda uygulanmaya devam etmektedir. Bu test sürecinde amaçlanan,

- a) Öğrencilerin okulda öğrendikleri bilgi ve becerileri günlük yaşamda kullanma yetisini ölçmek,
- b) Öğrencileri daha iyi tanımak; onların öğrenme isteklerini, derslerdeki performanslarını ve öğrenme ortamları ile ilgili tercihlerini daha açık bir biçimde ortaya koymak,
- c) Zorunlu eğitimin sonunda öğrencilerin hangi boyutlarda geliştiğini değerlendirmek, bilgilerini gerçek yaşam durumlarına uygulayabilmek ve topluma katılım için donanımlı hale gelip gelmediklerini kontrol edebilmek için tasarlanmıştır (PISA, 2017).

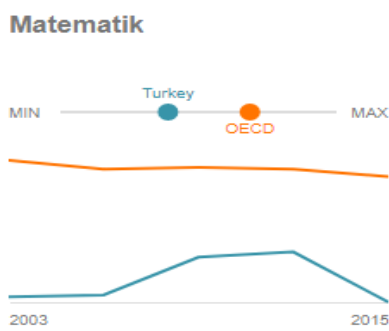
PISA testi ülkelerin okul programlarından bağımsız olarak hareket eder. Üç yılda bir tekrarlanan testin içeriğinde fen bilimleri, okuma becerisi ve matematik gibi ana konulara odaklanılır (2015 de işbirlikli problem çözme becerisi de dahil edilmiştir). PISA'nın her turunda, temel alanlardan bir tanesi, toplam test süresinin yaklaşık üçte ikisini alarak ayrıntılı olarak test edilir. 2000 ve 2009 yıllarında temel alan; Okuma becerisi, 2006 ve 2015 te temel alan; Bilim, 2003 ve 2012 yılları arasında temel alan; Matematik olmuştur. Araştırmanın geçtiği PISA 2015 testinde değerlendirme alanları; okuma, matematik ve işbirliğine dayalı problem çözme becerisi ile fen bilimleri olmuştur. Ayrıca isteğe bağlı olarak bazı ülkelerde finansal okuryazarlığa ilişkin bir değerlendirme yapılmıştır. 2015 yılında teste 72 ülke katılmış

ve yaklaşık 540 bin öğrenci, 29 milyon 15 yaş grubu öğrenciyi temsilen değerlendirmeyi tamamlamıştır (OECD, 2017). 2015 PISA uygulamasının bu seneki özelliği 57 ülkede ilk defa tamamen bilgisayar tabanlı uygulama olarak yapılmış olmasıdır. Bu uygulamada 66 kitapçık kullanılmıştır. 15 ülke ise kağıt- kalem değerlendirmesine katılmış ve 30 kitapçık uygulanmıştır (Taş, 2016).

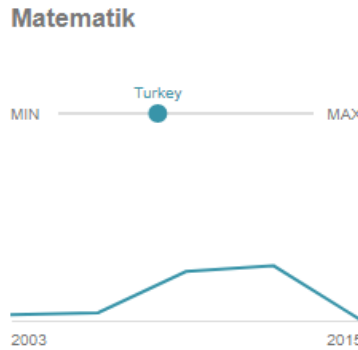
PISA 2015 uygulamasında Türkiye’den 61 il, 187 okul ve 5895 öğrenci katılmıştır. Bu örneklemin büyük çoğunluğu 9. ve 10. Sınıf öğrencilerinden oluşmaktadır. Uygulamaya ülkemiz açısından bakıldığında 2015 yılı değerlendirmesinden bir önceki değerlendirme olan 2012 PISA değerlendirmesinde ülkemiz, matematik okuryazarlığı alanında OECD ülkelerine göre düşük bir ortalama göstermiştir ve öğrencilerin büyük bir kısmı 2. düzeyin altındaki yeterlilik düzeyinde kalmıştır (Dibek, 2017). PISA 2012 uygulamasında OECD ülkelerinin matematik okuryazarlığı ortalaması 494 iken, Türkiye ortalaması 448 dir. 2015 değerlendirilmesi incelendiğinde ise OECD ülkelerinin matematik okuryazarlığı ortalama puanı 493 iken Türkiye verilerin ortalaması 425 olmuştur. Ancak ortalama tüm ülkeler için değerlendirildiğinde 2012 yılında 470 e ve 2015 yılında 461 e düştüğü görülmektedir (Taş, 2015).

PISA 2015 Türkiye sonuçları değerlendirilirken dikkate alınması gereken nokta; “Türkiye’de öğrenciler 8. Sınıf sonrasında liselere giriş sınavına tabi tutulmaktadır. PISA 2015 uygulamasında ise büyük oranda 10. Sınıf öğrenciler yer almaktadır. Bu uygulamasına giren öğrencilerin 8. Sınıf sınav başarısı, örneklemin temsil ettiği beklenen dağılımı göstermemektedir (Taş, 2015).

Aşağıdaki ise Türkiye örnekleminin yıllara göre dağılımı ve OECD ülkelerine göre durumu grafiklerle gösterilmiştir (PISA, 2018)



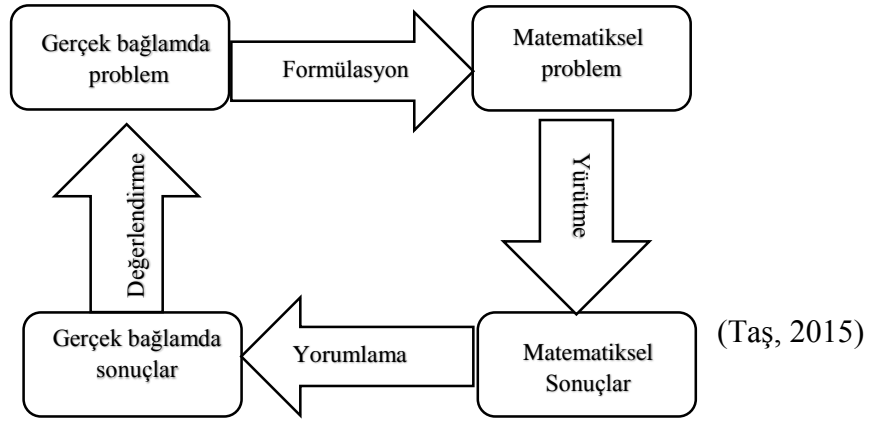
Grafik 1



Grafik 2

PISA, uzun vadede, çeşitli ülkelerdeki öğrencilerin bilgi ve becerileri ile her ülkenin farklı demografik alt gruplarındaki eğilimlerin izlenmesi için bir bilgi birikimi geliştirilmesine yol açan bir program olmuştur. Dünyanın dört bir yanındaki politikacılar, kendi ülkesindeki öğrencilerin bilgi ve becerilerini, diğer katılımcı ülkeler ile karşılaştırmak için PISA bulgularını kullanmakta ve iyileştirmeler için ölçütler oluşturmaktadır. Bu şekilde, kendi eğitim sistemlerinin göreceli güçlü ve zayıf yönlerini belirlemekte ve yeni politikalar oluşturmaktadır (OECD, 2017).

Milli Eğitim Bakanlığı'nın hazırlamış olduğu PISA 2015 ulusal raporunda matematik okuryazarlığı, bireylerin matematiği farklı şekillerde kullanma, yorumlama ve formüle etme kapasitesinin ölçülmesi olarak tanımlanmıştır. Bireylerin PISA testinde başarılı olması matematiksel mantık kurabilmeleriyle doğru orantılıdır. Bu alanda yeterlilik ise, bireylerin matematiğin dünyada oynadığı rolü fark etmesi ve yapıcı vatandaşlar olmak için sağlam dayanakları olan karar ve yargıları vermesiyle mümkün olur. Bu bağlamda PISA uygulaması soru maddelerini gerçek yaşam durumlarına dayanarak hazırlar. PISA testinin amacı, bireylerin öğrendiklerinden nasıl anlamlar çıkardıklarını, olası başka durumlarda bile matematiksel bilgilerini nasıl uyguladıklarını değerlendirmektir. Bu değerlendirme “Nicelik, belirsizlik ve veri, değişim ve ilişkiler, uzay ve şekil” olmak üzere dört farklı matematiksel içerik kategorileri ile yapılmaktadır. Ölçülen temel matematiksel yeterlilikler ise “Matematikselleştirme, Matematiksel araç kullanma, İletişim, Strateji üretme, Temsil biçimleri, Muhakeme ve argüman, Sembolik dil ve işlemler kullanma” olarak tanımlanmıştır. Ayrıca matematiksel süreç ise “Formülasyon, Yürütme, Yorumlama ve Değerlendirme” aşamalarıyla tamamlanmıştır. Matematik okuryazarlığının PISA tarafından değerlendirilmesinin genel özellikleri Şekil 1. de gösterilmiştir (Taş, 2015).



Şekil 1. Matematik Okuryazarlığı Değerlendirme Aşamaları

PISA, bir çok farklı kültürden öğrencilerin dahil olduğu ve ülkelerin dikkate aldığı bir test olduğu için sonuçların doğru ve yansız olması testin geçerliliği açısından çok önemlidir. Bu testin sonuçları, her bireyin kendine has özellikleri olduğu için farklı olacaktır. Ancak sonuçlardaki farklılığın sadece bireylerden kaynaklandığını düşünmek ve bu şekilde açıklamak çoğu zaman doğru değildir. Çünkü ortaya çıkan bu farklılık ölçme aracından kaynaklanıyor olabilir (Başusta, 2010). PISA gibi bireylerin duyuşsal ve bilişsel özelliklerinin değerlendirildiği uluslararası sınavlarda karşılanması gereken varsayımların en öncelikli olanlarından biri ölçme değişmezliğidir (Gierl, 2000). Bireylerin ve ülkelerin karşılaştırıldığı bu tür çalışmaların geçerli olabilmesi için ilgili yapılar bakımından gruplar arası ölçme değişmezliğinin sağlanmış olması gerekmektedir (Somer, 2009). Bahsi geçen Ölçme Değişmezliği, Meredith ve Millsap (1992) tarafından şu şekilde tanımlanmaktadır: “Rastgele seçilmiş bir bireyin belli bir puana sahip olma durumunun, hangi grupta yer aldığından bağımsız olmasıdır”. Herdman (1998) ise ölçme değişmezliğini, “Psikolojik bir ölçme aracının psikometrik özelliklerinin, farklı gruplar açısından denkliliğinin biçimsel olarak değerlendirilmesi” olarak tanımlamaktadır (Akt. Alatlı, 2018). Sadece geleneksel yaklaşımlar ile, gruplar arasındaki farklılığı değerlendirmek hatalı sonuçlar doğurabilir. Bu ise ortaya konmak istenen davranışın, yanlış tanımlanmasına neden olacaktır. Bu nedenle gruplar arası karşılaştırmaların ölçme değişmezliğinin aşamaları ile desteklenmesi gerekir. Ayrıca araştırmaların değişmezlik çalışmalarıyla desteklenmesi de önemli bilgiler sağlayabilir. Genele yönelik doğru bir değerlendirme yapılması, özelliklerin grup bakımından açıklanmasında yararlı olacaktır (Başusta, 2010).

Araştırmanın başında da belirtildiği gibi PISA 2015 matematik okuryazarlığı uygulamasına 72 ülke katılmış ve uygulama, katılımcı ülkelerin dilleri olmak üzere 82 ayrı dile çevrilmiştir (OECD, 2017). Farklı ülkelerde, farklı dillerde hazırlanmış ölçekler, dil uzmanlarıyla birlikte başka bir dile çevrilse ve dille bağlantılı olan ayrılıkların giderilmesi için çalışmalar yapılırsa bile, söz konusu ölçeğin farklı ülkelerde aynı anlama gelmesi ve bireyler tarafından aynı şekilde yorumlanması sağlanmış olmayacaktır (Cheug ve Rensvold; 1999). Bu bağlamda araştırmanın problemini, PISA 2015 uygulamasındaki matematik okuryazarlığının farklı dili konuşan ülkeler için ölçme değişmezliğinin incelenmesi oluşturmaktadır.

1.1.1. Problem Cümlesi

PISA 2015 matematik okuryazarlığı testinde farklı dilleri konuşan “Türkiye (Türkçe), Çin (Çince), Almanya (Almanca), İtalya (İtalyanca), İsrail (İbranice), Katar (Arapça), Litvanya (Litvanca), Rusya (Rusça), Singapur (İngilizce), Birleşik Krallık (Galce), Tayland (Tayca), Slovakya (Slovakça)” olmak üzere 12 ülke arasında “Ölçme değişmezliği” sağlanmakta mıdır? Bu doğrultuda şu sorulara cevap aranmıştır.

1.2. Araştırmanın Amacı

Bu araştırmanın amacı, PISA 2015 Matematik okuryazarlığı testinin, farklı dilleri konuşan ülkeler için ölçme değişmezliğini test ederek bütün gruplarda aynı yapıyı ölçüp ölçmediğini belirlemektir.

1.3. Araştırmanın Önemi

Ulusal ve uluslararası araştırmalar incelendiğinde farklı kültürlerde ve farklı gruplarda bulunan bireylerin karşılaştırıldığı çok sayıda çalışma olduğu görülmektedir. Ancak bu çalışmalarda kullanılan ölçme araçlarının eşdeğerliğine dair araştırmaların yeterince ortaya konmadığı dikkat çekmektedir (Yandı, 2017). Eğitim açısından bakıldığında ise öğrenci performansını değerlendirmek ve performans farklılıklarını açıklamaya yardımcı olabilecek öğrenci, aile ve kurumsal faktörler hakkında veri toplamak için kapsamlı ve titiz çalışan PISA programı, farklı kültürlerin karşılaştırıldığı çalışmaların en önemlilerinden biri sayılabilir. PISA programı hakkında alınacak kararlar, değerlendirmelerin kapsamı ve niteliği, toplanacak arka

plan bilgileri, katılımcı ülkelerin önde gelen uzmanları tarafından tartışılır ve ortak politika odaklı çıkarlar temelinde hükümetler tarafından ortaklaşa yönlendirilir (OECD, 2017). Söz konusu ülkeler, bu projeye maddi ve manevi anlamda destek verdikleri için, sonuçları ülke gündemlerinde önemli yer tutar. Ayrıca yüksek performans gösteren ülkeler, diğer ülkelerin eğitim politikalarına bir örnek teşkil etmeye başlar. “Eğitim” gibi önemli bir konuda bireyleri ve ülke yöneticilerini yönlendirebilen böyle bir testin son derece geçerli ve güvenilir olması gereklidir. Marsh ve ark. (2006) a göre; karşılaştırmalarda kullanılan ölçeklerin kültürlerarasında geçerli olması, bahsi geçen ölçeklerin güvenilirliğini ve yapılan araştırmaların genellenebilirliğini gösterir. Kültürlerarası geçerlilik ise, ölçme değişmezliği ile sağlanabilmektedir. Eğer kullanılan modelde ölçme değişmezliği sağlanamamışsa, 15 yaş grubu bireylerinin performansları arasındaki fark, ölçme aracının kendinden kaynaklanıyor olabilir. Bu bağlamda bu çalışmada, uluslararası testler için “Ölçme değişmezliği” nin önemi vurgulanmaya çalışılmıştır.

1.4. Araştırmanın Varsayımları

PISA 2015 matematik testine, farklı ülkelere öğrencilerin doğru ve samimi yanıtlar verdiği ve araştırma için seçilen örneklemin evreni en iyi biçimde yansıttığı varsayılmıştır.

1.5. Araştırmanın Sınırlılıkları

PISA 2015 uygulamasında yer alan alt boyutlardan yalnızca Matematik okuryazarlığı ölçekleri ele alınmıştır. PISA 2015 testi uygulamasına 72 ülke katılmış, araştırma 12 ülke (Türkiye, Çin, Almanya, İtalya, İsrail, Katar, Litvanya, Rusya, Singapur, Birleşik Krallık, Tayland, Slovakya) ile sınırlandırılmıştır. Ayrıca araştırma 82 ayrı dile çevrilmiş ancak 12 dil (Türkçe, Çince, Almanca, İtalyanca, İbranice, Arapça, Litvanca, Rusça, İngilizce, Galce, Tayca, Slovakça) ile araştırma yapılmıştır. PISA 2015 uygulamasında 96 ayrı kitapçık kullanılmış ancak araştırma tek kitapçık (54 no’lu) ile sınırlandırılmıştır. 54 numaralı kitapçıkta ise 22 madde analize dahil edilmiştir.

1.6. Tanımlar

Ölçme Değişmezliği: Araştırmada kullanılan ölçekte bulunan gizil yapının farklı gruplar arasındaki eşdeğerliğidir.

BÖLÜM II

KURAMSAL ÇERÇEVE İLE İLGİLİ ARAŞTIRMALAR

2.1. Ölçme Değişmezliği

Son yıllarda sosyal bilimler alanındaki gelişmelerin artmasıyla birlikte bireylerin bilişsel yetenekleri, düşünme stilleri ve kişilik özellikleri gruplar arası karşılaştırmalarda çokça kullanılır duruma gelmiştir. Bireylerde bulunan bu tip özellikler genellikle doğrudan gözlenemediği için bir anket yardımıyla en iyi şekilde ölçülebilmektedir. Bu anketler aracılığıyla ölçülen gizil yapının bütün gruplarda aynı anlama gelmesi gruplar arası karşılaştırmanın ön koşulu olmasının yanında, varyans / kovaryans değerleri gibi grupların örneklem kestirim karşılaştırmaları, grup farklılıklarını ortaya koyar nitelikte olmalıdır (Gregorich, 2006). Literatürde ölçme değişmezliği, ölçme eşdeğerliği ya da ölçüm değişmezliği olarak adlandırılan bu husus, ölçekte bulunan gizil yapının farklı gruplar veya alt gruplar arasındaki eşdeğerliği olarak tanımlanır. Grupların, gizil değişkenler modelinde birbirleriyle karşılaştırılması durumunda, bu modelin parametrelerinin, grup üyeleri arasında aynı kalması olarak ifade edilebilir (AERA, APA ve NCME, 1999). Başka bir deyişle ölçme değişmezliğinden bahsedilebilmesi için gözlenen değişkenin durumsal dağılımı ve gizil değişkenin değeri, bütün gruplarda veya alt örneklemelerde eşit olması gerekir (Başusta, 2010). Yapının işlevsel hale getirilmesinin ardından, ölçekte elde edilen puanların farklı şartlar altında da aynı anlama gelip gelmeyeceği ölçme değişmezliği tarafından sorgulanır. Bahsi geçen farklı şartlar; “ölçme teknikleri”, “ölçüm yapılan zaman” ve “ölçüm yapılan grubun kültürü” gibi durumlardır (Kline, 2016).

Bu bağlamda ölçüm değişmezliği,

- a. Psikolojik yapıların her alt grup için genelleştirilebilir olduğuna,
- b. Bu yapıların alt grup farklılıklarından etkilenmez olduğuna,
- c. Ölçme önyargılarının ve hatalarının en düşük düzeyde olduğuna işaret eder (Demir, 2017).

Ölçme değişmezliğinin en büyük avantajı, yapıların bütün gruplarda aynı şekilde çalıştığını sayısal verilerle göstermesidir. Bu şekilde; yanlılık, farklı gruplarda yapıya ilişkin örtük ortalama, kovaryans ve varyansların karşılaştırılabilirliği ortaya

konulabilmektedir (Little, 1997). Ölçme değişmezliği mevcut olduğunda, gizil değişken ile gözlenen değişken arasındaki faktör yükü gruplar arasında sabit kalır. Ölçme değişmezliğinin sağlanmadığı durumlarda, gözlenen puanların farklarını anlamlı bir şekilde yorumlamak zordur. Kültürler arası anlamlı karşılaştırmalarda ölçüm denkliği bir ön şart olduğu için bir araştırmacı bir ölçme aracının çeşitli sürümlerinin içerik, talimatlar ve yanıt seçenekleri açısından özdeş olduğundan emin olmalıdır (Raju, Laffitte ve Byrne, 2002).

Günümüzde araştırmacılar gruplar arası karşılaştırmaların ölçme değişmezliği çalışmalarlarıyla daha sık biçimde ilgilenmeye başlamışlardır. Ölçme değişmezliğinin sağlanmaması demek, anketlerden elde edilen sonuçların gerçekliğinin sorgulanması demektir. Çünkü ölçme eşdeğerliği sağlanamazsa gruplar arasındaki farklılıkların, ölçülmek istenilen davranışın farklılığından mı yoksa psikolojik etmenlerden mi kaynaklandığı saptanamaz (Ölçüoğlu, 2015)

Ölçme değişmezliğini oluşturmak için iki popüler yöntem kullanılmaktadır. Bu yöntemlerden biri Yapısal Eşitlik Modeli (YEM), diğeri Madde Tepki Kuramı (MTK) dır (Raju, Laffitte ve Byrne, 2002). Meade ve Lautenschlager (2004) e göre bu yöntemler genel anlamda benzer amaçlar için kullanılsa da, yöntemsel olarak farklıdırlar. Her iki yaklaşımda sınırlılıkları ve avantajları vardır (Akt. Kıbrıslıoğlu,2015). Ancak Ölçme değişmezliğini değerlendirmek için yapısal eşitlik modeli kullanılır. Bu değerlendirme için yapısal eşitlik modellemesinin kullanımına 1970 yılında İsveçli bir istatistikçi olan Karl Jöreskog öncülük etmiştir (Jöreskog, 1971). Jöreskog ve Sörbom (1993), gruplar arası ölçme değişmezliğini tespit etmeyi sağlayan faktör-analitik ve ki-kare testleri ile ilişkili bir sistem geliştirmiş ve bu yöntem Çok Gruplu Doğrulayıcı Faktör Analizi (ÇGDFA) olarak isimlendirilmiştir.

Ölçme Değişmezliğinin formülize edilmiş hali

Tanım 1. Rastgele seçilmiş X değişkeni, gizil W değişkenleri üzerindeki örnek uzayda bütün (x, w, v) değerleri için $|F(x, w, v) - F(x, w)| = 0$ koşulunu sağlıyorsa ölçme değişmezliği sağlanır (Meredith, 1993).

Sözel olarak ifade edilecek olursa, gizil W değişkenini ölçen rastgele seçilmiş bir X değişkeni var olsun. X ile W arasındaki ilişkiyi bir örnek verilerek olasılıkla ifade edelim. Gizil değişkenden W=3 puan alan bireyin testten X=10 alma olasılığı $F(X=10)$

W=3) olsun. Bu görünen F değeri örnek uzaydaki her bir w, x ve v değerleri için eşitse ölçme değişmezliği sağlanır.

Ölçmeye dahil edilen bireylere ilişkin “k” farklı grup tanımlanmış olsun. Bu gruplara göre bireylerin karşılaştırılması amaçlanmakta ise ifade edilen genel teorik yapı aşağıdaki gibidir.

$$X_k = T_k + \Lambda_k \xi_k + \delta_k$$

X_k = Gözlenen değişken matrisi

T_k = Örtük yapılar ile gözlenen değişkenler arasındaki katsayı vektörü

Λ_k =Faktör yükleri matrisi

ξ_k = Gözlenen bağımsız değişkenlerin hata vektörü

δ_k = Örtük faktörler vektörü

(Doğan, 2014).

2.2. Doğrulayıcı Faktör Analizi (DFA)

Maruyama (1998)’e göre; tanımlanmış ve sınırlandırılmış bir yapının, model olarak doğrulanıp doğrulanmadığını tespit etmek için geliştirilen analiz, doğrulayıcı faktör analizidir. Ayrıca bazen bu analiz, “Kuramsal yapı”nın doğrulanması anlamında da kullanılmaktadır (Akt. Şekercioğlu ve ark., 2016). Doğrulayıcı faktör analizinde amaç, değişkenler arasındaki ilişkilere ait yapısal hipotezlerin test edilmesi ve doğrulanmasıdır. Bu çerçeveden bakıldığında doğrulayıcı faktör analizinde, hipotezler doğrultusunda değişkenlerin faktörlerle ve faktörlerinde kendi aralarında kurulan ilişkilerin incelenmesine odaklanılır (Şekercioğlu ve ark., 2016). Başka bir deyişle DFA, önceden var olan faktör modeliyle verilerin arasındaki uyumu değerlendirir (Floyd, Widaman, 1995).

DFA’nın geleneksel kökeni genel faktör analizlerine dayanır. Ancak DFA geleneksel testlerin aksine tek bir anlamlılık değeri vermez (Şekercioğlu ve ark., 2016). DFA, faktör analizinden; farklı tahmin yöntemleri, maksimum benzerlik yöntemi ve faktör çözümünün yeterliliğini değerlendirme gibi farklı standartlarla ayrılır (Floyd, Widaman, 1995).

Brown (2006) a göre; bir modele DFA yapılmadan önce, modelde bulunan faktör sayısı, faktör çiftlerinin hangilerinin birbirleriyle ilişkili olduğu, gözlenen değişken yüklerinin hangi faktör altında toplanacağı bellidir. Kline (2005)’ e göre ise DFA’ nın adımları şu şekilde tanımlanmıştır:

- a. Modelin betimlenmesi
- b. Modelin tanımlanması için uygun koşulların sağlanması
- c. Model doğrultusunda verilerin toplanması
- d. Belirlenen modelin analizi
- e. Model uyumunun değerlendirilmesi
- f. Modelin uyumsuzluk sürecinde modifikasyon aşamalarının gerçekleştirilmesi
- g. Sonuçların yorumlanması (akt. Ayvallı, 2016).

2.3. Çoklu Grup Doğrulayıcı Faktör Analizi (ÇGDFA)

Jöreskog ve ark. (2001) e göre, gruplar, milletler, dinler, kültürler ya da ekonomilerden elde edilen verilerle geliştirilen ya da uyarlanan ölçeklerin faktör yapılarının birden fazla örneklemlerde eşitliği, yapısal eşitlik modeliyle test edilebilir (Akt. Şekercioğlu, 2009). Bu amaçla kullanılan ve yapısal eşitlik modelinin istatistiksel tekniklerinden biri olan, çoklu gruplu doğrulayıcı faktör analizi, birden fazla gruplar için yapısal parametrelerin denkleğinin aynı anda test edilmesini sağlar (Şekercioğlu, 2009).

Gruplar arası farklılığın belirlenmesinde kullanılan Çoklu gruplu DFA son derece güçlü bir yöntemdir. Bu analiz, her bir gruptan alınan ayrı ayrı örneklem gruplarının denkleğini belirlemek için karşılaştırmalar yapar. Aynı popülasyondan alınan örneklem gruplarının çapraz geçerlilikte olması, başka bir deyişle alınan örneklemlerin aynı özelliklere sahip olması beklenir (Hair ve diğ., 2010; Yıldız, 2017).

Çoklu grup modelinin iki önemli fonksiyonu vardır. Birincisi, denklik testlerini bütün gruplar arasında aynı anda yürütür. Başka bir deyişle, parametreler tüm gruplar için eş zamanlı olarak tahmin edilir. İkincisi, değişmezlik testi için belirlenen yapısal modelin uyumu, daha sonra belirtilecek olan tüm değişmezlik modellerinin karşılaştırılması için temel değeri sağlar (Byrne, 2008).

Gruplar arası ölçme denkleğinin gösterilmesi ve ölçeklerin yapısal eşdeğerliğinin incelenmesinde kullanılan çoklu gruplu doğrulayıcı faktör analizi, literatürde önerilen bir metod olup, çok faktörlü ölçeklere uygulanması ve çok kategorili verilerin analizinde kullanılabilmesi gibi birçok yönden diğer metotlara tercih edilebilmektedir. Ayrıca bu analizle ölçekte bulunan kültüre özgü maddeler tespit edilebilmektedir (Byrne ve Stewart, 2006). Madde fonksiyonu analizlerinde, ÇG-DFA'nın Madde Tepki Kuramı (MTK) ile etkisinin aynı olduğu, hatta çok

kategorili verilerde daha iyi performans sergilediği yapılan çalışmalarla kanıtlanmıştır (Stark, Chernyshenko ve Drasgow, 2006; Gelbal, Asil, 2012).

Jöreskog ve Sörbom (1993), farklı gruplar arasında ölçme değişmezliğini tespit etmeyi sağlayan faktör-analitik ve ki-kare testleri ile ilişkili bir sistem geliştirmiş ve bu yöntem Çok Gruplu Doğrulayıcı Faktör Analizi olarak isimlendirilmiştir.

Vanderberg ve Lance (2000) e göre ölçme değişmezliğinin test edilebilmesi için beş aşamalı bir süreç izlenmelidir. Aynı zamanda bu aşamalar ölçme değişmezliğinin türleridir. Aşamaların her biri sırasıyla test edilir. Bir hipotez test edildikten sonra yeterli sonuçlar alınırsa diğer hipoteze geçilir. Her aşamaya, bir önceki aşamaya göre artarak devam eden gruplar arası eşitlik sınırlamaları dahil edilmektedir. Bu aşamalar;

- 1.Şekilsel / Yapısal / Biçimsel Değişmezlik (Configural invariance)
- 2.Metrik / Zayıf Değişmezlik (Weak invariance / Metric invariance)
- 3.Ölçek / Güçlü / Skalar Değişmezlik (Scalar / Strong /invariance)
- 4.Katı / Tam Değişmezlik (Strict invariance)
- 5.Kısmi Ölçme Değişmezliği (Partial measurement invariance) dir.

2.4. Yapısal / Şekilsel / Biçimsel Değişmezlik (Configural invariance)

Vandenberg ve Lance (2000)' e göre, ölçme değişmezliğinin başlangıç ve en temel aşaması yapısal değişmezliktir. Bu aşamada, grupların faktör yapılarının eşit olup olmadığı incelenir. Yapısal değişmezlik, gruplar arasında aynı faktör modelinin olmasını gerektirir (Zumbo ve ark., 2007). Bryne (2008)'e göre ise ölçme değişmezliğinin bu aşaması, tahmin edilen parametrelerin gruplar arasında ne ölçüde geçerli olduğunu inceler. Yapısal değişmezliğinin sağlanması, gruplar arası karşılaştırmanın yapılabileceği anlamına gelmez ancak ölçme değişmezliği aşamalarının ilkinin sağlandığını ve diğer aşamaya geçilebileceğini gösterir (Kline, 2011; Akt. Kıbrıslıoğlu, 2015).

Yapısal değişmezlik araştırmacıları, farklı grupların, test maddelerini cevaplamak için aynı kavramsal yapıyı kullanıp kullanmadıklarını inceler. Faktör yükleri, regresyon sabitleri ve hata varyansları serbest bırakılarak, yalnızca gruplar için faktör sayısı ve faktör yapısı belirlenir (Vandenberg & Lance, 2000). Bu aşamada gruplar aynı anda tek bir model içinde test edilir ve bütün gruplarda aynı yapının var olup olmadığına bakılır. Bu modelde gruplar arası faktör yapısının benzer olması

yeterlidir, özdeş olması aranmaz (Byrne ve Steward, 2009). Eğer yapısal değişmezliğe dair bulgular elde edilirse, ölçme aracının bütün gruplar için aynı yapıyı ölçtüğü sonucuna varılır. Eğer sonuçlar, yapısal değişmezliği desteklemiyorsa, gözlemlenen ölçütlerin her grup içindeki farklı yapıları temsil ettiği anlaşılır. Bu nedenle yapısal değişmezlik modeli ölçme değişmezliğinin daha ileri testlerinin değerlendirildiği bir temel model olarak da görülmüştür. Çünkü yapısal değişmezlik gruplar arasında gösterilmezse daha ileri testler gerekçelendirilemez (Vandenberg & Lance, 2000).

2.5. Metrik/ Zayıf / Faktöriyel Değişmezlik (Weak invariance / Metric invariance)

Literatürde “Zayıf değişmezlik” olarak da geçen metrik değişmezlik aşamasında faktör yüklerinin gruplar arasında sabit olup olmadığı test edilmektedir. Ayrıca bu aşamada parametre sınırlaması ilk defa kullanılmaya başlanır (Alatlı, 2018). Bu değişmezlik modeli, modeldeki yapısal ilişkilerin gruplar arasında aynı olup olmadığıyla ve yapının, belirli değişkenler üzerine karşılaştırma yapılabilmesi için alt gruplarda anlamlı ve uygun olup olmadığıyla ilgilenir (Öğretmen, 2009). Byrne (1998)’e göre ise bu modelde, gruplardaki bireylerin maddeleri benzer şekilde yorumlayıp yorumlamadığı test edilmektedir. Zayıf değişmezlik aşamasında, faktör yapısına ek olarak faktör yükleri de sınırlandırılır ve ölçme birimlerinin yani ard arda iki kategori arası uzaklıkların gruplar arası eşitliği sınanır (Li, Wu, Zumbo, 2007).

Metrik değişmezlik şu şekilde belirlenir; bütün gruplar arasından rastgele bir biçimde referans olarak seçilen grubun faktör varyansı sabitlenir. Diğer grupların faktör varyansları serbest bırakılır (Millsap ve Olivera-Aguilar, 2012). Varyansların arasında anlamlı bir fark olmadığı tespit edildiğinde yapının gruplar arasında benzer şekilde açığa çıktığı sonucuna varılabilir (Kline, 2011).

2.6. Ölçek / Güçlü / Skalar Değişmezlik (Scalar / Strong / invariance)

Ölçme değişmezliğinin güçlü bir seviyesi olan ölçek değişmezliği, metrik değişmezliğinin sağlanmasından sonraki aşamadır. Ölçek değişmezliğinde, metrik değişmezliğinde aranan faktör yüklerinin denkliğine ek olarak regresyon sabitinin de gruplar arasında eşitliği gerekir (Wu ve ark.,2007). Başka bir deyişle; ölçme aracının maddeleri için oluşturulan regresyon denklemlerindeki sabit değer, bütün gruplar için değişmez olduğu test edilmektedir (Önen, 2007). Bu modelde, öncelikle grupların

faktör puanı sıfırlanır. Ardından regresyon sabitinin eşit olup olmadığına bakılır (Vanderbeg ve Lance, 2000).

Bu değişmezlik türü sağlandığında gözlenen değişkenlerin ve faktör yüklerinin ortalamaları karşılaştırılabilir demektir (Gregorich, 2006). Sağlanmadığı durumlarda ise “Maddelerin faktörleri açıklama düzeyi gruplar arasında farklılık gösterir” sonucuna varılabilir.

2.7. Katı / Tam Değişmezlik (Strict invariance)

Ölçme değişmezliğinin dördüncü aşaması olan Tam/Katı değişmezlik, “Bütün gruplarda ortak olan ve belirlenmiş faktörlere verilen tepkilerin varyanslarının gruplar arasındaki denklığı” anlamına gelir (Meredith ve Teresi, 2006). Faktör yüklerinin ve regresyon sabitlerinin denklığıne ek olarak regresyon kalıntı varyanslarının yani hata varyanslarının eşitliğini gerektirmektedir (Wu ve ark., 2007). Bu model, çok sınırlandırılmış bir model olduğu için uygulamada karşılanması zordur (Doğan, Uyar, 2011). Widemann ve Rice (1997), bu sınırlandırmanın fazla olmasının sebebinin, gizil değişkenden gelen varyans arttıkça hata varyanslarının da artmasından kaynaklandığını söylemiştir.

2.8. Kısmi Ölçme Değişmezliği (Partial Measurement Invariance)

Ölçme değişmezliğinin sağlanmadığı durumlarda başvurulan yöntem olan kısmi ölçme değişmezliği, değişmezliğin sağlanmamasının sebebi olan maddelerin belirlenmesi ve kontrol edilmesi imkanlarını araştırmacılara vermektedir (Vandenberg ve Lance, 2000). Yapısal değişmezlik aşaması haricinde, ölçme değişmezliğinin tüm aşamalarında “kısmi ölçme değişmezliği” nin incelenebileceği söylenir (Önen, 2009).

Modele parametreler eklenerek, çıkartılarak ya da birleştirilerek modifikasyon işlemi yapılır (Reise ve ark., 1993). Ancak bu araştırmada ölçme değişmezliğinin ilk dört aşaması incelenmiş olup, herhangi bir modifikasyon işlemi gerçekleştirilmemiştir. Bu nedenle kısmi ölçme değişmezliği diğer modeller kadar ayrıntılı açıklanmamıştır.

2.9. Yapısal Eşitlik Modeli

Bir araştırma yöntemi olarak kullanılan yapısal eşitlik modeli, araştırmacılara sunduğu avantajlar sebebiyle gün geçtikçe kullanımı artan bir teknik olmuştur. İlk

olarak sosyal bilimler alanında kullanılmaya başlasa da, günümüzde tıp, eğitim, ticaret gibi alanlarda da kullanım sıklığı giderek artmaktadır. Gün geçtikçe genişleyen bir çalışma alanı olan Yapısal Eşitlik Modeli bir çok farklı grupta aynı özelliğin ölçülebildiği varsayımı kabul edilerek kurulur (Başusta, 2010). Tamamen teoriye dayalı bir yöntem olan YEM'in amacı, daha önceden tahmin edilen nedensel ilişkilerin elde edilen verilerle doğrulanma düzeyini tespit etmektir (Yener, 2007).

YEM'in günümüzdeki kullanımına ilişkin çalışmalar ilk olarak 1970 li yıllarda görülmeye başlanmıştır ve hemen ardından sosyal bilimciler tarafından yoğun ilgi görmüştür. Bu gelişmeden sonra sosyal bilimler alanındaki araştırma sorularının karmaşıklığı ve özgünlüğü artmış, çalışmalar netlik kazanmıştır. Teknolojinin ilerlemesi ve bilgisayar programlarının çalışmalara dahil edilmesiyle yapısal eşitlik modeline ilgi, giderek hız kazanmıştır (Şekercioğlu ve ark., 2016). Yapısal Eşitlik modeli, ilk yıllarda karışık modellerde başarılı olması, modeldeki ilişkilere göre yeni tavsiyelerde bulunması ve ölçme hatalarını göz önünde bulunduruyor olması gibi nedenlerle araştırmacıların dikkatini çekmeyi başarmıştır (Dursun, Kocagöz, 2010). Gözlenen ve gözlenemeyen değişkenler arasındaki dolaylı veya doğrudan ilişkileri tek bir modelde analiz ediyor olması da zamanla onu günümüzün en çok kullanılan yöntemlerinden biri haline getirmiştir (Meydan, Şeşen 2015 akt. Demir,2016, Bahadır, 2012).

Psikolojik çalışmalar genellikle gizil değişkenlerden oluşur. Gözlenemeyen yapıların yorumlanması ve açıklanması gizil değişkenlerin analizi ile yapılır. Söz konusu analiz yapısal eşitlik modeli ile gerçekleşir (Kaynak, 2012).

Yapısal eşitlik modeli, sosyal ve davranış bilimlerindeki pek çok problemin çözüm kapısı olmuştur (Jöreskog ve Sorbom, 1982). Çünkü yapısal eşitlik modeli kavramı sosyal bilim araştırmalarında uzun süredir tartışılmalı nedensellik kavramına, yeni bir boyut kazandırmıştır. Bahsi geçen nedensellik kavramı, neden ve sonuç ilişkisine dayanan deneysel desene işaret etmekten ziyade araştırmacılar tarafından oluşturulmuş model bağlamında, değişkenler arasındaki dolaylı ve doğrudan etkilerin test edilmesidir (Şekercioğlu ve ark., 2016).

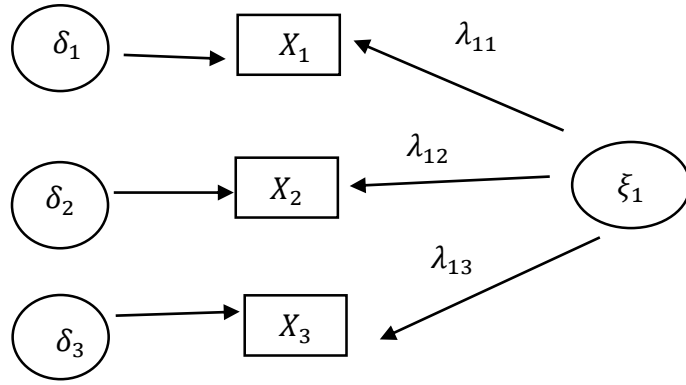
Bu şekliyle YEM, tek bir istatistik analizi olmaktan ziyade içinde birden fazla istatistik tekniğini barındıran bir yöntem olmuştur (Kline, 2011; Çokluk, Şekercioğlu ve Büyüköztürk, 2016). Başta, çoklu grup uygulamaları, doğrulayıcı faktör analizi, yol analizi gibi çeşitli istatistiksel tekniklerden oluşur. YEM daha çok açıklayıcı tekniklerden ziyade doğrulayıcı tekniklerden oluşur. Bu bağlamda araştırmacılar,

YEM de “Bu model geerli mi?” sorusuna cevap ararlar (Kline, 2005; okluk, Őekerciođlu ve Bykztrk, 2016)

2.9.1 Yapısal EŐitlik Modeline İliŐkin Temel Kavramlar

Jreskog (1973)’e gre, YEM iki paradan oluŐmaktadır. İlki, gzlenen deđiŐkenlerle gizil deđiŐkenlerin dođrulayıcı faktr analizi birbirine bađlanarak uygulandıđı “*lme modeli*”, diđeri ise gizil deđiŐkenlerin birbirine eŐ zamanlı eŐitlik sistemleri ile bađlanarak uygulandıđı “*yapısal model*” dir (Akt. Őekerciođlu ve ark.,2016).

lme Modeli: lme modeli, YEM analizlerinin baŐlangıcı kabul edilir (Anderson ve Gerbing, 1988). Ynsz iliŐkilerin ve gizil deđiŐkenlerin hesaplandıđı modeldir. Bu modelde btn parametreler serbest bırakılır (Őekerciođlu ve ark., 2016). lme modelinin yol Őemasıyla gsterimi Őekil 2.’deki gibidir.



Şekil 2. Ölçme Modeli

X = Gözlenen Dışsal Değişken

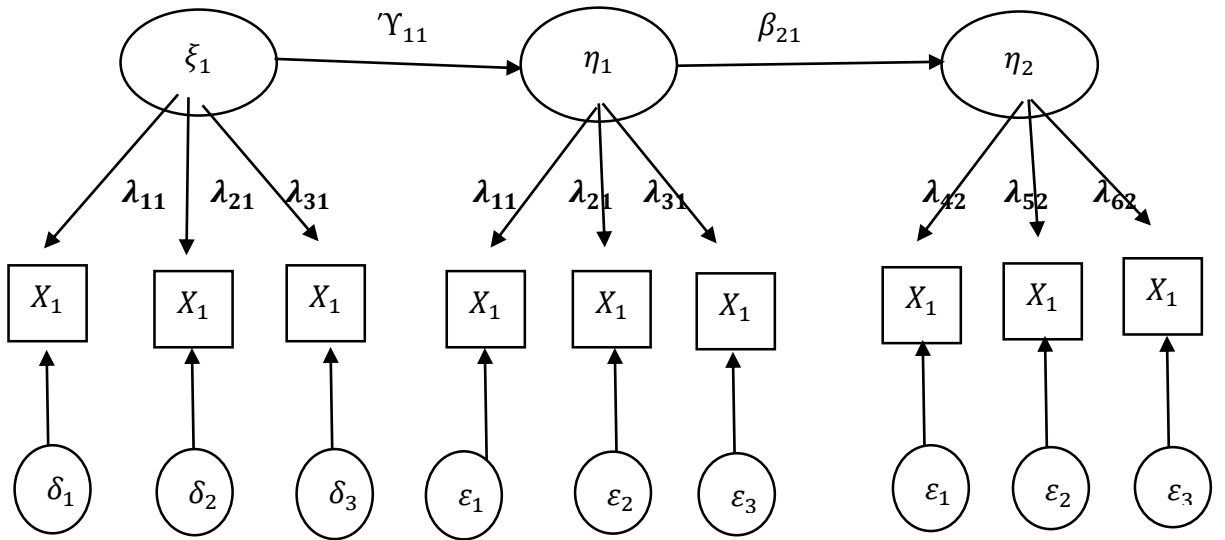
ξ (X_i / K_{si}) = Gizil Dışsal Değişken

λ (Lambda) = Gizil Değişken ve Gözlenen Değişken Arasındaki Bağı İlişkin Yapısal

Katsayı

δ (Delta) = Gözlenen Dışsal Değişkendeki Ölçme Hatası

Yapısal Model: Gözlenen ve örtük değişkenler arasındaki ilişkilerin düzeyini belirten genelleştirilmiş bir modeldir. Yapısal modelin yol şemasıyla gösterimi aşağıda Şekil 3.' deki gibidir.



Şekil 3. Yapısal Model

X = Gözlenen Dışsal Değişken

Y = Gözlenen İçsel Değişken

$\xi(X_i)$ = Gizil Dışsal Değişken

$\eta(\text{Eta})$ = Gizil İçsel Değişken

λ (Lambda) = Gizil Değişken ve Gözlenen Değişken Arasındaki Bağı İlişkin Yapısal Katsayı

δ (Delta) = Gözlenen Dışsal Değişkendeki Ölçme Hatası

ϵ (Epsilon) = Gözlenen İçsel Değişkendeki Ölçme Hatası

ζ (Zeta) = Gizil İçsel Değişkenle İlişkili Hata Terimi

Υ (Gamma) = Dışsal Bir Değişkenden, İçsel Bir Değişkene Olan Yapısal Etki

β (Beta) = İçsel Bir Değişkenin Diğer Bir İçsel Değişkene Olan Yapısal Etkisi

2.9.2. Yapısal Eşitlik Modellerinin Betimlenmesi ve Model Uyumunun Değerlendirilmesi

2.9.2.1. Modelin Betimlenmesi ve Tanımlanması:

Model betimleme, gizil değişken sayısının ve gizil değişkenlerin hangi gözlenen değişkenlere yükleneceğinin belirlenmesi ile ilgilidir. Burada dikkat edilmesi gereken, doğrulanmanın sağlanması için yeterli sayıda ve özgünlükte gizil değişkenin var olup olmadığıdır (Floyd ve Wideman, 1995). Betimleme işlemi modelin şeması çizilerek başlanabileceği gibi, bir dizi denklem ile de gösterilebilir. Bu denklemler daha sonra gözlenen değişken ile gizil değişken arasındaki var olduğu düşünülen ilişkilerle uyumlu model parametreleri şeklinde adlandırılır (Kline, 2005).

Modelin tanımlanabilmesi için, bütün parametrelere özgü tahminleri teorik olarak elde etmek gerekir. Yapısal eşitlik modelinin tanımlanmasında önemli iki ilke bulunmaktadır. Bunlardan ilki, gizil değişkenlerin herhangi bir ölçüğe atanmasıdır. Çünkü bu değişkenlerin sahip olduğu belirli bir ölçek bulunmamaktadır ve gözlenebilir bir yapıları yoktur. Bu işlem ise faktör varyansını belli bir değerde

sabitlerle oluşturulur. Diğer bir ilke ise kovaryans / varyans matrislerinde bulunan veri göstergelerinin sayısı, modelde bulunan parametre sayısından büyük ya da eşit olmalıdır (Khine, 2013).

2.9.2.2. Model Uyumunun Değerlendirilmesi:

Modelin tanımlanmasından sonra, elde edilen verilerle parametreler oluşturulur ve hesaplanır. Hesaplamanın ardından çözümde kullanılan temel teknik “Maksimum Olasılık” (Unweighted Least Square) yöntemidir. Farklılaşan amaçlara göre ise, “Genelleştirilmiş en küçük kareler” yöntemi de kullanılabilir. Bu yöntemlerde bakılan tek uyum ölçütü, elde edilen veriler ile önerilen modelin ne oranda biniştir (Sümer, 2000; akt. Şekercioğlu ve ark. 2016). Model uyumunun değerlendirilmesi için bilgi veren bazı istatistik testleri bulunmaktadır. Bu testler ve açıklamaları aşağıda verilmiştir:

Ki-Kare (X^2) İyilik Uyumu (Chi-Square Goodness of Fit):

Bu test, veriden elde edilen kovaryans matrisleri ile elde edilen modelin birbiriyle uyumunu inceler. Aynı zamanda, model kovaryans matrisi ile gözlenen kovaryans matrisi arasındaki farkın ve örneklem büyüklüğünün bir fonksiyonudur (Bayram, 2010). Ki kare değerinin istatistiksel olarak önemsiz çıkması; başka bir deyişle manidar olmaması, gözlenen ve beklenen matrislerin benzer olduğunu ve mükemmel bir uyumun bulunduğunu göstermektedir. Genellikle beklenen önemsiz bir p değeridir (Uyar, 2011).

Chou ve Bentler’a göre X^2 dağılımı, iki kovaryans arasındaki uyum değerinin, kullanılan örneklemdeki denek sayısı eksi bir ile çarpılmasından elde edilir (Şekercioğlu ve ark., 2016). Ancak Ki-Kare testinde örneklem büyüklüğü önemlidir. Örneklem hacminin 200 den büyük olduğu durumlarda istatistiksel olarak anlamlı sonuçlar elde edilirken, küçük verilerde anlamlılık sağlanamamaktadır (Bayram, 2010). Örneklem büyüklüğünün bu değer üzerindeki etkisine çözüm olarak, serbestlik derecesi kullanılmaktadır. Testteki Ki-kare değerinin serbestlik derecesine bölüldüğünde elde edilen değere bakılır. Eğer bu değer 3’den küçük ise Ki- kare değeri anlamlı bile çıksa model uyumunun yeterli olduğu sonucuna varılabilir (Meydan, Şeşen 2015; akt. Demir, 2016).

İyilik Uyum İndeksi (Goodness of Fit Index, GFI) ve Düzenlenmiş İyilik Uyum İndeksi (Adjusted Goodness of Fit Index, AGFI):

GFI, Ki-Kare'ye alternatif olarak örneklem büyüklüğünden etkilenmeyen bir indeks olarak geliştirilmiştir. Örneklemdeki kovaryans matrisinin model tarafından ne oranda ölçüldüğünü gösterir. Bu indeksin değer aralığı 0-1 dir. Değerin 1 e yakın olması model uyumunu gösterir. *AGFI*, *GFI*'nin düzenlenmiş halidir. *GFI*'ya benzer bir şekilde değer aralığı 0 ile 1 dir. Değerin 1 e yakın olması model uyumunu gösterir. Örneklem büyüklüğüne çok duyarlı bir indekstir. Bu yüzden büyük n'lerde daha uygun değerler verir (Şekercioğlu ve ark.,2016).

Yaklaşık Hataların Ortalama Karekökü (Root Mean Square Error Of Approximation, RMSEA):

RMSEA, merkezi olmayan ki-kare dağılımında popülasyon kovaryanslarını tahmin etmek amacıyla kullanılan bir indeks türüdür (Şekercioğlu ve ark., 2006). RMSEA modeldeki de yanlış tanımlanmış faktör yüklerine ve yanlış tanımlamalara duyarlıdır (Vandenberg ve Lance, 2000). GFI ve AGFI'nin tersine değer 0 a yakın olması uyumun mükemmelliğini gösterir. Örneklem ile evren kovaryansları arasında fark olmadığını ifade eder (Brown,2006; Thompson, 2004, Şekercioğlu ve ark., 2006).

Artık Ortalamaların Karekökü (Root Mean Square Residuals, RMR) Standartlaştırılmış Artık Ortalamaların Karekökü (Standardized Root Mean Square Residuals, SRMR):

Örneklemin kovaryans matrisleri ile evrenin kestirimsel kovaryans matrisi arasındaki artık kovaryans ortalamalarına verilen değerlerdir (Kline, 2005). Bu değerlerin 0' a yakın olması model uyumu için istenen bir durumdur. RMR, aynı verideki iki farklı modelin birbirleriyle karşılaştırılmasını sağlar (Schumacher ve Lomax, 2004 ;akt.Karaduman,2017). SRMR ise kovaryans matrisini ve örneklem tahminini, korelasyon matrisine dönüştürmeye dayanır (Kline, 2005)

Karşılaştırmalı Uyum İndeksi (Comparative Fit Index, CFI):

Karşılaştırma temelli bir indeks türüdür. İç içe modellerde NFI' da bulunan eksik yönleri kapatmak için Bentler (1990) tarafından geliştirilmiştir. Mevcut olan model ile, değişkenler arasındaki kovaryansı ve korelasyonu yok sayan model olan null (sıfır) hipotezi modeli arasında karşılaştırma yapmaktadır. Sonuç olarak, mevcut modelin sıfır hipotezi modeline göre daha iyi olması beklenir. Başka bir deyişle sıfır

hipotezinin χ^2 deęerinin anlamlı olması istenirken, var olan modelin görece anlamlı olmayan bir χ^2 deęeri vermesi beklenir (Sümer, 2000). CFI indeksi için örneklemin büyüklüęü de önemlidir. Bu indeks türü küçük örneklemlerde daha iyi çalıřır. 0 ile 1 arasında yer alan CFI'nın deęeri, 1'e yaklařıkça verinin uyumlu olduęunu gösterir. Deęer 0'a yaklařıkça veri uyumu azaldığı söylenebilir (Hooper, Cauglen ve Mullen, 2008; akt. Karaduman, 2017).

Normlařtırılmıř Uyum İndeksi (Normed Fit Index, NFI) ve Normlařtırılmamıř Uyum İndeksi (Non-normed Fit Index, NNFI):

NFI, serbestlik derecesini ve ki-kare deęerini hesaba katan indeks türüdür. Ki-kare deęerlerinin karřılařtırılmasıyla model tahminlemesi yapılır. 0 ve 1 arasında deęer alır. Deęerin 1'e yaklařması mükemmel uyuma iřaret eder (Tabachnick ve Fidell, 2007). Ancak NFI'nın, karřılařtırmasını yaparken ki-kare daęılımının sayıltılarına uyma zorunluluęu yoktur (řekercioęlu ve ark., 2007). NFI'nın küçük örneklemlere hassas olmaması durumlarında, serbestlik derecesi hesaba katılır. Yeni oluřan deęer NNFI olarak adlandırılır. Sonuç olarak iki indeks de birbirine benzer ancak NNFI'nın verdięi deęer model karmařıklığına da dikkate alır. (Sümer,2000; akt. řekercioęlu ve ark.,2016). Ancak çok küçük örneklemlerde dięer indekslere göre zayıf bir uyum indeksi vermesi, NNFI'nın bir sınırlılıęıdır (Anderson ve Gerbing, 1984).

Basitlik Uyum İndeksi (Parsimony Goodness Of Fit Index, PGFI): GFI indeksinin düzeltilmiř hali PGFI olarak adlandırılır. PGFI, GFI'yı yeniden yorumlar. Modelin yalınlığı hakkında bilgi verir. Deęerin 1 e yaklařması modelin yalınlığını gösterirken 0 a yaklařması karmařık bir model olduęunu gösterir (Marsh, Balla ve McDonald, 1988; Sümer,2000; akt. řekercioęlu ve ark., 2016).

Ancak YEM' de bahsi geçen uyum indekslerinin sınırlılıęları da bulunmaktadır. Bu sınırlılıęlar (Kline, 2005) tarafından belirlenmiřtir.

1. Uyum indekslerinin deęerleri, modelin toplam uyumunu gösterir. Deęerler olumlu çıksa bile modelin bazı bölümleri verilerle kötü uyum gösteriyor olabilir.
2. Bütün indeksler tek başına modeli deęerlendirmek için yeterli deęildir. Birden fazla indeksin yorumlanmasıyla sonuca ulařılabilir. Tüm modeller için en doęru sonucu verecek "sihirli indeks" bulunmamaktadır.

3. YEM’ de kullanılan uyum indekslerinin birçoğunun (RMSEA hariç) örneklem dağılımları bilinmemektedir. Ancak daha sonradan yorumlama kuralları belirlenmiştir.

Tablo1. Model Uyum İndekslerinin Kriterleri ve Kabulü

Uyum İndeksi	Kriterler	Kabul için Kesme Noktaları
X^2 X^2/sd	$p > 0.05$	--- ≤ 2 = mükemmel uyum $\leq 2,5$ =mükemmel uyum (Küçük Örneklerde) ≤ 3 = mükemmel uyum (Büyük Örneklerde) ≤ 5 = orta düzeyde uyum ≥ 0.90 = iyi uyum
GFI	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.95 = mükemmel uyum
AGFI	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.95 = mükemmel uyum
RMSEA	0 (mükemmel uyum) 1 (uyum yok)	≤ 0.05 = mükemmel uyum ≤ 0.07 = iyi uyum ≤ 0.10 = zayıf uyum
RMR / SRMR	0 (mükemmel uyum) 1 (uyum yok)	≤ 0.05 = mükemmel uyum ≤ 0.08 = iyi uyum
CFI	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.90 = iyi uyum
NFI / NNFI	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.90 = iyi uyum
PGFI	0 (uyum yok) 1 (mükemmel uyum)	≥ 0.95 = mükemmel uyum

2.10. İlgili Araştırmalar

Bu bölümde, ülkemizde ve dünyada yapılan kültürlere göre ölçme değişmezliği çalışmalarından bazılarına yer verilmiştir.

2.10.1. Türkiye’de ölçme değişmezliği ile ilgili araştırmalar

Kıbrıslıoğlu (2015), PISA 2012 verilerini kullanarak uygulanan matematik öğrenme alt boyutu anketleriyle oluşturulan matematik öğrenme alt modelinin, Türkiye, Çin-Şangay ve Endonezya örneklerinde ve cinsiyetler bazında ölçme değişmezliğini incelemiştir. Araştırmanın ilk aşaması, AFA ile oluşturulan matematik

öğrenme algısı modelinin 55 madde ve 9 faktör ile doğrulanmasıyla sonlandırılmıştır. İkinci aşamasında ise modelin cinsiyet ve ülke grupları arasında ölçme değişmezliğini sağlayıp sağlamadığına bakılmıştır. Bu aşamada ÇGDFA kullanılmış ve aşamalar halinde ölçme değişmezliğine bakılmıştır. Araştırmanın sonucunda, ülkeler arasında matematik öğrenme algısı modelinin sadece şekil değişmezliğini sağladığı görülmüştür. Bu bağlamda metrik ve ölçek değişmezliği sağlanamadığından gözlenen değişkenlerden alınan puanların karşılaştırılması doğru olmayacaktır. Ancak model, cinsiyet karşılaştırmalarında, ölçme değişmezliğini bütün aşamalarda sağlamıştır. Dolayısıyla değişkenlerin varyansları, kovaryansları ve gözlenen ortalamaları cinsiyet grupları arasında karşılaştırılabilir durumdadır ve örneklem arası farklılıklar, cinsiyet farklılıklarından kaynaklandığı kanısına varılabilir.

Asil ve Gelbal (2012), PISA 2006 verilerini kullanarak, öğrencilerin fen bilimlerine yönelik tutumlarından bilimsel sorgulamaya verilen destek alt boyutunun, Avustralya, Yeni Zelanda, ABD ve Türkiye örneklemi üzerinde diller ve kültürler arası eşdeğerliğini incelemiştir. Öğrenci anketinin faktör yapısı, araştırmanın birinci bölümünde Doğrulayıcı Faktör Analizi (DFA) yoluyla incelenmiştir. İkinci bölümde ise anket maddelerinin farklı diller ve kültürler arasında Değişen Madde Fonksiyonu (DMF) gösterip göstermediği incelemek için Çoklu Grup Doğrulayıcı Faktör Analizi kullanılmıştır. Son bölümde ise maddelerin DMF göstermesinin nedenlerini belirlemek için uzman görüşüne başvurulmuştur. DFA sonuçlarına göre ölçme modelinin, tüm örneklerde aynı faktör yapısına sahip olduğu görülmüştür. Ülkeler arasında değişen fonksiyon gösteren maddelerin olduğu, ÇGDFA bulgularında görülmüştür. Araştırmanın sonunda, ülkeler arasında kültürel ve dilsel farklılıkların artmasıyla, DMF gösteren maddelerin sayısını artırdığı ve maddelerin DMF göstermesinin başlıca sebeplerinden birinin çeviri problemleri ve kültürel farklılık olduğu sonucuna varılmıştır.

PIRLS 2001 başarı testinin ülkeler arası yapı geçerliliği, Akyıldız (2009) tarafından Çok Gruplu Doğrulayıcı Faktör Analizi aracılığıyla incelenmiştir. Araştırmaya 35 ülke dahil edilmiştir. Başarı testinin faktör yapısı dört ayrı ölçek aracılığıyla ayrı ayrı incelenmiştir. Elde edilen verilere göre PIRLS 2001 testinin faktör yapısının, uygulamaya dahil edilen bütün ülkelerde birbirine denk sayılabileceği sonucuna varılmıştır. Buna göre uygulamaya dahil edilen alt testlerin dil ve kültür geçerliğinin sağlandığı görülmektedir. Bu uygulamadan alınan verilere göre ülkelerin birbirleriyle karşılaştırılması anlamlıdır.

Ayvallı (2016), PISA 2012 matematik okuryazarlığının, OECD ülkeleri, Türkiye'deki bölge grupları ve cinsiyet bazında ölçme değişmezliğini incelemiştir. Veri seti olarak OECD ülkelerinden ve Türkiye'den 8 no'lu kitapçığı alan bütün öğrenciler çalışmaya dahil edilmiştir. Ülke, bölge ve cinsiyet grupları için ayrı ayrı DFA uygulayan Ayvallı, modelleri tüm gruplar için doğrulamıştır. Ardından ölçme değişmezliği aşamaları sonucunda Türkiye'deki bölge grupları ve cinsiyet için ölçme değişmezliğinin bütün aşamalarının sağladığını, ancak OECD ülkeleri için ölçme değişmezliğinin sağlamadığı sonucuna ulaşmıştır.

2.10.2. Dünya'da ölçme değişmezliği ile ilgili araştırmalar

Wu, Li ve Zumbo (2007) TIMSS 1999 matematik testinde ölçme değişmezliğini, ülkeler bazında incelemiştir. Araştırma, 7 ülke ile sınırlandırılmıştır. Benzer kültüre sahip olan ABD, Yeni Zelanda, Avustralya birinci grup olarak alınırken Japonya, Tayvan, Kore ise ikinci grup olarak ele alınmıştır. Ülkelerin ikiyeşerli olarak karşılaştırmaları yapılarak ölçme değişmezliği araştırılmıştır. Araştırmada temel amaç, farklı kültürler ve benzer kültürler arasında ölçme değişmezliğinin sağlanıp sağlanmadığını incelemektir. Araştırmanın sonunda benzer kültüre sahip ülkeler arasında ölçme değişmezliğinin sağlandığı görülürken, farklı kültürler arasında değişmezliğin, sadece metrik ve yapısal aşamada sağlandığı sonucuna ulaşılmıştır. Sonuç olarak kültür faktörünün bireylerin test maddelerini algılamaları üzerinde büyük ölçüde etkili olduğu görülmüştür.

Marsh, Hau, Artlet ve arkadaşları (2006), PISA (2000) testinde kullanılan öğrenme yaklaşımları ölçeğini incelemişler ve ülkeler arasında kültürel eşdeğerliğinin bulunup bulunmadığını araştırmışlardır. Araştırmaya 25 ülkenin verileri dahil edilmiştir. Ölçeğin ülkeler arası değişmezliğini belirlemek için çoklu gruplu doğrulayıcı faktör analizi kullanılmıştır. Araştırmanın sonucunda, ölçeğin 25 ülke arasında tam ölçme değişmezliğini sağladığı görülmüştür. Bu bağlamda modelin kültürler arası geçerliliği sağladığı belirlenmiştir.

TIMSS 2007 testinde bulunan fen bilimleri ve matematik motivasyon ölçeklerinin İngiliz kökenli ülkeler ve Arap ülkeleri arasındaki yakınsak, diskriminant ve faktör geçerliğini incelemek üzere, Abduljabbar, Marsh, Ebu-Hilal ve arkadaşları (2013) bir araştırma yapmışlardır. Araştırmaya TIMSS 2007 uygulamasına katılmış ve dili ana dili Arapça olan ülkelere Umman, Suudi Arabistan, Ürdün ve Mısır, ana dili

İngilizce olan ülkelerden ABD, Avusturya, İskoçya, İngiltere dahil edilmiştir. Araştırmanın faktör değişmezliği matematik -fen, cinsiyetler ve ülkeler bazında olmak üzere üç ayrı alanda araştırılmıştır. Araştırmanın sonucunda ulaşılan sonuca göre, cinsiyetler ve matematik-fen alanında ölçme değişmezliği sağlanırken, ülkeler arasında ölçme değişmezliği sağlanamamıştır. Bunun sonucu olarak kısmi ölçme değişmezliği çalışmalarına başvurulmuştur. Sonuç olarak TIMSS 2007 ölçeklerin matematik-fen alanları arasında ve cinsiyetler arasında geçerliği sağladığı görülürken, ölçeklerin ülkeler arasında geçerli olmadığı sonucuna ulaşılmıştır.

1995 TIMSS uygulamasının Fransızca ve İngilizce formlarının değişmezliği Ercikan ve Koh (2005) tarafından araştırılmıştır. Araştırmanın asıl amacı dilin etkisini incelemek olduğu için Kanada'dan öğrenciler uygulamaya dahil edilmiştir. Araştırmada DFA ve MTK birlikte kullanılmıştır. ÇGDFA ile uygulamanın Fransızca ve İngilizce formlarının ölçme değişmezliğini sağlayıp sağlamadığına bakılmıştır. Aynı zamanda maddelerin gruplar arasında DMF gösterip göstermediği incelenmiştir. Araştırmanın sonucunda fen bilimleri ve matematik alt gruplarının maddelerinin birçoğunun gruplar arası fonksiyonlarında önemli farklılaşmalarının olduğunu ve gruplar arasında ölçeklerin ölçme değişmezliğini sağlamadığı görülmüştür. Bu bağlamda farklı dillerde yapılan testlerin birbirleriyle denk olmadığı ve ülke sıralamalarında elde edilen sonuçların doğruluğunun sınırlandığı görülmüştür.

BÖLÜM III

YÖNTEM

Araştırmanın bu bölümünde araştırmanın türü, çalışma grubu, veri toplama aracı, verilerin analizi ve verilerin toplanmasından söz edilecektir.

3.1. Araştırmanın Türü

Bu araştırmada, PISA 2015 Matematik testinin faktör yapısını inceleyip, farklı gruplarda aynı yapıyı ölçüp ölçmediğine yönelik bir model ortaya koymaya çalışılmış, daha sonra modelin gruplar arası eşdeğerliği incelenmiştir. Bu çalışma, PISA 2015 testinin geçerliliğini saptamaya yönelik olduğundan bir ilişkisel tarama modelidir. Kıncal (2013)' e göre ilişkisel tarama modelleri, birden fazla değişken arasındaki ilişkinin saptanmasını ve derecesinin belirlenmesini, daha sonrasında bu ilişkilerle ilgili daha ayrıntılı çalışmalar yapılabilmesini sağlayan araştırma modelleridir. Neden sonuç ilişkisi aranmayan bu modeller, özellikle eğitim alanında birçok farklı şekillerde uygulanabilmektedir.

3.2. Çalışma Grubu

PISA 2015 uygulamasına 72 ülkeden 15 yaş grubu yaklaşık 540 bin öğrenci katılmıştır. PISA uygulamasında okul örnekleme, tabakalı seçkisiz örnekleme yöntemiyle belirlenmektedir. Balcı (2001)' e göre bu yöntem, en iyi temsillere ulaşabilmek için evreni homojen gruplara ayırmak olarak tanımlanmıştır. Bu araştırma için iki aşama uygulanmaktadır. İlk aşamada İstatistikî Bölge Birimleri Sınıflaması (İBBS), okul türü, eğitim türü, okulların idari biçimleri, okulların buldukları konum ve tabakaları kullanılarak tabakalı seçkisiz örnekleme yöntemiyle okullar belirlenmiştir, ikinci aşamada ise seçilen okullarda uygulamaya katılacak olan bireyler seçkisiz yöntemle belirlenmiştir (PISA, 2015). Bu çalışmaya PISA 2015 uygulamasına katılan 72 ülke de konuşulan 82 farklı dil ve 540 bin öğrenci evrenini temsilen 12 farklı dil ve 2816 öğrenci örnekleme dahil edilmiştir. Araştırmaya dahil edilen ülkeler random biçimde seçilmiştir. Çalışmaya Türkiye' den 165, Çin'den 212, Almanya' dan 178, İtalya' dan 310, İsrail'den 188, Katar'dan 467, Litvanya'dan 177, Rusya'dan 168,

Singapur'dan 162, Birleşik Krallık'dan 383, Tayland'dan 230 ve Slovakya'dan 176 öğrenci dahil edilmiştir.

PISA 2015 Matematik okuryazarlığı testinde kullanılan 96 kitapçık mevcuttur. Bu kitapçıklarda toplam 81 ayrı soru maddesi bulunmaktadır. Bu araştırmada ölçme değişmezliği analizleri yapılabilmesi için bütün ülkelerde ortak olarak kullanılmış 54 numaralı kitapçık seçilmiştir. Bu kitapçıkta ise matematik okuryazarlığı testine ait 25 madde bulunmaktadır. Verilere PISA'nın internet sitesinden ulaşılmıştır. Ancak kitapçıkta 22 madde analize dahil edilmiştir. 54 numaralı kitapçığı alan bireylerin ülkelere göre örneklem dağılımı Tablo 2'de verilmiştir.

Tablo 2. Ülkelere Göre Örneklem Dağılımı

Ülke	N	%
Çin	212	7.52
Almanya	178	6.32
Birleşik Krallık	383	13.6
Türkiye	165	5.85
İtalya	310	11
İsrail	188	6.67
Katar	467	16.58
Litvanya	177	6.28
Rusya	168	5.96
Singapur	162	5.75
Tayland	230	8.16
Slovakya	176	6.25

Tablo 2'ye göre araştırmaya katılan bireylerin dağılımı incelendiğinde, en yüksek katılımcıya sahip ülke %16.58 (467 kişi) ile Katar, en düşük katılımcıya sahip ülkenin ise %5.75 (162 kişi) ile Singapur olduğu görülmektedir. Araştırmaya dahil edilen diğer ülkelerin katılımcı yüzdeleri ise %13.6 ile %5.85 arasında değişmektedir.

3.3. Veri Toplama Araçları

Araştırmada PISA 2015 değerlendirilmesinde Türkiye, Çin, Almanya, İtalya, İsrail, Katar, Litvanya, Rusya, Singapur, Birleşik Krallık, Tayland, Slovakya örneklemelerinin 15 yaş grubu öğrencilerine uygulanan Matematik okuryazarlığı testinde yer alan maddelere verilen cevaplar kullanılmıştır. Madde formatlarının bir kısmı açık uçlu sorulardan, bir kısmı ise çoktan seçmeli sorulardan oluşmuştur. Araştırma için 54 numaralı kitapçık seçilmiş ancak sadece puan değerleri 1-0 olan

çoktan seçmeli sorulardan oluşan maddeler analize dahil edilmiştir. Bu bağlamda araştırmanın analizine 22 madde ile devam edilmiştir. Veriler 2016 yılında yayınlanmış olup, OECD'nin web sitesinden alınmıştır. Bu maddelerin içerikleri aşağıdaki Tablo 3'de ayrıntılı bir şekilde verilmiştir.

Tablo 3. Analizde Kullanılan Matematik Okuryazarlığı Maddeleri

Madde Kimliği	Ölçülen süreç	İçerik
CM305Q01S	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Uzay ve şekil
CM496Q01S	Durumları matematiksel olarak formüle etme	Sayı
CM496Q02S	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Sayı
CM423Q01S	Matematiksel sonuçları yorumlama, uygulama ve değerlendirme	Belirsizlik ve veriler
DM406Q01C	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Uzay ve şekil
DM406Q02C	Durumları matematiksel olarak formüle etme	Uzay ve şekil
CM603Q01S	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Sayı
CM571Q01S	Matematiksel sonuçları yorumlama, uygulama ve değerlendirme	Değişim ve ilişkiler
CM564Q01S	Durumları matematiksel olarak formüle etme	Sayı
CM564Q02S	Durumları matematiksel olarak formüle etme	Belirsizlik ve veriler
CM905Q01S	Matematiksel sonuçları yorumlama, uygulama ve değerlendirme	Sayı
DM905Q02C	Matematiksel sonuçları yorumlama, uygulama ve değerlendirme	Sayı
CM919Q01S	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Sayı
CM919Q02S	Durumları matematiksel olarak formüle etme	Sayı
CM954Q01S	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Değişim ve ilişkiler
DM954Q02C	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Değişim ve ilişkiler
CM954Q04S	Matematiksel kavramların, olayların ve prosedürlerin kullanılması	Değişim ve ilişkiler
CM943Q01S	Durumları matematiksel olarak formüle etme	Değişim ve ilişkiler
CM943Q02S	Durumları matematiksel olarak formüle etme	Uzay ve şekil
DM953Q02C	Matematiksel sonuçları yorumlama, uygulama ve değerlendirme	Belirsizlik ve veriler
CM953Q03S	Durumları matematiksel olarak formüle etme	Belirsizlik ve veriler
DM953Q04C	Durumları matematiksel olarak formüle etme	Belirsizlik ve veriler

3.4.Verilerin Analizi

Bu çalışmanın analizinde, doğrulayıcı faktör analizi yapmak, verilerin dağılımları belirlemek, model oluşturmak ve modelin gruplar arasında değişmezliğini test etmek için istatistik programlarından yararlanılmıştır. Araştırmaya; “Çin, Almanya, Birleşik Krallık, İsrail, İtalya, Litvanya, Katar, Rusya, Singapur, Slovakya, Tayland, Türkiye” ülkeleri dahil edilmiştir. Her ülke için ortak bir kitapçık olan 54 numaralı kitapçık seçilmiştir. Bu kitapçıkta 25 soruya ulaşılmıştır. Veriler analiz edilmeden önce, doğru sonuçlara ulaşabilmek için “Kayıp değer” ve “Uç Değer” incelemesi yapılmıştır. Kayıp değer; veri setinde ulaşılamayan veya kaybolmuş değerlerdir (Tabacnick ve Fidell, 2007). Literatürde az sayıda kayıp veriler için önerilegelen, değişkenlerin veya deneklerin silinmesidir (Büyüköztürk, 2016). Bu

çalışmada 3 madde kayıp değer olarak tespit edilmiştir. Yapılan analizlere dayanarak kayıp verilerin analize alınmamasına karar verilmiştir. Uç değer ise; örneklemin alındığı evrenin üyesi olmayan bireylerin tespiti olarak tanımlanır (Tabacnick ve Fidell, 2007). 12 ülke için yapılan araştırmada verilerinin tek değişkenli ve çok değişkenli uç değerlerine bakılmıştır. Bireylerin puanları, standart puana (Z puanına) çevrilmiş ve Z puanı +3, -3 aralığında olmayan bireylere rastlanmamıştır. Verilerin temizlenmesinin ardından ortak sorular belirlenmiştir. Bunlar içerikleriyle birlikte Tablo 2’ de verilmiştir. Modelin güvenilirliğini değerlendirebilmek için 1-0 lı verilerde kullanılan KR-20 güvenilirlik katsayısı değerleri incelenmiştir. Analize dahil edilen bütün ülkeler için KR-20 değeri .70 in üzerinde olduğu görülmüştür. Bu aralık literatürde test puanlarının güvenilir olduğunu söylemek için genellikle yeterli olarak görülmüştür (Büyüköztürk, 2016). Bu işlemlerin ardından veri dağılımının normalliğinin değerlendirilebilmesi için çarpıklık, basıklık ve anlamlılık değerlerine bakılmıştır. Olması gereken normal dağılımda, çarpıklık ve basıklık katsayıları sıfırdır. Ancak bu değerlerin +1, -1 aralığında olması normallik için yeterlidir. Bir diğer normallik testi ise Shapiro Wilks p anlamlılık değeridir. Eğer test sonucu $p < .05$ ise bu evren ile örneklemin varyanslarının eşit olduğunu ifade eder. Başka bir deyişle dağılım normalden aşırı bir sapma göstermez (Şekercioğlu, 2016). 12 ülkenin normallik testi sonuçları Tablo 3’de verilmiştir. İstatistikler hiçbir ülkenin dağılımının normalden aşırı bir sapma göstermediğini ortaya koymuştur.

Kayıp değer, uç değer ve normalliğin tespitinin ardından YEM çalışmalarına geçilmiştir. Bir istatistik programı yardımıyla 12 ülke için ayrı ayrı doğrulayıcı faktör analizi yapılmıştır. Bahsi geçen bütün ülkelerin modelleri için yapılan analiz sonucunda manidar olmayan t değerine rastlanmamıştır. Bütün modellerin RMSEA değerleri 0.08 in altında çıkmıştır. Bu değer literatürde iyi uyum denebilecek düzeyde kabul edilir. Değerlendirme için kontrol edilen diğer uyum indeksi X^2 istatistiğidir. Ancak bu istatistik serbestlik derecesi (sd) ile oranlanarak işleme alınır. Bu iki değer oranlandığında sonuçlar 3 ün altında ise mükemmel uyuma, 5 in altında ise orta düzey uyuma işaret eder (Büyüköztürk, 2016). Bu doğrultuda yapılan analiz sonucu bütün ülkelerinin X^2/sd değerlerin 3’ün altında olduğu görülmüştür. Böylece ölçme modeli tüm gruplar için doğrulanmıştır. İstatistikler Tablo 4’ de verilmiştir. Ayrıca analizlerin path diagramları Ek 1’ de verilmiştir.

Çoklu gruplarda yapılan analizlerde ölçme değişmezliğinin test edilebilmesi için beş aşamanın incelendiği bir analiz yapılmalıdır (Vandenberg ve Lance, 2000). Bu

doğrultuda bütün gruplar için aşamaların her biri sırayla test edilmiştir. Yöntem olarak ise “Genelleştirilmiş Küçük Kareler (Unweighted Least Square)” yöntemi uygulanmıştır. Model değişmezliğinin kabulü için ilk olarak uyum indekslerinin istenilen aralıkta olup olmadığı incelenmiştir. Model uyumu değerlendirilirken RMSEA, CFI, NFI ve SRMR değerleri dikkate alınmıştır. Kline (2005)’e göre, ölçme modellerini karşılaştırırken uyum indeksleri arasındaki fark değerlendirilir. Bahsi geçen indeksler arasındaki fark incelenirken öncelikle X^2 testine (Δx^2) ve serbestlik derecelerine (Δsd) bakılır . Ayrıca ΔCFI , ΔNFI , ΔGFI , $\Delta SRMR$ $\Delta RMSEA$, ΔTLI değerlerinin de incelenmesi önerilmiştir (Cheng ve Rensvold, 2002; akt. Ayvally, 2016). Bu araştırmada $\Delta SRMR$ ve ΔCFI değerleri temel alınmıştır. Karşılaştırma sonucunun, $n > 300$ verileri için metrik değişmezliğinde $\Delta SRMR \geq .015$ ve $.01 \geq \Delta CFI \geq -.01$ olması modelin kabul edilebilirliğini gösterir. Ölçek değişmezlik ve katı değişmezlik karşılaştırılmalarında $n > 300$ verileri için kabul aralığı $\Delta SRMR \geq .01$ ve $\Delta CFI \geq -.01$ dir (Chen, 2007).

Bir hipotez incelenip yeterli sonuçlar alınınca diğer aşamaya geçilmiştir. İlk olarak bütün gruplar için Ölçme değişmezliğinin ilk aşaması olan “Yapısal / Şekilsel Değişmezlik” hipotezine bakılmıştır. Bu aşamada uyum istatistikleri kabul edilebilir sonuçlar verdiği için ($X^2/sd < 5$, $RMSEA < 0.1$ ve CFI , NFI , $NNFI > 0.90$) modelin yapısal / şekilsel değişmezlik aşamasını sağladığı sonucuna ulaşılmıştır. Böylelikle metrik değişmezlik aşamasının ilk ön koşulu sağlanmıştır. Bir sonraki aşama olan metrik değişmezlik aşamasında ise model uyum istatistiklerinin (X^2/sd , $RMSEA$, CFI , NFI , $NNFI$) istenilen düzeyde olduğu ve $\Delta SRMR \geq .01$ ile $\Delta CFI \geq -.01$ değerlerinin de kabul edilebilir aralıkta olduğu görülmüştür. Bu bulgular ışığında modelin, yapısal değişmezliğin ardından metrik değişmezlik aşamasını da sağladığı görülmüştür. Ardından ölçek değişmezliği aşamasına geçilmiştir. Bu aşamada ΔCFI sonuçlarının kabul edilen aralıkta olmamasıyla modelin ölçme değişmezliğinin ölçek değişmezlik aşamasını sağlamadığı yargısına varılmıştır. Ölçme değişmezliği aşamalarında herhangi bir analiz yeterli sonuçlar vermediğinde diğer analiz de yetersiz sonuç vereceğinden analiz o aşamada sonlandırılır. Katı değişmezlik aşaması için değerler kontrol edildiğinde zaten değerlerin kabul edilebilir aralıkta olmadığı görülmüştür. Böylelikle analiz sonlandırılmıştır.

BÖLÜM IV

BULGULAR

Araştırmanın bu bölümünde PISA 2015 matematik okuryazarlığı testinin, 12 ülke için test istatistiklerine, güvenilirlik katsayılarına ve ölçme değişmezliği sonuçlarına, yapısal değişmezlik, metrik değişmezlik, ölçek değişmezliği ve katı değişmezliği bulgularına yer verilmiştir.

4.1. Araştırmaya dahil olan Ülkeler İçin Ölçme Değişmezliğine İlişkin Bulgular

Bu başlık altında farklı dilleri konuşan 12 ülke için ölçme değişmezliğinin sağlanıp sağlanmadığı değerlendirilmiştir. İlk olarak kayıp değer olarak tespit edilen CM967Q03S CM949Q02S CM967Q01S maddeleri veri setinden çıkarılmıştır. Ardından araştırmaya dahil edilen ülkeler için Cronbach Alfa katsayıları, basıklık (kurtosis) ve çarpıklık (skwness) hesaplanarak Tablo 3' de verilmiştir.

Tablo 4. Analize Dahil Edilen Ülkeler İçin Normallik Testi Sonuçları

	Çarpıklık	Basıklık	Test of Normality	KR-20
CHNATAP	-.229	-.911	.000	.863
DEU	-.049	-.716	.024	.827
GBR	.154	-.559	.002	.808
ISR	.493	-.563	.000	.843
ITA	.169	-.689	.000	.823
LTU	.605	.065	.000	.814
QAT	.980	.977	.000	.820
RUS	.240	-.424	.036	.786
SGP	-.264	-.589	.015	.843
SVK	.126	-.579	.030	.810
THA	.925	.591	.000	.774
TUR	.492	-.380	.000	.763

Tablo 3' deki sonuçlar incelendiğinde çarpıklık ve basıklık değerlerinin ± 1 aralığında yer alması veri setinin normal dağılıma yakın bir dağılım sergilediği anlamına gelmektedir (Büyüköztürk, 2016). Güvenirlilik tespiti için incelenen KR-20 katsayılarında ise bütün ülkeler için değerlerin .70 den büyük olduğu görülmektedir. Bu aralık test puanlarının güvenirliliği için yeterli kabul edilmiştir (Büyüköztürk, 2016).

Test istatistiklerinin hesaplanmasının ardından, modelin arařtırmaya dahil edilen bütün gruplarda dođrulanıp dođrulanmadığını tespit etmek için, dođrulayıcı faktör analizi uygulanmıştır. Sonuçlar Tablo 4’ de verilmiştir.

Tablo 5. Analize Dahil Edilen Ülkeler İçin DFA Testi Sonuçları

	X^2	p-value	RMSEA	X^2/sd
CHNATAP	284.20	0.000	0.010	1.34
DEU	218.64	0.009	0.016	1.23
GBR	417.28	0.000	0.051	1.09
ISR	380.42	0.000	0.066	2.03
ITA	387.68	0.000	0.053	1.25
LTU	273.74	0.001	0.042	1.55
QAT	573.84	0.000	0.061	1.23
RUS	255.29	0.015	0.036	1.52
SGP	258.80	0.010	0.038	1.60
SVK	333.43	0.000	0.058	1.90
THA	411.49	0.000	0.065	1.79
TUR	505.01	0.000	0.053	1.85

Tablo 4’ deki uyum indeksleri incelendiğinde bütün gruplar için X^2/sd oranının 2 nin altında olduğu ve RMSEA değerlerinin .08 den küçük olduğu görülmektedir. Bu değerler literatürde iyi uyuma işaret etmektedir (Büyüköztürk, 2016). Bu durumda, modelin arařtırmaya dahil edilen bütün ülkeler için dođrulandığı söylenebilir.

Son olarak ise arařtırmada ölçme deđişmezliđinin dört aşaması sırasıyla test edilmiştir (Wu ve ark., 2007). Bütün aşamalar için hata varyanslarının standartlaştırılmış değerleri sırasıyla Tablo 6’ da verilmiştir. Ayrıca madde faktör yükleri ve parametrelerin standartlaşmamış durumları Ek 2’deki yol diyagramında gösterilmiştir.

Tablo 6. Ölçme Değişmezliği Aşamalarına İlişkin Uyum İndeksileri

	<i>M.K*</i>	<i>SB_{x²}</i> (<i>sd</i>)	<i>X²/sd</i>	ΔX^2	$\Delta X^2/\Delta sd$	<i>RMSEA</i>	<i>SRMR</i>	<i>CFI</i>	<i>NNFI</i>	$\Delta SRMR$	ΔCFI
<i>Model 1^A</i>	-	8816.51 (2750)	3.20	-	-	0.097	0.13	0.96	0.96	-	-
<i>Model 2^B</i>	M1-M2	8089.51 (2750)	2.94	727.0	0.26	0.091	0.11	0.96	0.96	0.02	0
<i>Model 3^C</i>	M2-M3	5347.01 (2508)	2.13	2693,4	0.81	0.16	0.13	0.98	0.98	0.02	0.02
<i>Model 4^C</i>	M3-M4	8816.64 (2750)	3.20	3469.63	1.07	0.097	0.13	0.96	0.96	0	0.02

A *Yapısal Değişmezlik (Faktör korelasyonları, faktör yükleri, ve hata varyansları sabit)*

B *Metrik Değişmezlik (Faktör yükleri serbest, hata varyansları ve faktör korelasyonları sabit)*

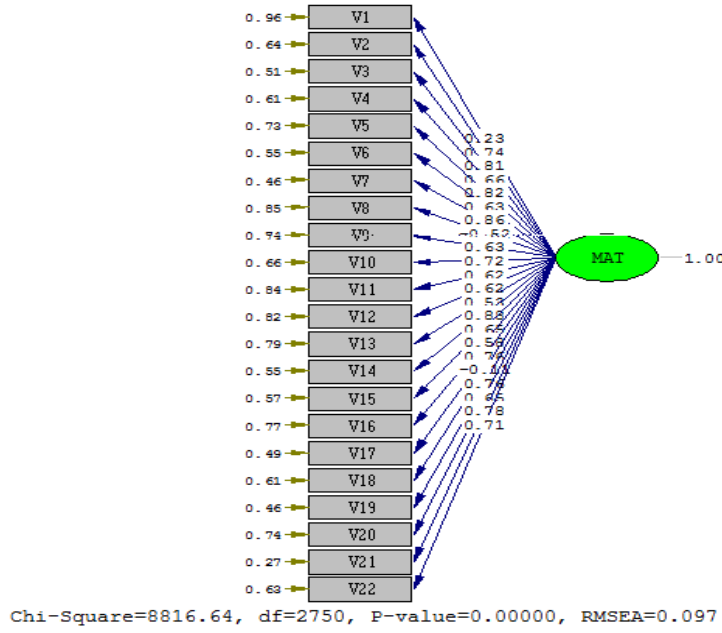
C *Ölçek Değişmezliği (Hata varyansları ve faktör yükleri serbest, faktör korelasyonları sabit)*

D *Katı Değişmezlik (Hata varyansları serbest, faktör korelasyonları ve faktör yükleri sabit)*

4.1.1. Şekilsel Değişmezlik

Bu başlık altında araştırmanın alt problemlerinden ilki olan PISA 2015 matematik okuryazarlığı testinin farklı dilleri konuşan ülkeler arasında şekil değişmezliğinin sağlanıp sağlanmadığı sorusuna yanıt aranmıştır. Bu bağlamda, modeldeki madde gruplarının ve faktör yapısının araştırmaya dahil edilen ülkeler arasında değişmezliği test edilmiştir. Analiz sonuçları Tablo 4' de verilmiştir.

Tablo 6' da bulunan uyum indekslerine göre ölçme değişmezliğinin şekil değişmezliği aşamasında X^2/sd istatistiği incelendiğinde sonucun iyi uyum düzeyinde olmasa da kabul edilebilir aralıkta olduğu görülmektedir ($2 < X^2/sd \leq 8$). Diğer uyum indeksleri incelendiğinde ise model uyumunun kabul edilebilir değer aralığında olduğu görülmektedir ($RMSEA < 0.1$, CFI , NFI , $NNFI > 0.90$, $.05 < SRMR \leq 1$). Bu durumda modelin 12 ülke arasında şekil değişmezliğini sağladığı görülmektedir. Analizin path diagramı sonuçları Şekil 4 te verilmiştir. Yani ölçme modelinin farklı dili konuşan her grupta aynı olduğu söylenebilir. Bu aşamanın sağlanması bir sonraki aşamanın ön koşulu olduğu için metrik değişmezliği analizine geçilmiştir.

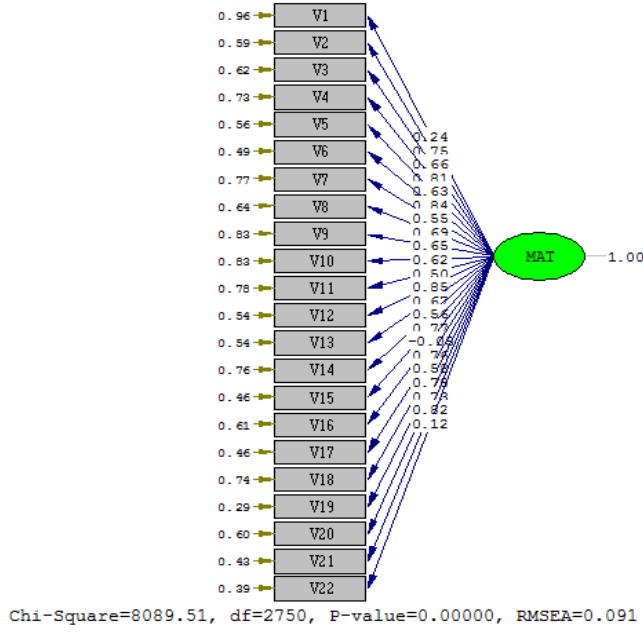


Şekil 4. Şekil Değişmezliği Path Diagramı Sonuçları

4.1.2. Metrik Değişmezlik

Bu başlık altında araştırmamızın ikinci alt problemi olan PISA 2015 matematik okuryazarlığı testinin farklı dilleri konuşan ülkeler arasında metrik değişmezliğinin sağlanıp sağlanmadığı sorusuna yanıt aranmıştır. Bu aşamada ülkeler arasında faktör yapısının ve madde gruplarının eşliğinin yanı sıra faktör yükleri de eşitlenmiş ve metrik değişmezliğin sağlanıp sağlanmadığı kontrol edilmiştir. Analiz sonuçları Tablo 5’ de verilmiştir.

Tablo 5’ deki uyum istatistiklerine göre metrik değişmezlik aşamasında modelin, veriyle uyumlu olduğu görülmektedir. ($2 < X^2/sd \leq 8$, $RMSEA < 0.1$, CFI, NFI, NNFI > 0.90 , $.05 < SRMR \leq 1$). Ancak metrik değişmezliğinin sağlanabilmesi için, yapısal değişmezlik aşamasında elde edilen uyum indeksleriyle, metrik değişmezlik aşamasında elde edilen uyum indekslerinin arasındaki farkların kabul edilebilir aralıkta ($-0.01 \leq \Delta CFI \leq 0.01$, $\Delta SRMR \geq 0.01$) olması gerekmektedir. Bu farklar incelendiğinde ise sonuçların istenilen düzeyde olduğu görülmektedir. Analiz sonuçlarının path diagramı Şekil 5. te verilmiştir. Bu sonuçlardan yola çıkarak, modeldeki değişkenlerin faktör yükleri, alt gruplarda sabit kaldığı yargısına varılabilir. Yani farklı dilleri konuşan bölgelerde ölçeğin boyutları ile ölçülmek istenen özellikler arasındaki ilişkilerin benzer olduğu sonucuna ulaşılabilir. Bu aşamanın sağlanmasından sonra ölçek değişmezliği aşamasına geçilmiştir.



Şekil 5. Metrik Değişmezliği Path Diagramı Sonuçları

4.1.3. Ölçek Değişmezliği

Bu başlık altında araştırmanın üçüncü alt problemi olan PISA 2015 matematik okuryazarlığı testinin farklı dilleri konuşan ülkeler arasında ölçek değişmezliğinin sağlanıp sağlanmadığı sorusuna yanıt aranmıştır. Bu bağlamda; regresyon denklemlerinde bulunan sabit değerlerin bütün gruplar için eş değeri test edilmiştir (Önen, 2007). Tablo 6’ daki uyum istatistiklerine göre ölçek değişmezliği aşamasındaki uyum indekslerinden X^2/sd , CFI, NFI, NNFI, SRMR ve $\Delta SRMR$ değerlerinin kabul edilebilir sınırlar içerisinde olduğu görülmektedir ($2 < X^2/sd \leq 8$, CFI, NFI, NNFI > 0.90 , $.05 < SRMR \leq .1$, $\Delta SRMR > 0.01$). Ancak RMSEA değerinin 0.16 ile kabul edilen aralığın dışına çıktığı (RMSEA < 0.1) ve metrik değişmezlik aşamasında elde edilen CFI değeri ile ölçek değişmezliği aşamasında elde edilen CFI değerleri arasındaki farkın kabul edilen aralıkta ($-0.01 \leq \Delta CFI \leq 0.01$) olmadığı görülmektedir. Bu bağlamda, modelin ölçek değişmezliği aşamasını sağlamadığı ifade edilebilir. Elde edilen bulgulara dayanılarak, maddelerin faktör yükleri, farklı dilleri konuşan ülkeler arasında değişmez iken, faktörler arası korelasyonlar için aynı durumun söz konusu olmadığı söylenebilir. Başka bir deyişle, örtük yapıların ortalamalarının farklı gruplarda aynı sonuçların vermediği görülmektedir. Bu durumda farklı dilleri konuşan bireylerin ortalamalarının karşılaştırılmasının anlamlı olmayacağı yargısına ulaşılabilir.

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

Bu başlık altında araştırmanın bulgularına dayanarak sonuçlar ve öneriler raporlanmıştır.

5.1. Sonuç ve Tartışma

“Uluslararası Öğrenci Değerlendirme Programı” olarak adlandırılan PISA projesinin 2015 uygulamasında Matematik okur-yazarlığı testi ile 22 maddelik bir model elde edilmiş ve modelin farklı dilleri konuşan ülkeler arasındaki ölçme değişmezliği incelenmiştir. Elde edilen bulgulara dayanılarak ulaşılan sonuçlar şu şekildedir:

1. PISA 2015 matematik okuryazarlığı testinin ölçme değişmezliğinin tespiti için öncelikle normallik tespiti yapılmıştır. Hesaplamalar sonucunda oluşturulan modelin normale yakın bir dağılım gösterdiği sonucuna varılmıştır. Ardından 22 maddeden oluşan model doğrulayıcı faktör analizi ile doğrulanmıştır.
2. Farklı dillerin konuşulmasına dayalı olarak yapılan ölçme değişmezliği testi sonuçları, yapısal ve metrik değişmezlik koşullarını sağlarken modelin ölçek değişmezliği aşamasının koşullarını sağlamadığı görülmüştür. Bu sonuçlara göre, modelin farklı dilleri konuşan gruplar arasında yapısı değişmezdir ve modeldeki değişkenlerin faktör yükleri alt gruplarda eşittir. Ancak faktörler arası korelasyonlar gruplar arasında sabit değildir.

Bu sonuçlara dayanarak, PISA 2015 matematik okuryazarlığı testinin değişkenlerinden bir kısmının gruplardan birine ya da birden fazlasına yanlı davrandığı ve farklı dilleri konuşan bireyler için bazı değişkenlerin aynı anlamlara gelmediği sonucuna ulaşılabılır. Benzer sonuçlara farklı örneklemeler ve farklı çalışma alanlarıyla diğer araştırmacılar tarafından ulaşılmıştır. Örneğin; Kıbrıslıoğlu (2015)'in PISA 2012 matematik öğrenme modelinin kültürlere ve cinsiyete göre ölçme değişmezliği çalışmasında analiz sadece şekil değişmezliğini sağlamıştır. TIMSS 2011 verilerini inceleyen Ölçüoğlu (2015) ise matematik başarısını etkileyen değişkenleri

bölgelere göre incelediğinde analiz sonuçlarının ölçek değişmezliği aşamasını sağlamadığı görülmüştür. Benzer şekilde Uyar (2011) tarafından PISA 2009 sonuçları farklı gruplarda incelendiğinde ise analiz sonuçlarının ölçek değişmezliği aşamasını sağlamadığı görülmüştür. Bu bağlamda yapılan araştırma önceki çalışmalar ile paralellik göstermektedir. Dolayısıyla farklı dilleri konuşan veya farklı gruplarda bulunan bireyler arasında gözlenen değişkenlerin ortalamalarının farklılıkları, ölçeğin özelliklerinden meydana gelmiş olabilir. Bu nedenle aynı dili konuşmayan grupların aynı ölçeklerden veya aynı testlerden aldıkları puanlar tartışmaya açık hale gelmektedir. Büyük çaplı ölçme ve değerlendirme uygulamalarında elde edilen sonuçlara göre yorumlama yapılmadan önce bahsi geçen ölçeklerin değişmezlik ve yanlılık çalışmalarının yapılması ve sonuçların bu bulgulara dayanarak yorumlanması gerekmektedir.

5.2. Öneriler

Araştırma sonuçlarından çıkan uygulamaya ve araştırmaya yönelik olarak öneriler şu şekilde sıralanabilir:

1. Özellikle geniş ölçekli ölçme ve değerlendirme çalışmalarında ölçeğin geliştirilmesi ve uyarlanması aşamasında, uygulanacağı grupların özellikleri dikkate alınmalıdır. Yapılan karşılaştırmaların sonuçları yorumlanmadan önce ölçme aracının gruplar arası geçerliliğine ilişkin ölçme değişmezliği test edilmelidir.
2. Bu araştırma yalnızca PISA 2015 matematik okuryazarlığı testinin 54. kitapçığıyla gerçekleştirilmiş bir çalışmadır. Sonraki araştırmalarda farklı kitapçıklar kullanılabilir veya farklı çalışma alanlarıyla (fen okuryazarlığı, okuma becerileri) sonuçların genellenebilmesine yardımcı olunabilir.
3. Araştırma için kullanılan model ölçme değişmezliğinin yapısal ve metrik değişmezliği aşamalarını sağlamıştır. Daha sonraki araştırmalarda grupların kısmi değişmezlik aşaması test edilebilir.
4. Araştırmaya farklı dili konuşan 12 ülke dahil edilmiştir. Daha sonraki araştırmalarda farklı dilleri konuşan başka ülkeler veya daha farklı demografik özellikler karşılaştırılabilir.

5. Arařtırmada kullanılan modelin ölçme deęişmezlięi saęlamadıęı sonucuna varılmıřtır. Yani, ölçek maddelerinin bir kısmının bazı gruplara yanlı davrandıęı düşünölmektedir. Daha sonraki arařtırmalarda yanlılık gösteren maddelerin neler olduęunun belirlenmesi için yanlılık çalıřmaları yapılabilir.
6. Arařtırmada ölçme deęişmezlięinin tespiti için çoklu gruplu doęrulatoryıcı faktör analizi kullanılmıřtır. Yöntem olarak ise, genelleřtirilmiř küçük kareler (unweighted least square) yöntemi kullanılmıřtır. Daha sonraki arařtırmalarda maksimum likelihood yöntemi kullanılıp yöntemler arası karşılařtırma imkanı oluşturulabilir.

1.9. Kaynakça

- AERA, A. P. A. (2005). NCME (1999). *Standards for educational and psychological testing*, 146.
- Akyıldız, M. (2009). Pirls 2001 testinin yapı geçerliliğinin ülkelerarası karşılaştırılması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(1).
- Alatlı, B. K., & Bökeoğlu, Ö. Ç. (2018). Uluslararası Öğrenci Değerlendirme Programı (PISA-2012) Okuryazarlık Testlerinin Ölçme Değişmezliğinin İncelenmesi. *İlköğretim Online*, 17(2).
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*, 103(3), 411.
- Asil, M., & Gelbal, S. (2012). PISA öğrenci anketinin kültürler arası eşdeğerliği. *Eğitim ve Bilim*, 37(166).
- Ayvallı, M. (2016). PISA 2012 matematik okur-yazarlığı testinin ölçme değişmezliğinin incelenmesi. *Akdeniz Üniversitesi, Antalya*
- Bahadır, E. (2012). Uluslararası öğrenci değerlendirme programına (PISA 2009) göre Türkiye'deki öğrencilerin okuma becerilerini etkileyen değişkenlerin bölgelere göre incelenmesi. *Unpublished master's thesis, Hacettepe University, Ankara*.
- Balcı, A. (2001). *Sosyal bilimlerde araştırma: Yöntem, teknik ve ilkeler*. Pegem A Yayıncılık.
- Bayram, N. (2010). *Yapısal Eşitlik Modellemesine Giriş AMOS Uygulamaları*. Ezgi Kitabevi
- Başusta, N. B. U. (2010). Ölçme eşdeğerliği. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2).
- Başusta, N. B., & Gelbal, S. (2015). Gruplararası karşılaştırmalarda ölçme değişmezliğinin test edilmesi: PISA öğrenci anketi örneği. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30(4), 80-90.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Newyork: Guilford Publications.
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20(4), 872-882.

- Byrne, B. M., & Stewart, S. M. (2006). Teacher's corner: The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling*, 13(2), 287-321.
- Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of management*, 25(1), 1-27.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling*, 14(3), 464-504.
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2016). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Pegem Akademi.
- Demir, B. (2016). *Yapısal eşitlik modeli ile öğrencilerin iş bulma kaygılarına yönelik ölçek geliştirme: Cumhuriyet Üniversitesi İİBF'de bir uygulama* (Yayınlanmamış yüksek lisans tezi). Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, Sivas
- Demir, E. (2017). Testing Measurement Invariance of the Students' Affective Characteristics Model across Gender Sub-Groups. *Educational Sciences: Theory and Practice*, 17(1), 47-62.
- Dibek, M. İ., & Demirtaşlı, R. N. (2017). Öğrenme ve Öğretme Süreci Değişkenleri ile PISA 2012 Matematik Okuryazarlığı Arasındaki İlişkiler. *İlköğretim Online*, 16(3).
- Dural, S., Somer, O., Korkmaz, M., Can, S., & Öğretmen, T. (2011). İkinci derece örtük gelişme modelleri ve ölçme eşdeğerliği. *Eğitim ve Bilim*, 36(161).
- Dursun, Y., & Kocagöz, E. Yapısal Eşitlik Modellemesi Ve Regresyon: Karşılaştırmalı Bir Analiz. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, (35), 1-17.
- Ercikan, K., & Koh, K. (2005). Examining the construct comparability of the English and French versions of TIMSS. *International Journal of Testing*, 5(1), 23-35.
- Fidell, S., Tabachnick, B., & Pearsons, K. S. (2010). The state of the art of predicting noise- induced sleep disturbance in field settings. *Noise and Health*, 12(47), 77.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment*, 7(3), 286.

- Gierl, M. J. (2000). Construct equivalence on translated achievement tests. *Canadian Journal of Education/Revue canadienne de l'éducation*, 280-296.
- 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(11 Suppl 3), S78.
- Herdman, M., Fox-Rushby, J., & Badia, X. (1998). A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Quality of life Research*, 7(4), 323-335.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), 409-426.
- Jöreskog, K. G., & Sörbom, D. (1982). Recent developments in structural equation modeling. *Journal of marketing research*, 404-416.
- Karaduman, B., & Kilmen, S. (2018). Sınav Stresi Ölçeğinin Türkçeye Uyarlanması ve Ölçme Değişmezliğinin İncelenmesi Adaptation of the Examination Stress Scale into Turkish and Examination of Measurement Invariance. *Journal of Measurement and Evaluation in Education and Psychology*, 9(2), 101-115.
- Kaynak, Z. N. (2012). *Yapısal eşitlik modelleri* (Master's thesis, İstanbul Ticaret Üniversitesi).
- Khine, M. S. (Ed.). (2013). *Application of structural equation modeling in educational research and practice* (Vol. 7). Rotterdam, NL: Sense Publishers.
- Kıbrıslıoğlu, N. (2015). Pısa 2012 Matematik Öğrenme Modelinin Kültürlere ve Cinsiyete Göre Ölçme Değişmezliğinin İncelenmesi: Türkiye-Çin (Şangay)-Endonezya Örneği.
- Kıncal, R. Y. (2013). *Bilimsel Araştırma Yöntemleri*. Nobel Yayıncılık
- Kline, R. B. (2005). Principles and practice of structural equation modeling. *New York: Guilford*.
- Korkmaz, M., Somer, O., & Güngör, D. (2013). Ergen örnekleme beş faktör kişilik envanteri'nin cinsiyetlere göre ortalama ve kovaryans yapılarıyla ölçme eşdeğerliği. *Eğitim ve Bilim*, 38(170).
- Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate behavioral research*, 32(1), 53-76.
- Maruyama, G. (1997). *Basics of structural equation modeling*. Sage. 139-140

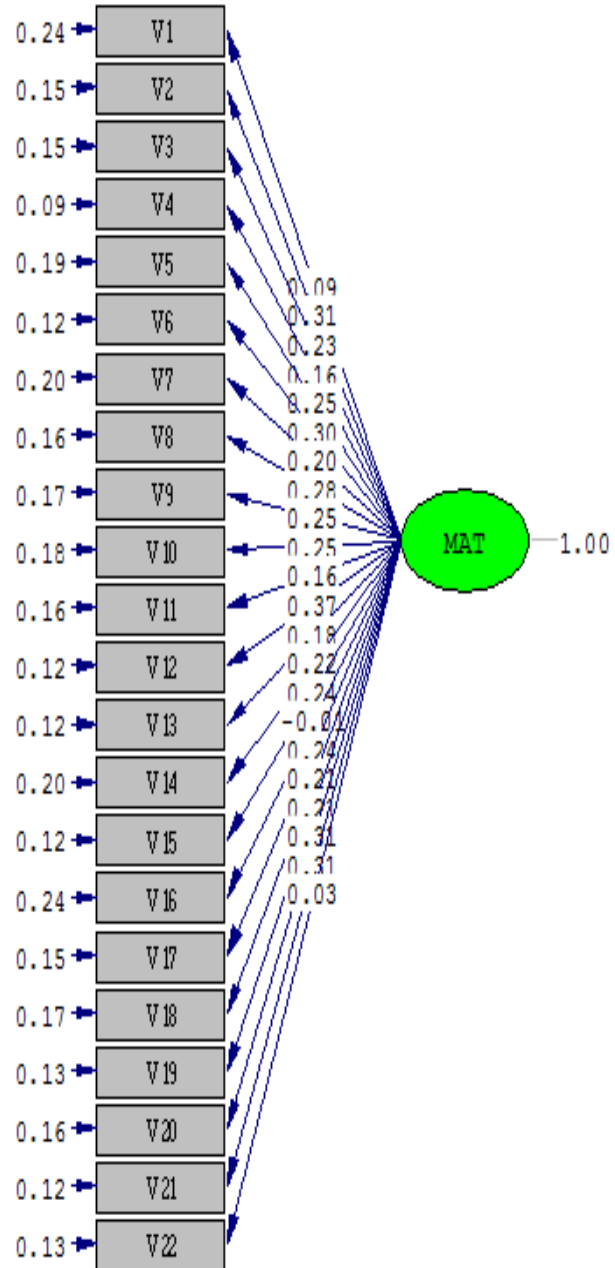
- Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J., Abdelfattah, F., Leung, K. C., ... & Parker, P. (2013). Factorial, convergent, and discriminant validity of timss math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology, 105*(1), 108.
- Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing, 6*(4), 311-360.
- 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.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-543.
- Meredith, W., & Millsap, R. E. (1992). On the misuse of manifest variables in the detection of measurement bias. *Psychometrika, 57*(2), 289-311.
- Organisation for Economic Co-operation and Development (2016). *PISA 2015 assessment and analytical framework: Science, reading, mathematics and financial literacy*. OECD publishing.
- Öğretmen, T. (2009). A comparison of the parametric methods based on the item response theory in determining differential item and test functioning. *Eğitim ve Bilim, 34*(152), 113.
- Ölçüoğlu, R. (2015). Tıms 2011 Türkiye Sekizinci Sınıf Matematik Başarısını Etkileyen Değişkenlerin Bölgelere Göre İncelenmesi.
- Önen, E. (2007). Gruplar Arası Karşılaştırmalarda Ölçme Değişmezliğinin İncelenmesi: Epistemolojik İnançlar Envanteri Üzerine Bir Çalışma. *Ege Eğitim Dergisi; Yıl: 2007 Cilt: 8 Sayı: 2*.
- Önen, E. (2009). Ölçme değişmezliğinin yapısal eşitlik modelleme teknikleri ile incelenmesi. *Yayınlanmamış Doktora Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara*.

- 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.
- 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.
- Somer, O., Korkmaz, M., Dural, S., & Can, S. (2009). Ölçme Eşdeğerliğinin Yapısal Eşitlik Modellemesi ve Madde Cevap Kuramı Kapsamında İncelenmesi. *Türk Psikoloji Dergisi, 24*(64).
- Şekercioğlu, G. (2009). Çocuklar için benlik algısı profilinin uyarlanması ve faktör yapısının farklı değişkenlere göre eşitliğinin test edilmesi. *Yayınlanmamış Doktora Tezi, Ankara Üniversitesi, Ankara.*
- Şeyma, U. Y. A. R., & DOĞAN, N. (2011). PISA 2009 Türkiye örnekleminde öğrenme stratejileri modelinin farklı gruplarda ölçme değişmezliğinin incelenmesi. *Uluslararası Türk Eğitim Bilimleri Dergisi, 2014*(3), 30-43.
- Taş, U. E., Arıcı, Ö., Ozarkan, H. B., & Özgürlük, B. (2016). PISA 2015 ulusal raporu. *Ankara: MEB.*
- Uzun, B., & Öğretmen, T. (2010). Fen başarısı ile ilgili bazı değişkenlerin TIMSS-R Türkiye örnekleminde cinsiyete göre ölçme değişmezliğinin değerlendirilmesi. *Eğitim ve Bilim, 35*(155).
- 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.
- 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), 1-26.
- Yandı, A., Köse, İ. A., & Uysal, Ö. (2017). Farklı Yöntemlerle Ölçme Değişmezliğinin İncelenmesi: Pisa 2012 Örneği. *Mersin Üniversitesi Eğitim Fakültesi Dergisi, 13*(1), 243-253.
- Yener, H. (2007). Personel performansına etki eden faktörlerin yapısal eşitlik modeli (YEM) ile incelenmesi ve bir uygulama. *Yayınlanmamış Doktora Tezi.*

Yildiz, M. A. (2017). Adaptation of General Belongingness Scale into Turkish for Adolescents: Validity and Reliability Studies. *Journal of Education and Training Studies*, 5(2), 223- 231.

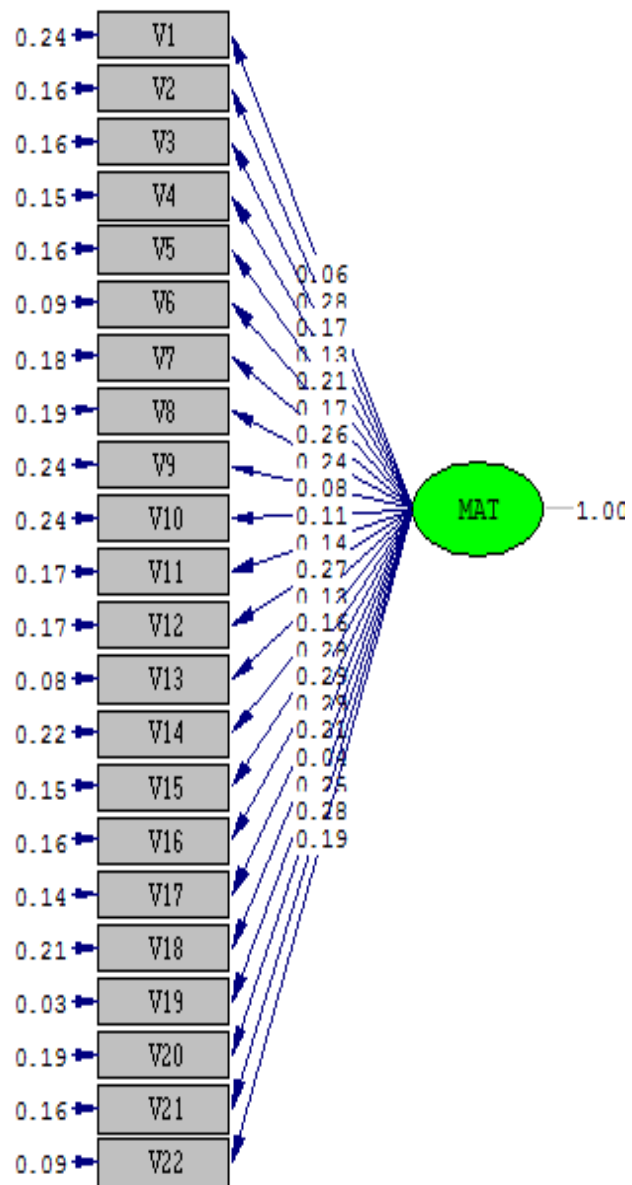
EK 1. 12 ÜLKE İÇİN DOĞRULAYICI FAKTÖR ANALİZİ YOL DİAGRAMLARI

Çin



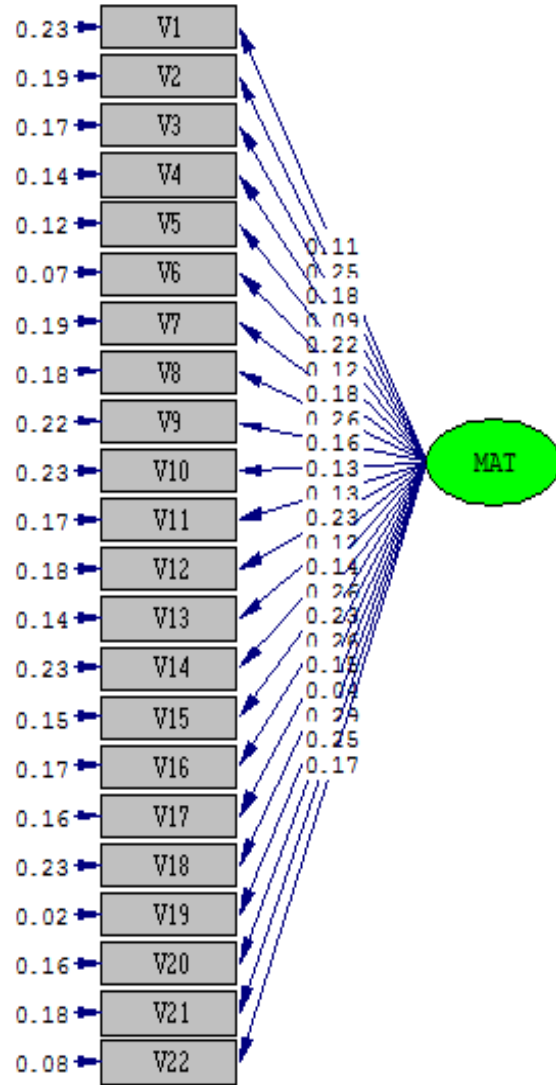
Chi-Square=284.20, df=209, P-value=0.00041, RMSEA=0.041

Almanya



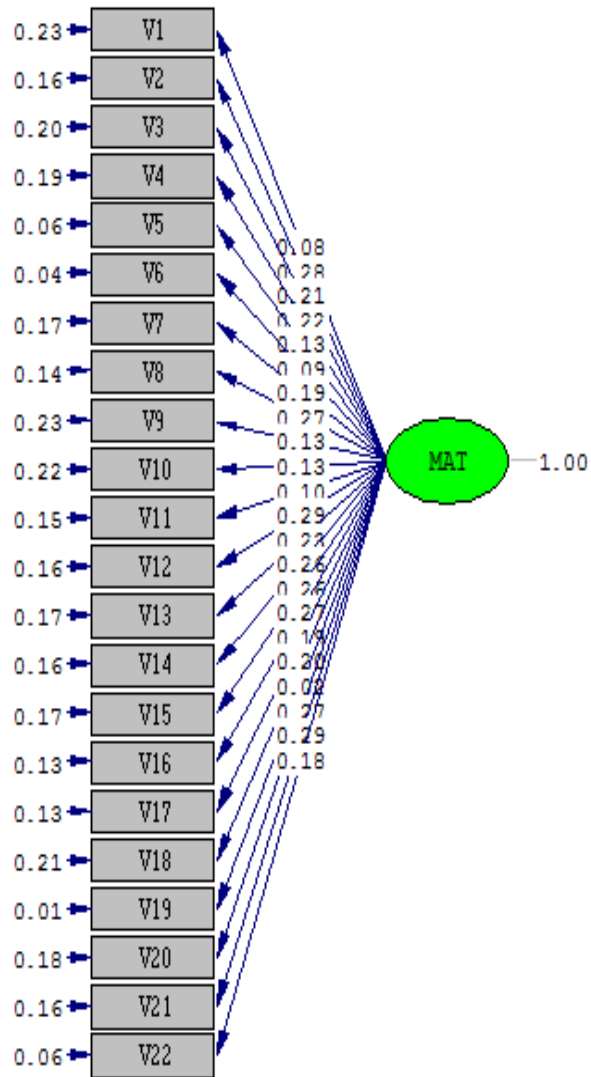
Chi-Square=218.64, df=209, P-value=0.30967, RMSEA=0.016

Birleşik Krallık



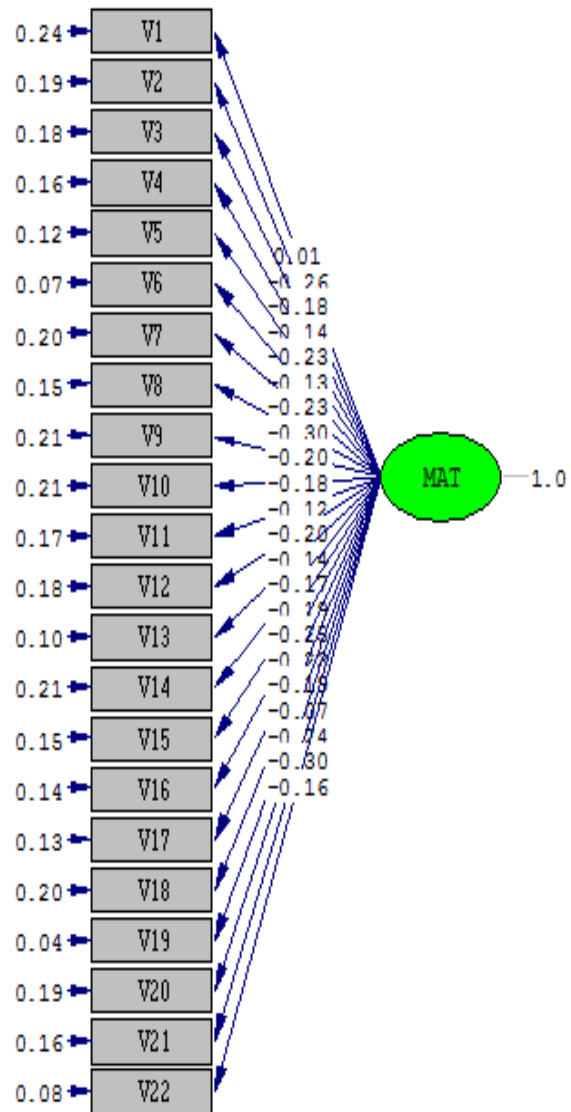
Chi-Square=417.28, df=209, P-value=0.00000, RMSEA=0.

İsrail



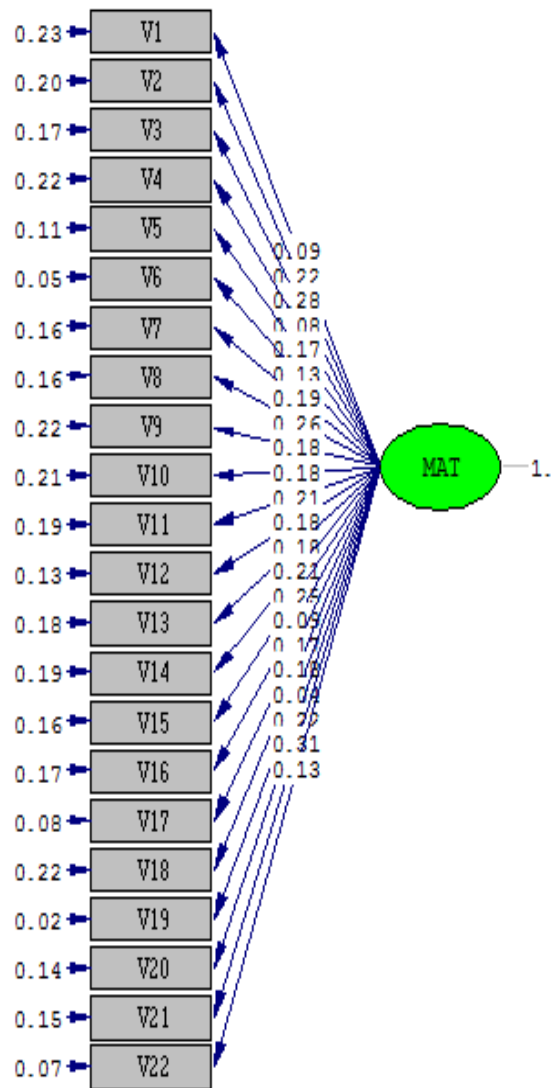
Chi-Square=380.42, df=209, P-value=0.00000, RMSEA=0.066

Italya



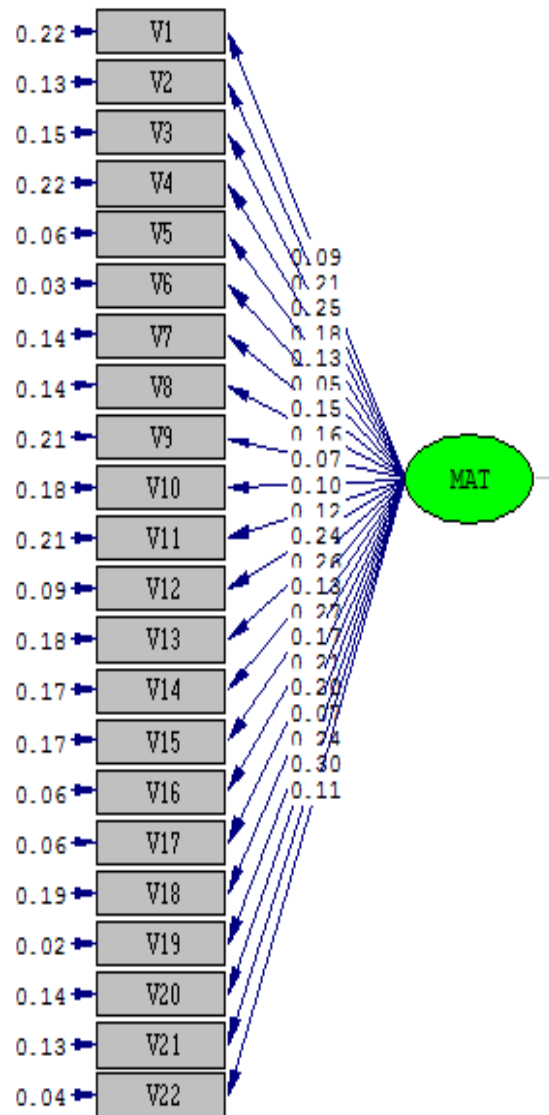
Chi-Square=387.68, df=209, P-value=0.00000, RMSEA=0.053

Litvanya



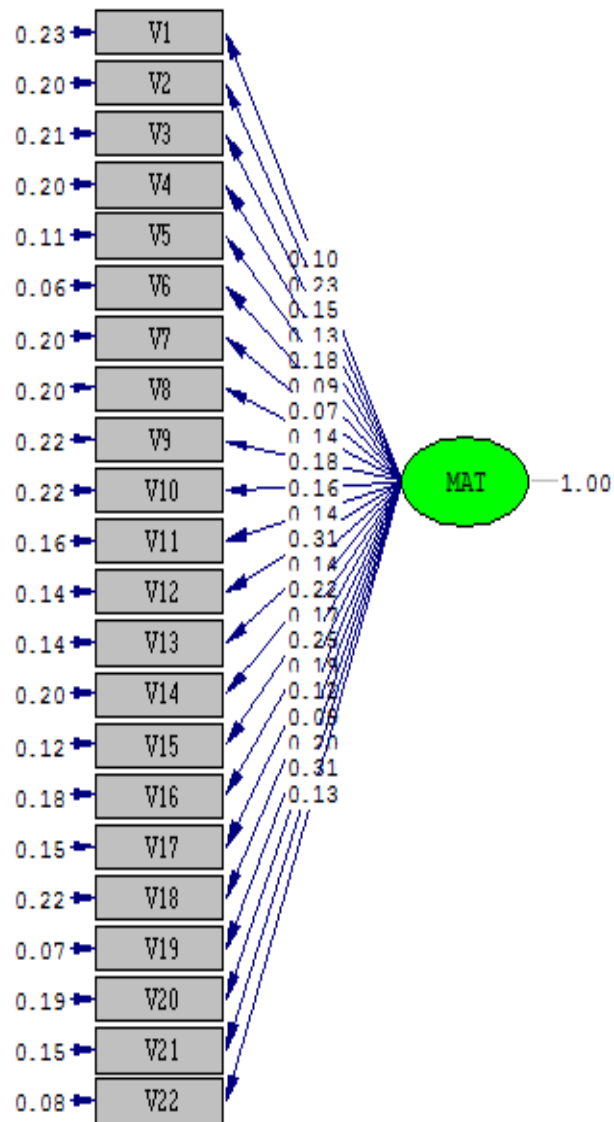
Chi-Square=273.74, df=209, P-value=0.00175, RMSEA=0.042

Katar



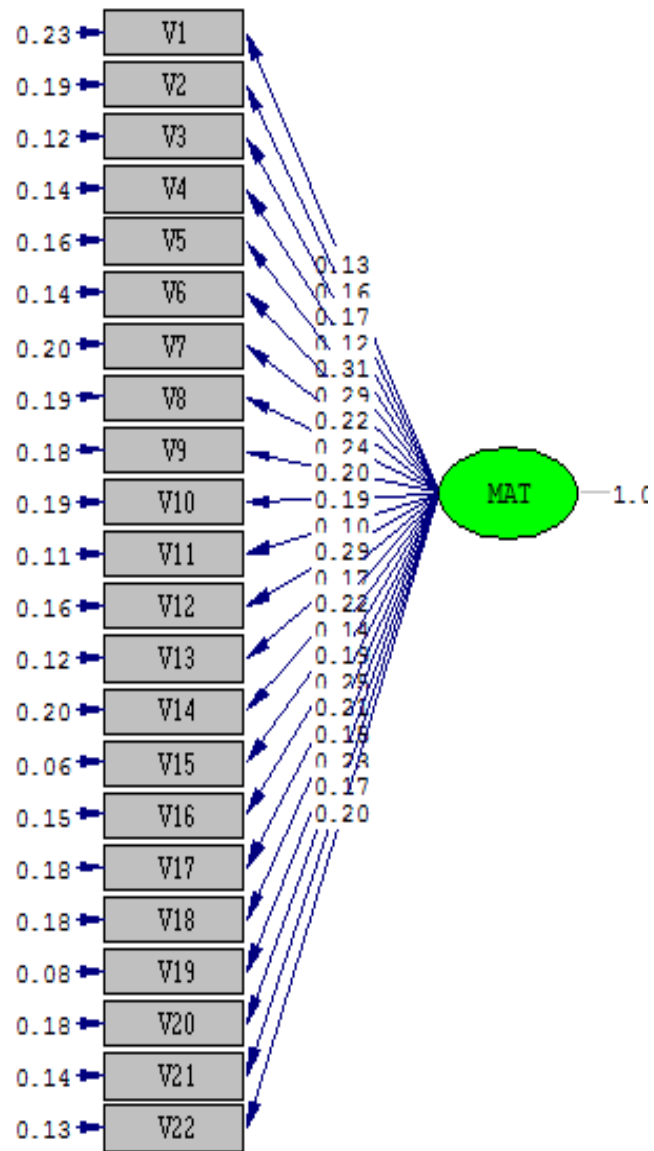
Chi-Square=573.84, df=209, P-value=0.00000, RMSEA=0.06

Rusya



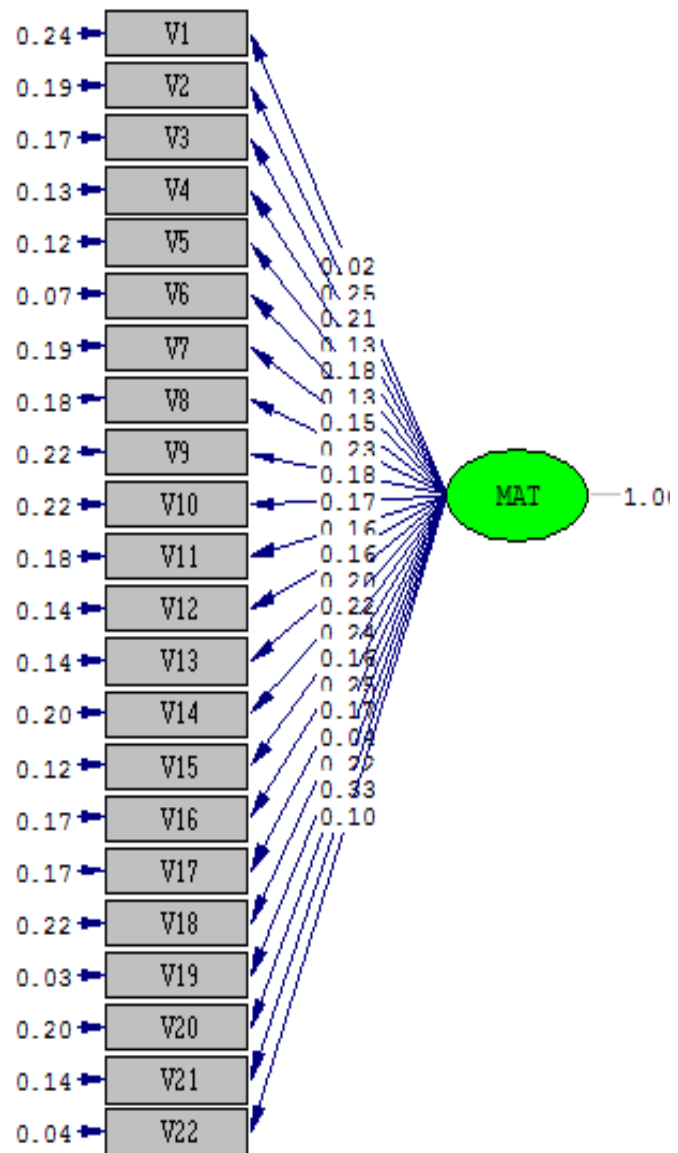
Chi-Square=255.29, df=209, P-value=0.01587, RMSEA=0.036

Singapur



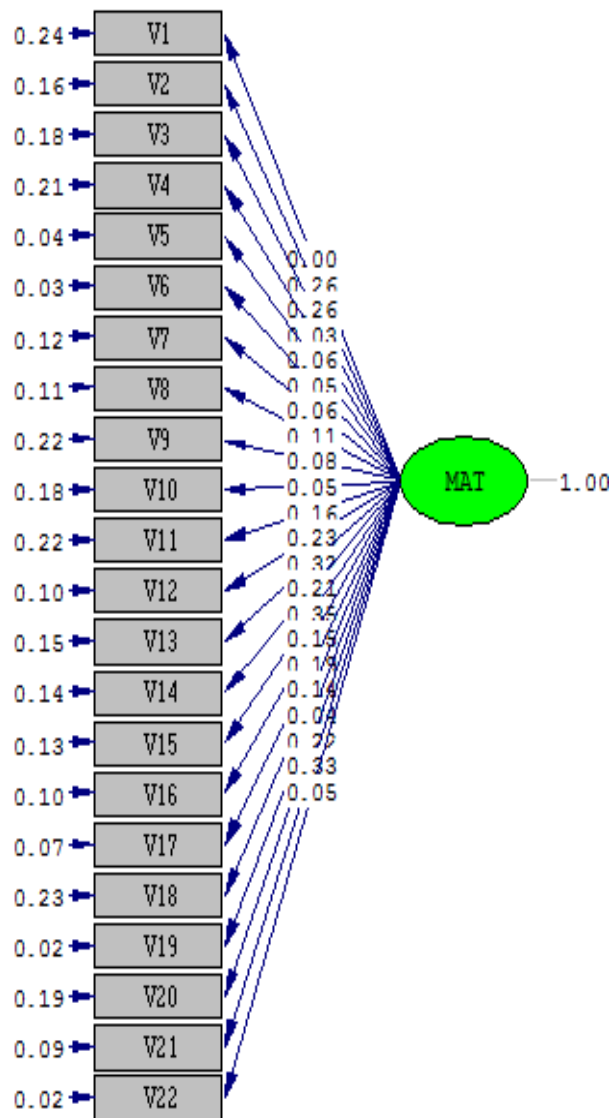
Chi-Square=258.80, df=209, P-value=0.01079, RMSEA=0.038

Slovakya



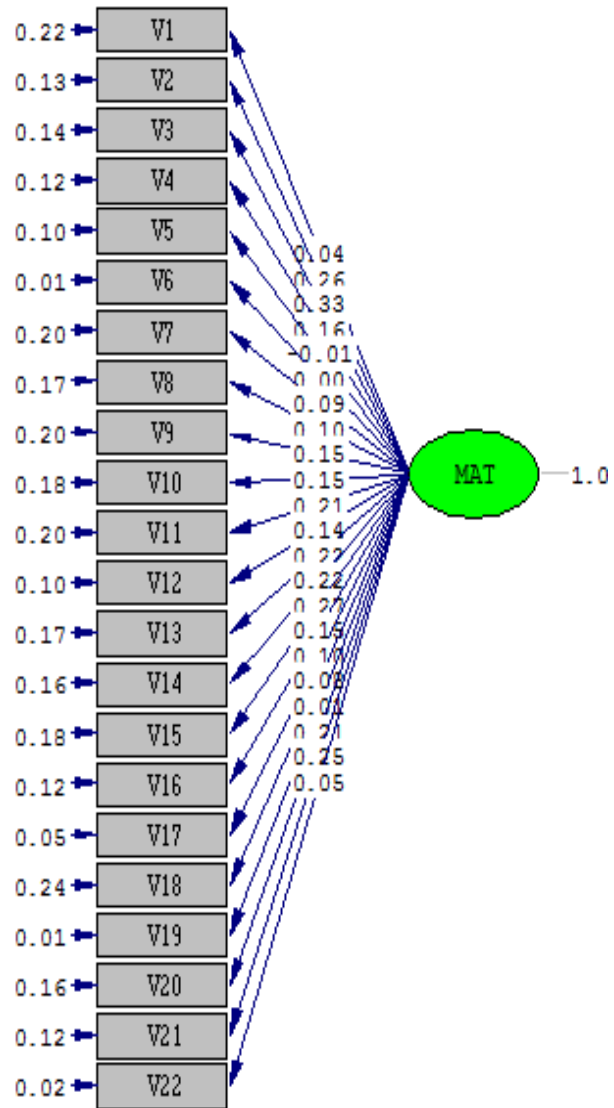
Chi-Square=333.43, df=209, P-value=0.00000, RMSEA=0.058

Tayland



Chi-Square=411.49, df=209, P-value=0.00000, RMSEA=0.065

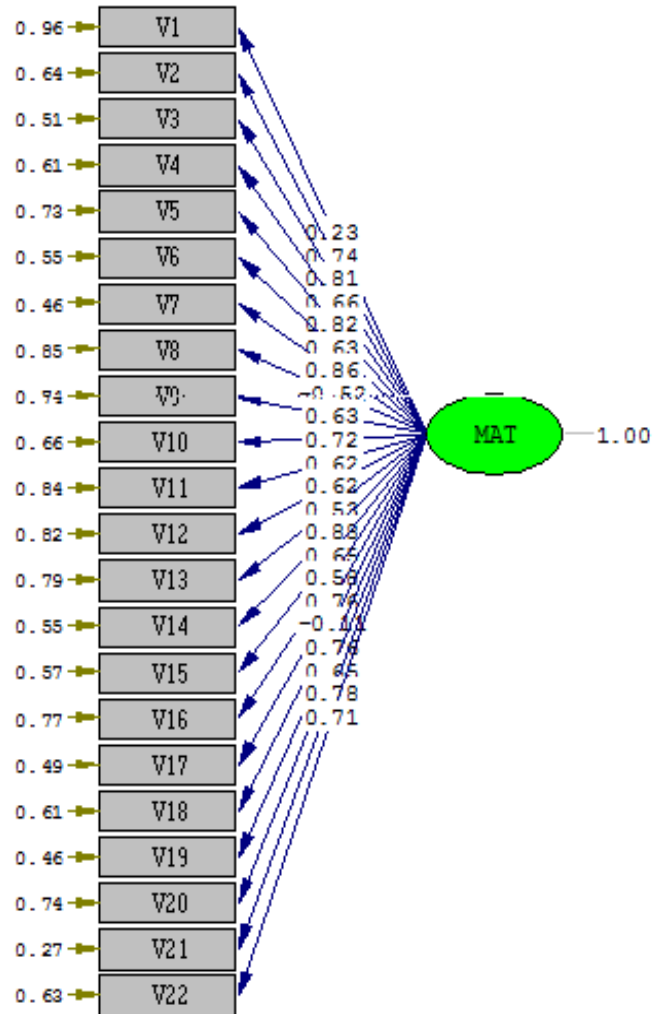
Türkiye



Chi-Square=305.01, df=209, P-value=0.00002, RMSEA=0.053

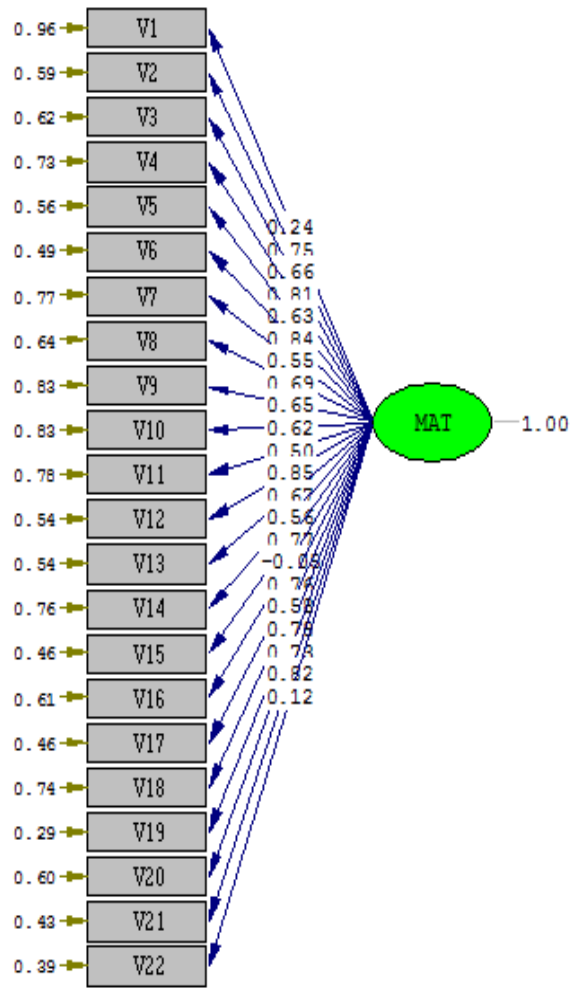
EK 2. DEĞİŞMEZLİK AŞAMALARINA İLİŞKİN YOL DİYAGRAMLARI

A. ŞEKİL DEĞİŞMEZLİĞİ



Chi-Square=8816.64, df=2750, P-value=0.00000, RMSEA=0.097

B. METRİK DEĞİŞMEZLİK



Chi-Square=8089.51, df=2750, P-value=0.00000, RMSEA=0.091

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Esmâ ASAR

Doğum Yeri ve Tarihi : ALTINDAĞ 05.01.1993

Eğitim Durumu

Lisans Öğrenimi: Akdeniz Üniversitesi İlköğretim Matematik Öğretmenliği

Yüksek Lisans Öğrenimi : Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim

Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme

Bildiği Yabancı Diller : İngilizce

Bilimsel Faaliyetleri :

İletişim

E-Posta Adresi : esmaolgun2020@gmail.com

Tarih: 03.05.2018

İNTİHAL RAPORU



Hakan Koçar | Kullanıcı Dâğıtı | Mesajlar | Öğrenciler | Türkçe | Topluak | Yardım | Çık

Ödevler Öğrenciler Not Defteri Kütüphaneler Takvim Tartışma Tesitler

GÖRÜNTÜLENİYOR: ANASAYFA - ESMA - ESMA TEZ

Bu sayfa hakkında

Bu sizin ödev kutunuzdur. Bir yazılı ödevi görüntülemek için yazılı ödevin başlığını seçin. Bir Benzerlik Raporunu görüntülemek için yazılı ödevin benzerlik sütunundaki Benzerlik Raporu ikonunu seçin. Taklanabilir durumda olmayan bir ikon Benzerlik Raporunun henüz oluşturulmadığını gösterir.

esma tez

GELEN KUTUSU | GÖRÜNTÜLENİYOR: YENİ ÖDEVLER

Dosyayı Gönder

Çevrimiçi Derecelendirme Raporu | Ödev ayarlarını düzenle | E-posta bildirmeyenler

TAZAR	BAŞLIK	BENZERLİK	PUANLA	CEYAP	DOĞYA	ÖDEV NUMARASI	TARİHİ
<input type="checkbox"/>	Esmâ Ölgün	Esmâ tez	%12	--	--	1110325289	11-Nis-2019

Danışman Dr. Öğr. Üyesi:

Hakan KOÇAR