



HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı

Eğitimde Ölçme ve Değerlendirme Bilim Dalı

TIMSS 2019 MATEMATİK BAŞARI TESTİ VE MATEMATİĞE İLİŞKİN DUYUŞSAL ÖZELLİKLERİN ÖLÇME DEĞİŞMEZLİĞİNİN İNCELENMESİ

Ferda KOÇAK

Yüksek Lisans Tezi

Ankara, 2024

Liderlik, arařtırma, inovasyon, kaliteli eđitim ve deęiřim ile

Daha ileriye... En İyiyeye...



Eğitim Bilimleri Ana Bilim Dalı

Eğitimde Ölçme ve Değerlendirme Bilim Dalı

TIMSS 2019 MATEMATİK BAŞARI TESTİ VE MATEMATİĞE İLİŞKİN DUYUŞSAL
ÖZELLİKLERİN ÖLÇME DEĞİŞMEZLİĞİNİN İNCELENMESİ

INVESTIGATION OF MEASUREMENT INVARIANCE OF TIMSS 2019
MATHEMATICS ACHIEVEMENT TEST AND MATHEMATICAL ATTITUDINAL
CHARACTERISTICS

Ferda KOÇAK

Yüksek Lisans Tezi

Ankara, 2024

Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

Ferda KOÇAK'ın hazırladıđı “TIMSS 2019 Matematik Başarı Testi ve Matematiđe İlişkin Duyuşsal Özelliklerin Ölçme Deđişmezliđinin İncelenmesi” başlıklı bu çalıřma j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalında Yüksek Lisans** olarak kabul edilmiřtir.

J¼ri Başkanı	İmza
J¼ri Üyesi (Danıřman)	İmza
J¼ri Üyesi	İmza

Bu tez Hacettepe Üniversitesi Lisansüstü Eđitim, Öğretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 13/02/2024 tarihinde uygun gör¼lmüş ve Enstitü Yönetim Kurulunca / / tarihi itibarıyla kabul edilmiřtir.

Prof. Dr. İsmail Hakkı MİRİCİ
Eđitim Bilimleri Enstitüsü M¼d¼r¼

Öz

TIMSS uygulamaları öğrencilerin matematik ve fen alanlarında sahip oldukları bilgi ve becerileri ortaya koymakta ve ülkelere bu alanlardaki eğitimlerini geliştirmelerine ve başarılarını yordamalarına destek olmaktadır. TIMSS uygulamalarından elde edilen sonuçlarla farklı ülkelerdeki öğrencilerin matematik ve fen alanlarındaki başarıları ve çeşitli özellikleri karşılaştırılmaktadır. Bu noktada karşılaştırmalara ilişkin yorumların isabetli olması için TIMSS uygulamalarında kullanılan başarı testi ve ölçeklerin ölçme değişmezliğine dair kanıtların toplanması gerekmektedir. Söz konusu ihtiyaçtan hareketle bu araştırmada, TIMSS 2019 uygulamasındaki 8.sınıf matematik başarı testinin ve duyuşsal özellikler ile ilgili olan “Matematik Öğrenmeyi Sevme”, “Matematikte Kendine Güven” ve “Matematiğe Değer Verme” ölçeklerinden oluşturulan modelin ve ayrı ayrı her bir ölçeğin Türkiye, Güney Kore ve Finlandiya ülkelerine göre ölçme değişmezliğinin test edilmesi amaçlanmıştır. Başarı testi, duyuşsal özelliklerle ilgili oluşturulan model ve belirtilen her bir ölçek için önce açımlayıcı faktör analizleri ve ardından doğrulayıcı faktör analizleri yapılmıştır. Ayrıca her bir ülke için doğrulama işleminin yapılmasının ardından üç ülkeye ait veriler birlikte kullanılarak ölçme değişmezliği Yapısal Eşitlik Modeli'ne dayalı Çok Gruplu Doğrulayıcı Faktör Analizi ile test edilmiştir. Araştırma sonucunda başarı testi, duyuşsal özelliklerle ilgili oluşturulan model ve her bir ölçek için “ölçek değişmezliği” aşamasına ilişkin kanıt sağlanamamış, duyuşsal özelliklerle ilgili oluşturulan model ve her bir ölçek için “metrik değişmezlik” aşaması sağlanmıştır. Bu nedenle ortaya çıkan farklılıkların ülkelere kaynaklandığını belirtmek doğru olmayacaktır. Dolayısıyla ülkelere ilgili yorumda bulunurken bu durum ihmal edilmemelidir.

Anahtar sözcükler: TIMSS, ölçme değişmezliği, YEM, ÇGDFA, matematik başarıları, Matematik Öğrenmeyi Sevme, Matematikte Kendine Güven ve Matematiğe Değer Verme

Abstract

TIMSS implementations reveal the knowledge and skills of students in mathematics and science and support countries in improving their education in these fields and predicting their achievements. With the results obtained from TIMSS implementations, the achievements and various characteristics of students in mathematics and science in different countries are compared. At this point, in order for the interpretations regarding the comparisons to be accurate, it is necessary to collect evidence on the measurement invariance of the achievement tests and scales used in TIMSS implementations. Based on this need, this study aimed to test the measurement invariance of the 8th grade mathematics achievement test in the TIMSS 2019 implementation and the model created from the scales of "Students Like Learning Mathematics", "Students Confident in Mathematics" and "Students Value Mathematics" related to affective traits and each scale separately according to the countries of Turkey, South Korea and Finland. First, exploratory factor analyses and then confirmatory factor analyses were conducted for the achievement test, the model of affective traits and each scale. In addition, after the validation process for each country, measurement invariance was tested with Multiple Group Confirmatory Factor Analysis based on Structural Equation Modeling by using data from three countries together. As a result of the research, the evidence for the "scale invariance" stage was not provided for the achievement test, the model for affective traits and each scale, while the "metric invariance" stage was provided for the model for affective traits and each scale. Therefore, it would not be correct to state that the differences that emerged are due to the countries. Therefore, this situation should not be neglected when making comments about the countries.

Keywords: TIMSS, measurement invariance, SEM, MGCFA, mathematics achievement, Students Like Learning Mathematics, Students Confident in Mathematics, Students Value Mathematics

Teşekkür

Tez yazım sürecinde akademik gelişimime önemli katkılar sunduğu ve bana sağladığı destek ve rehberlik için kıymetli danışmanım Prof. Dr. Nuri DOĞAN'a,
Tez jürimde görev alarak tezime sundukları değerli ve yapıcı dönütleri için sevgili hocalarım Doç. Dr. Sevda ÇETİN ve Dr. Öğretim Üyesi Cansu AYAN'a,
Ellerindeki her türlü kaynağı benimle paylaşan ve destekleriyle beni motive eden sevgili arkadaşlarım Dr. Emine ŞİMŞEK ve Demet KIL'a,
Araştırma sürecinde bilgi birikimiyle araştırmama destek veren kıymetli eşim KORKUT KOÇAK'a,
Varlığı ile hayatıma anlam ve neşe katan biricik oğlum Batu Mert KOÇAK'a en içten teşekkürlerimi sunarım.

İçindekiler

Kabul ve Onay	ii
Öz.....	iii
Abstract	iv
Teşekkür.....	v
Tablolar Dizini.....	viii
Şekiller Dizini	ix
Simgeler ve Kısaltmalar Dizini	x
Bölüm 1 Giriş	1
Problem Durumu	1
Araştırmanın Amacı ve Önemi	7
Araştırma Problemi.....	10
Sınırlılıklar	11
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar	12
Ölçme Değişmezliği.....	12
Yapısal Eşitlik Modeli	18
Doğrulamalı Faktör Analizi	21
Çok Gruplu Doğrulamalı Faktör Analizi	22
İlgili Araştırmalar.....	24
Bölüm 3 Yöntem	29
Araştırmanın Türü	29
Çalışma Grubu	29
Veri Toplama Süreci.....	30
Veri Toplama Araçları.....	31
Verilerin Analizi.....	38
Bölüm 4 Bulgular, Yorumlar ve Tartışma	61
Araştırmanın 1. Alt Problemine İlişkin Bulgular	61

Araştırmanın 2. Alt Problemine İlişkin Bulgular	64
Araştırmanın 3. Alt Problemine İlişkin Bulgular	68
Bölüm 5 Sonuç ve Öneriler	76
Sonuçlar	76
Öneriler	78
Kaynaklar.....	81
EK-A: Yığılma Grafikleri.....	92
EK-B: Tetrakorik Korelasyon Tablosu.....	94
EK-C: Polikorik Korelasyon Tablosu	94
EK-Ç: Araştırma Etik Komisyon İzin Muafiyeti Formu.....	96
EK-D: Etik Beyanı	97
EK-E: Yüksek Lisans Tez Çalışması Orijinallik Raporu	98
EK-F: Thesis/Dissertation Originality Report	98
EK-G: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı.....	100

Tablolar Dizini

Tablo 1 Ölçme Değişmezliği Aşamalarına Göre Yapılan Kısıtlamalar	15
Tablo 2 Ülkelere Göre Öğrenci Sayıları	30
Tablo 3 Veri Toplama Araçları	31
Tablo 4 Öğrenme ve Konu Alanlarının Dağılımı	32
Tablo 5 Bilişsel Alanların Dağılımı.....	33
Tablo 6 Başarı Testindeki Maddeler ve Maddelere İlişkin Bilgiler	34
Tablo 7 Duyuşsal Özelliklerle İlgili Oluşturulan Modelde Yer Alan Ölçeklerin Maddeleri.....	37
Tablo 8 Başarı Testindeki Maddelerin Çarpıklık ve Basıklık Değerleri.....	39
Tablo 9 Başarı Testindeki Maddelerin Faktör Yükleri	41
Tablo 10 Başarı Testi Uyum Değerleri	43
Tablo 11 Duyuşsal Özellik Modelindeki Maddelerin Çarpıklık ve Basıklık Değerleri	45
Tablo 12 Duyuşsal Özellik Modelindeki Maddelerin Faktör Yükleri.....	47
Tablo 13 Duyuşsal Özellik Modeli Uyum Değerleri.....	50
Tablo 14 Matematik Öğrenmeyi Sevme Ölçeğindeki Maddelerin Faktör Yükleri.....	51
Tablo 15 Matematik Öğrenmeyi Sevme Ölçeği Uyum Değerleri.....	53
Tablo 16 Matematikte Kendine Güven Ölçeğindeki Maddelerin Faktör Yükleri.....	54
Tablo 17 Matematikte Kendine Güven Ölçeği Uyum Değerleri.....	56
Tablo 18 Matematiğe Değer Verme Ölçeğindeki Maddelerin Faktör Yükleri.....	57
Tablo 19 Matematiğe Değer Verme Ölçeği Uyum Değerleri	58
Tablo 20 Ülkelere Göre Başarı Testinin Uyum Değerleri	61
Tablo 21 Ülkelere Göre Başarı Testinin ÇGDFA Uyum Değerleri.....	62
Tablo 22 Ülkelere Göre Duyuşsal Özellik Modelinin Uyum Değerleri	65
Tablo 23 Ülkelere Göre Duyuşsal Özellik Modelinin ÇGDFA Uyum Değerleri	65
Tablo 24 Ülkelere Göre Matematik Öğrenmeyi Sevme Ölçeğinin Uyum Değerleri	69
Tablo 25 Ülkelere Göre Matematik Öğrenmeyi Sevme Ölçeğinin ÇGDFA Uyum Değerleri ..	69
Tablo 26 Ülkelere Göre Matematikte Kendine Güven Ölçeğinin Uyum Değerleri	71
Tablo 27 Ülkelere Göre Matematikte Kendine Güven Ölçeğinin ÇGDFA Uyum Değerleri	71
Tablo 28 Ülkelere Göre Matematiğe Değer Verme Ölçeğinin Uyum Değerleri	73
Tablo 29 Ülkelere Göre Matematiğe Değer Verme Ölçeğinin ÇGDFA Uyum Değerleri	74

Şekiller Dizini

Şekil 1 <i>Başarı Testi Yol Diyagramı</i>	42
Şekil 2 <i>Duyuşsal Özellik Modeli Yol Diyagramı</i>	49
Şekil 3 <i>Matematik Öğrenmeyi Sevme Ölçeği Yol Diyagramı</i>	52
Şekil 4 <i>Matematikte Kendine Güven Ölçeği Yol Diyagramı</i>	55
Şekil 5 <i>Matematiğe Değer Verme Ölçeği Yol Diyagramı</i>	58

Simgeler ve Kısaltmalar Dizini

ÇGDFA: Çok Gruplu Doğrulayıcı Faktör Analizi

DFA: Doğrulayıcı Faktör Analizi

IEA: International Association for the Evaluation of Educational Achievement

MTK: Madde Tepi Kuramı

OECD: The Organisation for Economic Co-operation and Development

PIRLS: Progress In International Reading Literacy Study

PISA: Programme for International Student Assessment

TALIS: The Teaching and Learning International Survey

TIMSS: Trends In International Mathematics And Science Study

YEM: Yapısal Eşitlik Modeli

Bölüm 1

Giriş

Bu bölümde problem durumu, araştırmanın amacı ve önemi, problem cümlesi, alt problemler ve sınırlılıklar yer almaktadır.

Problem Durumu

Eğitim bir insan hakkı, kalkınmanın ise ana öğelerinden biridir. Fırsat eşitliğini sağlamak, yoksulluğu azaltmak ve sağlığı, barışı, istikrarı iyileştirmek için en önemli araçlardan biri eğitimidir. Ülkeler insanların eğitimlerine yatırım yaparak beşeri sermayelerini güçlendirmekte, bireylerin yaşadığımız çağda başarılı olmalarını sağlayacak bilişsel, sosyal, duygusal, teknik ve dijital beceriler gibi 20.yy becerilerini edinmelerine destek olmaktadır. Beşeri sermayesi yüksek bireyler katma değeri yüksek ürün ve hizmet sunacak önemli bir değer olarak görülmektedir (OECD, 2023). Ülkeler, alanında etkin ve yetkin bireyler yetiştirerek geleceklerini güvenli bir biçimde inşa edebilmekte ve böylelikle diğer ülkelerle rekabet edebilir duruma gelmektedir. Bu nedenle ülkelerin yetişmiş insan gücü önemli görülmektedir. Dolayısıyla ülkeler yetiştirdikleri bireylerin sahip olduğu becerileri tespit etmek ve hatta gelecekteki potansiyellerini öngörmek istemektedir. Bu nedenle ülkeler eğitim sistemlerinin değerlendirilmesine özen göstermektedir.

Ülkeler eğitim sistemlerinin mevcut durumunu ortaya koymak ve eğitim sisteminin farklı bileşenlerine göre ülkeler arasında karşılaştırmalar yapmak amacıyla uluslararası araştırmalara gönüllü olarak katılmaktadır. Böylelikle ülkeler eğitim sistemlerinin küresel çağın ihtiyaçlarını karşılama kapasitelerini öngörebilmekte ve dolayısıyla eğitim politikalarının hangi yönde geliştirilmesi gerektiğine ilişkin önemli veriler elde etmektedir. Bu bağlamda Türkiye, OECD (The Organisation for Economic Co-operation and Development) tarafından yapılan PISA (Programme for International Student Assessment), TALIS (The Teaching and Learning International Survey); IEA (International Association for the Evaluation of Educational Achievement) tarafından yapılan TIMSS (Trends In International Mathematics

And Science Study), PIRLS (Progress In International Reading Literacy Study) gibi birçok uluslararası arařtırmada yer almaktadır. Dünya apında geniř lekte uygulanan zellikle PISA ve TIMMS arařtırmaları, arařtırma sonularına iliřkin lkeler arası karřılařtırmalar lkemizde kamuoyu tarafından yakından takip edilmektedir. Ancak bu uygulamalardan elde edilen sonulara dayanan yararlanılabilecek karřılařtırmalar yapılabilmesi iin karřılařtırılan lkeler baėlamında lme deėiřmezliėinin saėlandıėını kanıtlamak gerekmektedir (Chen, 2007; Cheung & Rensvold, 2002).

Uluslararası arařtırmalar, arařtırmacılar tarafından analiz edilmekte ve arařtırmaların sonuları toplumun geniř kesimi tarafından takip edilmektedir. Elde edilen sonular eėitim politikaları geliřtirenler iin kanıtlar oluřturmakta ve politikaların řekillenmesine katkı sunmaktadır. Alan yazında lkelerin uluslararası arařtırmaların sonularından faydalanma dzeylerinin dolayısıyla sonuların politikaları zerindeki etkisinin farklılařtıėını ortaya koyan arařtırmalar vardır. rneėin Grek (2009) arařtırmasında uluslararası arařtırma sonularının Avrupa'daki eėitim sistemlerinde yapılacak deėiřikliklere yn verdiėini belirtirken; Tan (2019) arařtırmasında Shanghai'da PISA sonularından eėitim reformu yapılırken faydalanıldıėını ifade etmiřtir. Ritzen (2013) arařtırmasında ise PISA uygulamalarına (2000-2006) katılan lkelerin elde ettikleri sonulardan hareketle politika geliřtirme dzeylerinde farklılařma olduėunu aıklamıř ve sz konusu arařtırmada Trkiye, PISA sonuları doėrultusunda dřk dzeyde politika reten lkeler arasında yer almıřtır. lkelerin uluslararası arařtırmalara olan ilgileri, buldukları sıralamalardan memnun olma durumları veya yksek sıralarda olma arzuları lkelerin arařtırma sonularından politika retme dzeylerini deėiřtirmektedir (Aydın, 2020). Nitekim uluslararası arařtırmalar lkelerin gndeminde yer almakta ve eėitim sistemleri ile politikalarını etkilemektedir (Grlen ve diėerleri, 2019).

Uluslararası arařtırmalara verilen nemden hareketle geniř lekli bu arařtırmalara ynelik eřitli raporlar hazırlanmaktadır. Raporlarda incelenen durumlardan biri de hem deėerlendirme erevelerinde yer alan ėrenme ve biliřsel alanlara ynelik elde edilen veriler hem de anketlerden ulařılan duyuřsal alana ynelik veriler kullanılarak eřitli karřılařtırmalar yapmaktadır. Bu alıřmalarda lkeler arasında sıralama alıřması, grupta alıřması,

başarı ve duyuşsal özelliklerin karşılaştırılması, başarıya etki eden faktörlerin karşılaştırılmasının yapıldığı, başarıyı yordayan değişkenlerin belirlendiği görülmektedir (Atılğan, 2022; Çakıcı Eser, 2021; Öncü, 2019). Oysaki karşılaştırma yapılan ülkeler arasında birçok farklılık vardır. Kültürel açıdan bakıldığında ülkeler dil, değer ve düşünme yapısı gibi farklı kültürel öğelere sahiptir. Sosyo-ekonomik açıdan ise ülkelerin refah seviyeleri, kişi başına düşen milli gelirleri, eğitime ayrılan bütçeleri, öğrenci başına yapılan yatırımları gibi birçok durum farklılık göstermektedir. Zira her bir milletin anıldığı özellikler de farklıdır. Örneğin Almanlar çalışkan, düzenli, dakik olarak değerlendirilirken Japonlar disiplinli, özdenetim ve çalışma etiği yüksek olmaları ile akıllara gelmektedir. Eğitim ile kültür arasındaki güçlü bağ düşünüldüğünde, sosyo-ekonomik düzey ile başarı arasındaki ilişki göz önüne alındığında ülkelerin bu özelliklerinin eğitim sistemlerine de yansıtacağı aşikârdır. Dolayısıyla eğitim sistemleri birbirinden farklılaşacaktır.

Bu karşılaştırmalarda ulaşılan farklılığın sadece bireyden ya da örneklemden kaynaklandığının düşünülmesi uygun olmayabilir. Bu nedenle farklılığın hangi değişkenden kaynaklandığını ifade etmeden önce farklılığın ölçme aracından meydana gelip gelmediğinin araştırılması gerekmektedir. Nitekim bireyler arasındaki farklılık ölçme aracının kendinden de kaynaklanabilir (Başusta & Gelbal, 2015). Dolayısıyla uluslararası araştırmalar üzerinden yapılacak karşılaştırma çalışmalarında kullanılacak ölçme araçlarının gruplar için aynı anlama geldiğini ortaya koymak zorunludur. Bu nedenlerle karşılaştırmaların yapılabilmesi için ölçme sonuçları karşılaştırılan ülkeler açısından ölçme aracının aynı özelliği ölçebildiği yönünde kanıtların toplanması gerekmektedir. Bu amaçla yararlanılacak çalışmalardan biri de ölçülen psikolojik özelliğin ülkelere göre aynı yapıyı gösterdiğinin kanıtı olan “ölçme değişmezliği” çalışmalarıdır. Ölçme değişmezliğinin sınanması ölçme aracının farklı ülkeler arasında benzer yapıya sahip olup olmadığı ortaya koymakta ve böylelikle ölçme aracının geçerliğine dair kanıt sağlamaktadır (Widaman & Reise, 1997).

TIMSS araştırmaları 4 ve 8. sınıf öğrencilerinin matematik ve fen başarılarındaki eğilimlerini ortaya koymaktadır. Benzer şekilde PISA araştırmaları da 15 yaşındaki öğrencilerin matematik ve fen okuryazarlıkları ile okuma becerilerine odaklanmaktadır. Fakat

PISA arařtırmalarında öğrencilerin gerçek hayatta karşılařması olası becerileri deęerlendirilirken TIMSS arařtırmalarında matematik ve fen alanlarında öğrencilerin kazandıkları bilgi ve becerilerin deęerlendirilmesi saęlanmaktadır. PISA arařtırmaları öğretim programlarından ziyade öğrencilerin okulda öğrendiklerinin günlük hayatta beceriye dönüşme durumlarını incelerken TIMSS arařtırmalarının öğretim programlarında yer alan bilgi ve becerileri yokladıęı dolayısıyla öğretim programlarıyla daha yakından iliřkili olduęu ifade edilebilir. Zira TIMSS arařtırmaları kapsamında öğretim programı anketi ile ülkelerin eğitim sistemleri, ülkelerde uygulanan matematik ve fen öğretim programlarının yapısı, içerięi ile ilgili politikalar ve uygulamalar hakkında bilgiler toplanmaktadır. Bu kapsamda her ülke matematik ve fen alanında ulusal olarak belirlenmiř (veya resmi) öğretim programını detaylandırmaktadır. TIMSS uygulamasına katılan öğrencilerin öğretmenlerine mevcut veya önceki okul yıllarında hangi TIMSS konularının işlendięi sorulmaktadır (Martin ve dięerleri, 2020). Böylelikle TIMSS arařtırmaları öğretim programlarına dair detaylı bilgiler elde edilmesine fırsat sunmaktadır. Dolayısıyla TIMSS arařtırmaları kanıta dayalı veriler sunarak öğretim programlarının geliřtirilmesine yönelik bilgiler saęlamakta ve böylelikle programlarda yapılacak deęiřikliklere yön verebilmektedir. Bu açıklamalar çerçevesinde TIMSS'e dayalı karşılařtırmaların özellikle öğretim programları üzerinde daha fazla etkisi olması beklendięinden ve alan yazında PISA uygulamalarına yönelik daha fazla arařtırma olduęundan (Asar; 2019; Ayvalli, 2016; Baędu Söyler ve dięerleri, 2021; Bařusta & Gelbal, 2015; Gönen, 2021; Kankaras & Moors, 2014; Karakoç Alatlı, 2016; Kıbrıslıoęlu, 2015) arařtırma kapsamına TIMSS sonuçları alınmıřtır.

TIMSS, ilki 1995 yılında uygulanan, dört yıllık aralıklarda yapılan ve 2019 yılında yedinci uygulaması gerçekteřtirilen uluslararası çapta geniř ölçekli bir arařtırmadır. TIMSS uygulamaları hem eğitim sistemlerinin uluslararası karşılařtırmalarına imkân tanımakta hem de ülkelere kendi eğitim sistemleri hakkında detaylı bilgiler vermekte ve kendi başarılarını önceki uygulamalardaki başarıları ile kıyaslama fırsatı sunmaktadır. En son gerçekteřtirilen TIMSS 2019 uygulamasında 4.sınıf düzeyinde 58 ülke ve 6 karşılařtırma katılımcısı ve 8. sınıf düzeyinde 39 ülke ve 7 karşılařtırma katılımcısı olmak üzere toplamda 64 ülke ve 8

karşılaştırma katılımcısı yer almıştır (Mullis ve diğerleri, 2020). Türkiye öğrencilerin matematik ve fen başarılarındaki eğilimlerini görmek için her iki düzeyde TIMSS 2019 uygulamasına katılmıştır (MEB, 2020).

TIMSS 2019 matematik ve fen değerlendirme çerçeveleri katılımcı ülkelerle işbirliği yapılarak içeriği belirten öğrenme alanları ve düşünme süreçlerini açıklayan bilişsel alanlar şeklinde iki boyutta tasarlanmıştır. Sınıf düzeylerine göre öğrenme alanları değişiklik gösterebilmekte, aynı öğrenme alanlarının içeriği ise sınıf düzeyine göre farklılaşmaktadır. Bilişsel alanlar her iki değerlendirme çerçevesinde de bilme, uygulama ve akıl yürütme şeklinde düzenlenmiş olup sınıf düzeylerine göre oranlarında farklılıklar vardır (Mullis ve Martin, 2017). Öte yandan TIMSS 2019 uygulamasında katılımcı ülkelere bilgisayar tabanlı değerlendirmeye (eTIMSS) geçiş fırsatı sunulmuş ve katılımcı ülkelerin yaklaşık yarısı eTIMSS'e geçmiştir. Türkiye, bilgisayar tabanlı değerlendirmeye geçiş yapan ülkelerden biridir (Mullis ve diğerleri, 2020). Bunun yanı sıra TIMSS 2019 uygulamasında ilk kez tanıtılan, yeni bir girişim olan Problem Çözme ve Sorgulama (Problem Solving and Inquiry-PSI) görevleri (tasks) yer almış ve bu görevler doğası gereği bilgisayar tabanlı olduğundan eTIMSS değerlendirmesinde yer almıştır. PSI'nın kâğıt-kalem karşılığı olmadığı için TIMSS 2019 uygulamasında ana raporlama ölçeklerine dâhil edilmemiştir. Bunun yerine sonuçları ayrı raporlanmıştır. Hiçbir öğrenci hem eTIMSS hem de PSI maddelerini almamıştır (Foy ve diğerleri, 2020).

TIMSS'e katılan dördüncü ve sekizinci sınıf düzeylerindeki öğrenciler matematik ve fen öğrenmeye yönelik deneyimleri, öğretimi ve tutumları hakkında; öğrencilerin öğretmenleri ve okul müdürleri okul, sınıf kaynakları ve yaklaşımları hakkında; dördüncü sınıf öğrencilerinin ebeveynleri öğrencilerin öğrenimi ve evdeki kaynaklar hakkında çeşitli anketler doldurmaktadır. Elde edilen anket verileri ev, okul ve sınıf hakkında zengin bilgiler sunmaktadır (Mullis & Martin, 2017). Böylelikle öğrenci başarıları, çeşitli değişkenler açısından bağlamsal olarak değerlendirilmekte ve tedbir alınması gereken eğitim politika alanları anlaşılabilir.

Geniş ölçekleri sınavların eski tarihli uygulama sonuçları üzerinden yapılan bazı araştırmalar dil, değer, düşünme yapısı, inanç, eğitim sistemi vb. bakımlardan farklı olan ülkeler için geniş ölçekli sınavların değişmezlik özelliğini sağlamadığını ya da belirli düzeylerde sağladığını göstermektedir (Hansson & Gustafsson, 2013; He ve diğerleri, 2018; Öncü, 2019; Polat, 2019; Tavlıca, 2019; Wu ve diğerleri, 2007). Bu çerçevede görece yakın bir zamanda sonuçları açıklanan TIMSS 2019 uygulamasına yönelik karşılaştırmaların yapılabilmesi için “ölçme değişmezliği” çalışmasının yapılması gerekliliği ortaya çıkmaktadır. Alan yazında çoktan seçmeli ve açık uçlu maddeleri içeren başarı testlerine yönelik değişmezlik çalışmalarının nispeten az olması sebebiyle bu araştırmada her iki soru türü de dâhil edilerek matematik başarı testinin ülkelere göre ölçme değişmezliği ele alınmıştır.

Öğrencilerin duyuşsal özelliklerinin onların öğrenme süreçlerini olumlu etkileyeceğinden ve başarılı olmalarına katkı sağlayacağından hareketle alan yazında geniş ölçekli sınavlarda kullanılan öğrencilerin ilgi, tutum gibi duyuşsal özellikleri ile ilgili olan araştırmalar da vardır (Atılğan, 2022; Demir, 2020; Lay ve diğerleri, 2013; Polat, 2019; Uyar & Doğan, 2014). TIMSS uygulamalarında da öğrencilerin duyuşsal özelliklerinden olan matematik başarısına yönelik tutumları 1995 yılından beri ölçülmektedir (Mullis ve diğerleri, 2017). TIMSS 2019 uygulamasında öğrencilerin tutumları “Matematik Öğrenmeyi Sevme (Students Like Learning Mathematics)”, “Matematikte Kendine Güven (Students Confident in Mathematics)” ve “Matematiğe Değer Verme (Students Value Mathematics)” şeklinde bir dizi ölçek aracılığıyla ölçülmüştür. Matematik Öğrenmeyi Sevme ölçeği ile öğrencilerin konuları öğrenmeye yönelik içsel motivasyonu ölçülmekte ve matematik öğrenmeye içsel olarak motive olan öğrencilerin konuyu ilginç ve eğlenceli buldukları ifade edilmektedir. Matematiğe Değer Verme ölçeği ise dışsal motivasyonu ölçmektedir. Dışsal motivasyonun övgü, kariyer başarısı, para ve diğer teşvikler gibi dışsal ödüllerden gelen güdüyü ifade ettiği belirtilmektedir. Matematikte Kendine Güven ölçeği ise konuya özgü benlik kavramını ölçmektedir. Öğrencilerin farklı konulardaki yeteneklerine ilişkin farklı görüşlere sahip olma eğiliminde oldukları ve öz değerlendirmelerinin genellikle geçmiş deneyimlerine ve bu deneyimlerin nasıl sonuçlandığına dayandığı belirtilmektedir. TIMSS verileri, bu ölçekler ile

öğrenci başarısı arasında güçlü bir ilişki olduğunu göstermekte ve TIMSS raporlarında öğrencilerin kendilerine yönelik farkındalık kazanmaları ile olumsuz tutumlarının iyileştirilebileceği açıklanmaktadır (Hooper ve diğerleri, 2017). Bu nedenle öğrencilerin duyuşsal özelliklerinden olan tutumlarına yönelik ölçeklerin ölçme değişmezliklerinin test edilmesinin önemli olduğu düşünülmektedir.

Alan yazında ülkelere göre yapılan ölçme değişmezliği çalışmalarında ya sadece başarı testinin ya da sadece duyuşsal özellikler kapsamında oluşturulan modellerin test edildiği görülmektedir. Bu noktada seçilen ülkeler için hem başarı testi hem de duyuşsal özelliklerle ilgili oluşturulan modelin ölçme değişmezliği sonuçlarının benzer mi yoksa farklı mı olacağını ortaya koymanın da gerekliliği anlaşılmıştır. Ayrıca alan yazında duyuşsal özelliklerle ilgili ölçme değişmezliği incelenen araştırmalarda birkaç ölçeğin bir araya getirilerek ölçme değişmezliğinin daha sık test edildiği görülmektedir (Başusta & Gelbal, 2015; Demir, 2020; Kıbrıslıoğlu, 2015; Polat, 2019). Bu tarz araştırmalarda her bir ölçeğin farklı bir boyuta yükleneceği beklenmektedir. Oysaki TIMSS 2019 teknik raporunda duyuşsal özelliklerle ilgili analizlerin ölçek bazında yapıldığı görülmektedir. Her bir ölçeğin Cronbach's Alpha Güvenilirlik Katsayıları, Temel Bileşenler Analizi gibi analiz yöntemleri ölçek bazında uygulanmış ve bu analizlere dayalı olarak tüm katılımcı ülkelerle ilgili sonuçlar sunulmuştur. Buradan hareketle yapılan bu araştırmada her bir ölçekten elde edilen ülkeler bazındaki ölçme değişmezliği değerleri ile üç ölçeğin bir arada bulunduğu modelden elde edilen ölçme değişmezliği değerlerinin kıyaslama ihtiyacı duyulmuş ve modelde yer alan her bir ölçek için seçilen ülkeler bazında ölçme değişmezliği de test edilmiştir. Sonuç olarak bu araştırma sayesinde ölçeklerin bir arada ya da ayrı ayrı test edilmeleri durumunda elde edilen değerlerin karşılaştırılmasına imkân sunulmaktadır.

Araştırmanın Amacı ve Önemi

TIMSS uygulamalarının asıl odağı öğrencilerin matematik ve fen alanlarında sahip oldukları bilgi ve becerileri ortaya koymak ve ülkelere bu alanlardaki eğitimlerini geliştirmelerine ve başarılarını yordamalarına destek olmaktır. Bu amaçla uygulamalarda

başarı testleri yer alır. Başarı testleri ile öğrencilerin performansları izlenmektedir. Elde edilen sonuçlar, öğrencilerin güçlü yönlerini daha da geliştirmede ve zayıf yönlerinin tespit edilmesinde kullanılmaktadır. Ayrıca TIMSS uygulamalarında çeşitli bağlam anketleri (context questionnaire) yer almaktadır. Bunlar sayesinde öğrencilerin matematik ve fen öğrenimleriyle ilgili ev, okul ve sınıf bağlamlarında çeşitli bilgilere ulaşılmaktadır. TIMSS uygulamaları sayesinde ülkelerin eğitim sistemlerinin yapısı ile matematik ve fen öğretim programları hakkında zengin bilgiler elde edilmektedir. TIMSS uygulamaları, matematik ve fen öğretimini ve öğrenimini geliştirmeye yardımcı olmak için çeşitli bilgiler sunmaktadır. TIMSS sonuçları, ülkelerin eğitim sistemlerindeki güçlü ve zayıf yönlerin belirlenmesine imkân sağlamaktadır. Öte yandan TIMSS farklı ülkeler arasında işbirliği yapmak ve bilgi paylaşmak açısından önemli görülmektedir. TIMSS, farklı ülkelerdeki öğrencilerin matematik ve fen alanlarındaki başarılarını ayrıca çeşitli özelliklerini kıyaslamak için ölçme araçlarının kullanıldığı uluslararası bir uygulamadır.

Farklı ülkeler arasında başarı testleri ve bağlam anketi ölçeklerinin (context questionnaire scales) karşılaştırılabilmesi için ölçme değişmezliği test edilmelidir. Ölçme değişmezliği, bir ölçme aracının farklı gruplarda veya zamanlarda kullanıldığında sonuçların aynı şekilde yorumlanabileceği anlamına gelmektedir. Başarı testleri ve ölçeklerin ölçme değişmezliği açısından test edilmesi, sonuçların ülkeler arasında farklı yorumlanmasını önleyerek doğru karşılaştırmalar yapılmasına olanak sağlamaktadır. Eğer bir ölçme aracı ölçme değişmezliği özelliklerine sahipse, farklı ülkelerde kullanılsa bile sonuçlar tutarlı olacaktır. Bu nedenle ölçme değişmezliği çalışmaları, başarı testleri ve ölçekler gibi ölçme araçlarının farklı ülkeler arasında kullanıldığı durumlarda sonuçların güvenilir ve geçerli, yorumların isabetli olmasını sağlamaktadır. Başarı testleri ve ölçeklerin ölçme değişmezliği test edilmeden kullanılması ise sonuçların yanlış yorumlanmasına neden olabilmekte ve ölçme sonuçlarının güvenilirliğini azaltabilmektedir. Bu nedenle bilimsel araştırmalarda ölçme değişmezliği testlerinin yapılması, değişmezlik kanıtlarının toplanması önemlidir.

Uluslararası karşılaştırmalarda başarı testlerinin ülkelere göre ölçme değişmezliğinin test edilmesi çeşitli araştırmalara (Asar, 2019; Ayvalı, 2016; Kankaras & Moors, 2014;

Karakoç Alatlı, 2016; Öncü, 2019; Tavlıca, 2019; Wu ve diğerleri, 2007) konu olmuştur. TIMSS uygulamalarında başarı testlerinin yanı sıra ölçeklere yönelik de ölçme değişmezliği çalışmaları yürütülmektedir. Bu araştırmalarda özellikle öğrencilerin duyuşsal özellikleri ile ilgili olan ölçekler için ölçme değişmezliği test edilmiştir (Demir, 2017; Kıbrıslıoğlu, 2015; Ölçüoğlu & Çetin, 2016). Ölçekler ile ilgi, güven, değer verme gibi öğrencilerin duyuşsal özelliklerine dair veriler toplanmakta (Mullis ve diğerleri, 2020) ve bu veriler sayesinde duyuşsal özelliklere ilişkin yorumlar yapılmaktadır. Duyuşsal özelliklerin öğrenci başarısı üzerindeki etkisini (%25 açıklama düzeyi) ortaya koyan Bloom (2012) tarafından yapılan araştırma dikkate alındığında, alan yazında duyuşsal özelliklere ilişkin ölçeklerin tercih edilmesi anlaşılmaktadır. TIMSS uygulamalarında duyuşsal özelliklerle ilgili ölçekler birçok araştırmada (Çakıcı Eser, 2021; Demir, 2020; Doğan & Barış, 2010; Ertürk & Erdinç Akan, 2018; Lay ve diğerleri, 2013; Ölçüoğlu & Çetin, 2016; Polat, 2019) kullanılmıştır. Duyuşsal özelliklerin öğrenci başarısını etkilemesinden dolayı öğrencilerin duyuşsal özellikleri ile ilgili olan ölçeklerin araştırma kapsamında ele alınması önemli görülmektedir. Araştırmalarda test edilen başarı testi ya da ölçeklerin sağladığı ölçme değişmezliği aşaması farklılık gösterebilmektedir.

Bu araştırmada, 8.sınıf matematik başarı testinin ve duyuşsal özelliklerle ilişkili “Matematik Öğrenmeyi Sevme”, “Matematikte Kendine Güven” ve “Matematiğe Değer Verme” ölçeklerinden oluşturulan modelin ve ayrı ayrı her bir ölçeğin ülkelere göre ölçme değişmezliğinin irdelenmesi amaçlanmıştır. TIMSS teknik raporunda öğrencilerin duyuşsal özelliklerine ilişkin ölçekler belli başlıklar altında sunulsa da ölçeklere ilişkin analizler ayrı ayrı yapılmaktadır. Örneğin araştırma kapsamında alınan “Matematik Öğrenmeyi Sevme”, “Matematikte Kendine Güven” ve “Matematiğe Değer Verme” ölçekleri öğrenci tutumları başlığı altında yer almasına rağmen analizler ölçek bazında yapılarak tüm katılımcı ülkelerin verileri sunulmuştur. Bu nedenle araştırma kapsamında her bir ölçeğin ülkelere göre ölçme değişmezliğinin incelenmesi de önemli görülmüştür.

Araştırmanın yapılacağı ülkeler farklı dilleri konuşan, toplumlarına uygun farklı tipik davranış örnekleri olan, farklı düşünme yapılarına sahip kısaca kültürel öğelerinde farklılık

gösteren ve dolayısıyla farklı yaşam biçimlerine sahip öte yandan farklı gelir düzeylerindeki Türkiye, Güney Kore ve Finlandiya olarak seçilmiştir. Araştırma kapsamında ele alınan matematik başarı testinin, oluşturulacak duyuşsal özellik modelinin ve ölçeklerin ülkelere göre eşdeğer olması beklenmektedir. Aksi takdirde ülkelere göre değişkenlik tespit edildiğinde ülkeler arasında sınırlı düzeyde karşılaştırma yapmak mümkün olacak, kapsamlı karşılaştırmalar doğru olmayacak, kullanılan ölçme araçlarının geçerliği değişmezlik özelliği bakımından sağlanamayacaktır. Dolayısıyla güvenilirliği ve geçerliği birçok yönden ispat edilmemiş ölçme araçlarından elde edilen sonuçlara göre yapılan yorumlar yanlış olabilecektir (Vandenberg & Lance, 2000). Başarı testinin ve oluşturulacak duyuşsal özellik modelinin/ölçeklerinin ülkelere göre ölçme değişmezliği sağlandığında ise ülkeler arasında yapılacak karşılaştırmalar anlamlı ve doğru olacaktır. Bu durum öğrencilerin hem performanslarına hem de duyuşsal özelliklerine yönelik karşılaştırma yapma imkânı sunacaktır. Diğer bir deyişle aynı ülkelerde farklı ölçme araçlarının karşılaştırılması sayesinde ülkelerin eğitim sistemlerinin farklı yönlerinin anlaşılması için önemli veriler sağlanacaktır. Böylelikle ülkelerin eğitim sistemlerinin nasıl işlediği hakkında daha fazla bilgi edinilecek ve eğitim politikalarının, sistemlerinin, eğitim sisteminin önemli bir bileşeni olan öğretim programlarının geliştirilmesi için faydalı bilgiler elde edilebilir. Öte yandan öğrencilerin küresel çapta rekabet edebilme durumlarına ilişkin bilgi elde edilecek ve ülkelerin gelecekteki potansiyelleri öngörülebilecektir.

Araştırma Problemi

TIMSS 2019 uygulamasında yer alan 8.sınıf matematik başarı testinin, duyuşsal özelliklerle ilgili oluşturulan üç ölçeğin bir arada bulunduğu modelin ve ölçeklerin her biri için ülkeler (Türkiye, Güney Kore ve Finlandiya) arasında ölçme değişmezliği sağlanmakta mıdır?

Bu kapsamda aşağıda yer alan alt problemler yanıtlanacaktır:

Alt Problemler

1. TIMSS 2019 uygulamasında yer alan 8.sınıf matematik başarı testinin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği sağlanmakta mıdır?

2. TIMSS 2019 uygulamasında yer alan ve 8.sınıf öğrenci anketi kapsamında öğrencilerin matematik alanındaki duyuşsal özelliklerine yönelik hazırlanan üç ölçeğin bir arada bulunduğu modelin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği sağlanmakta mıdır?

3. TIMSS 2019 uygulamasında yer alan ve 8.sınıf öğrenci anketi kapsamında öğrencilerin matematik alanındaki duyuşsal özelliklerine yönelik hazırlanan ölçeklerin her biri için Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği sağlanmakta mıdır?

Sınırlılıklar

Araştırmada TIMSS 2019 uygulamasından ulaşılan Türkiye, Güney Kore ve Finlandiya ülkelerine ait 8.sınıf düzeyindeki öğrencilerin 14 no'lu kitapçıktaki matematik başarı testine ve duyuşsal özellikler ile ilgili seçilen “Matematik Öğrenmeyi Sevme”, “Matematikte Kendine Güven” ve “Matematiğe Değer Verme” ölçeklerine ait verileri kullanılarak sonuçlar üç ülkenin verisiyle sınırlandırılmıştır.

Başarı testinin ölçme değişmezliği analizinde metrik değişmezlik aşamasının yapılmaması araştırmanın sınırlılığı olarak belirtilebilir.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Kültürün en kapsayıcı tanımlarından biri "onun bir yaşam biçimi" oluşudur (Varış, 1994). Kültürün kuşaktan kuşağa aktarılması eğitim yoluyla sağlanmaktadır. Ültanır'a göre (2003) "dil, hoşgörü ve kendini tanıma, değerler, topluma uygun tipik davranış örnekleri, düşünme yapısı, teknoloji, ekonomi, sanat, spor ve oyunlar, din" şeklindeki kültür öğeleri eğitimle aktarılmalıdır. Buradan hareketle ülkelerin eğitim sistemlerinin bu kültür öğelerini yansıttığı ve böylelikle birbirinden farklılaştığı söylenebilir. Dolayısıyla uluslararası eğitim araştırmalarında ülkeler karşılaştırıldığında ülkelerin kendine has özelliklerinin etkileri bertaraf edilmelidir. Bu tesir ortadan kaldırılmadığında ulaşılan sonuçlar üzerinden yapılan yorumlar doğru olmayabilir. Uluslararası araştırmalarda yer alan gerek başarı testleri gerekse ölçeklerin farklı ülkelerdeki öğrenci grupları için aynı anlama gelip gelmediği belirlenmelidir. Bu noktada söz konusu test, ölçek veya oluşturulacak modellerin farklı ülkelerdeki öğrenci gruplarına göre ölçme değişmezliği test edilerek geçerliği ortaya koyulmalıdır. Aşağıda sırasıyla ölçme değişmezliği, yapısal eşitlik modeli, doğrulayıcı faktör analizi, çok gruplu doğrulayıcı faktör analizi başlıkları altında kuramsal temel verilmiştir. Ayrıca ilgili alan yazın sunulmuştur.

Ölçme Değişmezliği

Araştırmalarda ele alınan değişkenler doğrudan ölçülemeyen gizil yapılar olduğunda, gruplar arasında karşılaştırmalar yapmak için gizil değişkenlerin ortalama farklılıklarının kullanılması önerilmiştir (Hong ve diğerleri, 2003; Kuhn & Holling, 2009). Karşılaştırmalı araştırmalar genellikle nicel yöntemler kullanılarak yapıldığından, kullanılan ölçme araçlarının grup karşılaştırmaları için geçerli bir temel sağlayıp sağlamadığı konusunda endişeler ortaya çıkabilir (Gregorich, 2006). Gruplar arasındaki gizil değişkenlerin ortalama farklarını karşılaştırmadan önce, ölçme değişmezliğini test etmek önemli bir ön koşuldur (Chen, 2007; Cheung & Rensvold, 2002). Ölçme aracının farklı gruplardaki bireyler tarafından aynı

biçimde anlaşılıp anlaşılmadığı, farklı gruplar arasında yapının eş olup olmadığını belirlemek ölçme değişmezliği analizleri ile mümkün olmaktadır (van de Schoot, ve diğerleri, 2012). Yani belirli bir özellik bakımından farklı gruplarda yer alan aynı seviyedeki bireylerin ölçme aracındaki maddeleri aynı biçimde kavraması ve değerlendirmesi gerekir (Bryne & Watkins, 2003). Ölçme aracının gruplarda aynı şeyi ifade etmediği ortaya koyulursa gruplarda farklı bir psikolojik özelliğin ölçüldüğüne dair fikir oluşur ve her bir gruptan ulaşılan sonucun aynı anlamda olmadığı ifade edilir. Bu durumda gruplar arasında karşılaştırma yapılamaz ve dolayısıyla gruplar arasındaki fark yorumlanamaz (Cheung & Rensvold, 2002).

Uluslararası araştırmalarda gruplar arası karşılaştırmalar yapılmaktadır. Karşılaştırmalarda gruplar arasında farklılıklar tespit edilebilir. Ulaşılan bu farklılıklar gruplardaki bireylerin özelliklerinden kaynaklanmayabilir. Bu farklılıklar ölçme aracının kendisinden de kaynaklanabilir (Başusta & Gelbal, 2015). Ölçme aracı gruplardan herhangi birinin lehine ve buna bağlı olarak yanlılık oluşmasına neden olabilir. Dolayısıyla farklılıklara dayalı olarak yapılacak yorumların temeli olmalıdır. Aksi halde yapılacak yorumlar yanlış olabilir. Nitekim ölçme sonuçlarının yorumlarının geçerli olması ve dolayısıyla karşılaştırmaların doğru olması için ölçme değişmezliğinin sağlanması gerekmektedir.

Meredith (1993), ölçme değişmezliğinin gösterilmesi için Yapısal Değişmezlik (Configural Invariance), Metrik Değişmezlik (Metric Invariance), Ölçek Değişmezliği (Scalar Invariance) ve Katı Değişmezlik (Strict Invariance) olarak dört aşamalı mantıksal bir sürecin izlenmesini ve kurulan hipotezlerin test edilmesini önermektedir.

Yapısal Değişmezlik

Yapısal değişmezlik aşamasında temel olarak, aynı faktör yapısının gruplar için geçerli olup olmadığı test edilir. Farklı gruplardan bireylerin maddeleri yanıtlamak için aynı kavramsal çerçeveyi kullanıp kullanmadığı araştırılır (Cheung & Rensvold, 2002). Bu değişmezlik aşaması yalnızca genel yapıyı inceler. Her bir grup için aynı sayıda faktörün olup olmadığı ve her bir faktörde aynı maddelerin yer alıp almadığı değerlendirilir. Yapısal değişmezlik sağlanırsa ölçme aracının gruplar için aynı faktör yapısını ölçtüğü anlaşılır.

Gruplar için faktör sayısı ve her bir faktördeki madde sayısı eşit olur (Bryne, 2008). Bu aşamanın sağlanması grupların karşılaştırılabilir olduğunu ifade etmez, fakat diğer aşamaları sınamak için bu aşamanın sağlanması ön koşul niteliğindedir. Bu değişmezliğin sağlanamaması ise gruplar arasında farklı yapıların ölçüldüğü anlamındadır ve diğer aşamaların test edilmesi gerekli ve uygun değildir (Cheung & Rensvold, 2002).

Metrik Değişmezlik

Metrik değişmezlik aşamasında faktör yüklerinin gruptaki değişmezliği test edilir. Yani faktörler ve maddeler arasındaki ilişkilerin gruplar arasında ne ölçüde eşdeğer olduğu incelenir. Bu aşama sağlanıyor ise maddelerin her bir gruptaki bireyler tarafından aynı şekilde anlaşıldığı ve yorumlandığı belirtilir (Kline, 2016). Metrik değişmezlik sağlanmıyor ise bir veya daha fazla gruptaki bireylerin bazı maddeleri farklı anlamlandırdıkları söylenebilir (Gregorich, 2006).

Ölçek Değişmezliği

Bu aşamada metrik değişmezlik aşamasında test edilen faktör yüklerine ilaveten madde sabitlerinin gruptaki eşitliği test edilir. Gruplar arasındaki ortalama farkları yorumlamadan önce bu aşamanın karşılanması zorunlu olduğundan ölçek değişmezliği önemlidir (Meredith, 1993). Ölçek değişmezliğinin sağlanması, gizil yapıda aynı puanı alan bireylerin buldukları gruptan bağımsız olarak gözlenen değişkende aynı puana ulaşmaları anlamındadır (Milfont & Fischer, 2010). Bu durumda gözlenen değişkenlerdeki ortalama farklılıkların, gizil yapıların ortalamalarından meydana geldiği ve gruptan herhangi biri(leri) lehine yanlılık olmadığı dile getirilebilir (Başusta & Gelbal, 2015). Bu değişmezlik aşamasının belirlenmesi geçerlik anlamında çok güçlü kanıt olarak kaydedilir (Vandenberg & Lance, 2000). Ölçek değişmezliği desteklenmediğinde ise, gruplar arası karşılaştırmalara ilişkin çıkarımların geçerliği sorgulanır (Cheung & Rensvold, 2002). Bu durumda değişen madde fonksiyonu şüphesi ortaya çıkar ve madde yanlılığı ile ilgili incelemeler yapılması gerekir.

Katı Değişmezlik

Katı değişmezlik aşamasında önceki aşamaya ek olarak hata varyanslarının gruplar arasında eşitliği sınanır. Bu aşamanın karşılanmasının uygulamada oldukça güç olduğu belirtilmektedir (Cheung & Rensvold, 2002). Alan yazında bu aşamanın gerekliliği konusunda farklı fikirler vardır. Brown (2006) bu aşamanın değerlendirilmesini araştırmacının takdirine bırakırken Meredith (1993) gruplar arası karşılaştırma yapmak için bu aşamanın sağlanmasının önemli olduğunu belirtmiştir. Katı değişmezliğin sağlanması gruplar arasında gizil ve gözlenen değişkenlerin ortalamalarının karşılaştırılmasına ek olarak gözlenen varyans ve kovaryansların da karşılaştırılabilmesi anlamındadır (Gregorich, 2006).

Yukarıda açıklandığı üzere ölçme değişmezliğinde test edilen her aşama, bir önceki aşamaya göre yeni kısıtlamalar gerektirir. Yani değişmezlik aşamaları giderek kısıtlanmaktadır. Tablo 1’de ölçme değişmezliği aşamalarına göre yapılan kısıtlamalar sunulmuştur:

Tablo 1

Ölçme Değişmezliği Aşamalarına Göre Yapılan Kısıtlamalar

Değişmezlik Aşamaları	Kısıtlamalar
Yapısal Değişmezlik	Kısıtlama Yok
Metrik Değişmezlik	Faktör Yükleri
Ölçek Değişmezliği	Faktör Yükleri + Madde Sabitleri
Katı Değişmezlik	Faktör Yükleri + Madde Sabitleri + Hata Varyansları

Ölçme değişmezliği popülasyonlar arası (ülkeler arası vb.) değerlendirme için önemli bir araştırma konusu olarak tasarlanmıştır (Cheung & Rensvold, 2002; Kline, 2016). Ölçme değişmezliğinin test edilmesi için sıklıkla iki analiz yöntemi kullanılmaktadır. Bu yöntemlerden biri Madde Tepi Kuramına (MTK) dayalı iken diğeri Yapısal Eşitlik Modeline (YEM) dayalıdır. Madde tepki kuramında doğrusal olmayan yöntem olan madde ve test işlev farklılıkları analizi (Değişen Madde Fonksiyonu - Değişen Test Fonksiyonu) yapılırken yapısal eşitlik modelinde doğrusal bir yöntem olan doğrulayıcı faktör analizi kullanılmaktadır (Raju ve diğerleri, 2002).

Alan yazında ölçme değişmezliğini inceleyen araştırmalarda YEM kapsamında yer alan Çok Gruplu Doğrulayıcı Faktör Analizi (ÇGDFA) sıklıkla tercih edilmektedir (Atılğan, 2022; Başusta & Gelbal, 2015; Çakıcı Eser, 2021; Demir, 2020; Gönen, 2021; Güngör & Atalay Kabasakal, 2020; Ölçüoğlu & Çetin, 2016; Öncü, 2019; Polat, 2019; Tekin & Çobanoğlu Aktan, 2021; Uyar & Doğan, 2014; Ülkü, 2019). Ölçme değişmezliği bu araştırmada YEM'e dayalı ÇGDFA yapılarak inceleneceğinden YEM'e ilişkin bilgiler ayrıca sunulmuştur.

YEM'e yönelik analizleri yapmadan önce eldeki verilerin analizlere uygunluğunu tespit etmek için verilerin bazı gereklilikleri sağlayıp sağlamadığı kontrol edilmelidir. Bu kapsamda Çokluk vd. (2010) tarafından örneklem büyüklüğü, kayıp veriler, uç değerler, normallik ve çoklu bağlantıya ilişkin hipotezlerin test edilerek doğrulanması gerektiği belirtilmiştir.

Örneklem Büyüklüğü

Bir araştırmada doğru sonuçlara ulaşmak için belirli bir sayıda örneklemin alınması anlamındadır. Bu sayı genellikle araştırmanın türüne, popülasyonun büyüklüğüne, örnekleme yöntemine ve istatistiksel analizlerin hassasiyetine bağlı olarak değişebilir. Genel olarak örneklem ne kadar büyük olursa sonuçlar o kadar doğru ve güvenilir olur. Ancak daha büyük bir örneklem daha fazla kaynak, zaman ve çaba gerektirir. Bu nedenle araştırmalarda güvenilir şekilde tahminler yapabilmek için yeterli büyüklükte örnekleme çalışmak önemlidir (Creswell & Plano Clark, 2014). Alan yazında örneklem büyüklüğünün yeterliğine ilişkin farklı görüşler vardır. Kline (1994) örneklem büyüklüğü için ölçme aracında yer alan madde sayısının 10 katı kadar kişinin alınmasını yeterli görürken Bryman ve Cramer (2001) madde sayısının 5 ya da 10 katının yeterli olacağını belirtmektedir. Örneklem büyüklüğünün yeterliğini yorumlamak için Kaiser-Meyer-Olkin testinin sonuçlarına da bakılabilir.

Kayıp Veriler

Bir araştırmada veri setinin eksiksiz olması oldukça güçtür. Verilerin eksik olması kayıp veri olarak ifade edilmektedir. Araştırmalarda kayıp veri sorununun görülmesi olası bir durumdur. Gözlenen değişkenlerde yaşanan kayıplar, gizil değişkenlerle ilgili yapılan kestirimlerin doğruluğunu etkileyebilir ve araştırma sonuçlarını yanıltıcı hale getirebilir (Hohensinn & Kubinger, 2011). Bu nedenle söz konusu sorun araştırmanın başında öncelikle

engellenmeli ya da karşılaşıldığında uygun analiz yöntemi ile giderilmelidir. Kayıp verilerin analiz dışı bırakılması verilerin ihmal edilebilir olma durumlarına bağlıdır. Uygun kararlar alabilmek için kayıp verinin yapısı ve mekanizması incelenmelidir.

Uç Değerler

Uç değerler veri setinde diğer verilere kıyasla oldukça farklı olan ve genellikle diğer verilerle uyuşmayan, istatistiksel analizlerde yanıltıcı sonuçlara neden olabilen değerlerdir. Uç değerler, bir veri setinde yüksek veya düşük bir değer olarak ortaya çıkabilir. Analizlerde uç değerlerin varlığı sonuçların doğruluğunu etkileyeceğinden uç değerler belirlenmelidir. Uç değer sorunu araştırmacının veri girişinden kaynaklanabileceği gibi, örneklem içindeki oldukça farklı bir durumdan da oluşabilir (Çokluk ve diğerleri, 2010).

Normallik

Verilerin normalliği tek ve çok değişkenli olarak ele alınır. Tek değişkenli normallik, bir değişkenin normal dağılım gösterip göstermediğini analiz etmektir. Normal dağılım, çarpıklık ve basıklık katsayılarına bakılarak değerlendirilebilir. Normal dağılımda çarpıklık ve basıklık katsayıları sıfır olmalıdır. Fakat bu katsayılar genellikle sıfırdan farklı olduğundan bu değerler için alan yazında kabul edilebilir farklı aralıklar belirtilmiştir. Örneğin George ve Mallery (2010) bu aralıkların ± 2 ; Tabachnick ve Fidell (2013) $\pm 1,5$ olduğunu kabul etmiştir. Ayrıca normallik çarpıklık ve basıklık katsayılarının standart z-skorları, histogram grafikleri ve normallik testleri ile kontrol edilebilir. Normallik kontrolünde örneklem büyüklüğü önemlidir. Demir (2022) simülatif olarak elde edilen farklı basıklık ve çarpıklık katsayıları altında normal dağılıma sahip verilerle farklı örneklem büyüklüklerindeki normallik testlerini karşılaştırdığı araştırmasında çarpıklık ve basıklık katsayılarının sıfıra eşit veya sıfıra yakın olduğu durumlarda normallik testlerinin örneklem büyüklüğünden etkilenmediği; fakat çarpıklık ve basıklık katsayılarının sıfırdan uzaklaştığı durumlarda normallik testlerinin örneklem büyüklüğünden etkilendiği ve bu tür testlerin anlamlı sonuçlar verme eğiliminde olduğunu tespit etmiştir.

Çok değişkenli normallik ise birden fazla değişkenin normal dağılım gösterip göstermediğini analiz etmektir. Çok değişkenli normallik, çok değişkenli istatistik

yöntemlerinin uygulanmasında önemli bir önkoşuldur. Çok değişkenli normallik, her bir değişkenin tek değişkenliği normalliği sağlaması, her bir değişkene ait saçılım grafiklerinin doğrusal olması, artıkların eşvaryanslı olması ile mümkündür (Kline, 2016). Çok değişkenli normallik mahalnobis uzaklıkları yardımıyla saptanabilir.

Çoklu Bağlantı

Çoklu bağlantı, bağımsız değişkenler arasında yüksek korelasyonlar olduğunda ortaya çıkan istatistiksel bir problemdir. Başka bir deyişle çoklu bağlantı, iki veya daha fazla bağımsız değişken birbiriyle yüksek oranda doğrusal olarak ilişkili olduğunda ortaya çıkar ve her bir değişkenin bağımlı değişken üzerindeki bağımsız etkisini belirlemek zorlaşır. Değişkenler arasında 0.90'dan büyük korelasyon değeri elde edilmesi çoklu bağlantı sorununu ifade eder. Bu duruma neden olan değişkenlerin veri setinden silinebileceği belirtilmiştir (Tabachnick & Fidell, 2013).

Yapısal Eşitlik Modeli

Davranış bilimlerinde sıklıkla tercih edilen Yapısal Eşitlik Modellemesi (YEM), faktör analizi, regresyon veya yol analizinin bir birleşimi olarak ele alınabilir. Ayrıca YEM çeşitli değişkenler arasındaki karmaşık ilişkileri modellemek için kullanılan çok değişkenli regresyon analizlerine kıyasla daha esnek bir modeldir. YEM gizil faktörlerle ilişkilendirilen teorik yapıları temsil eder. Teorik yapılar arasındaki ilişkiler, faktörler arasındaki regresyon veya yol katsayıları ile temsil edilir (Hox & Bechger, 1999). YEM, gözlenen veya gizil değişkenler arasındaki nedensel ilişkilere dair kurulan bir hipotezi değerlendirmek için istatistiksel bir araçtır (Mueller & Hancock, 2018). Bu model, karmaşık verileri modellemeye ve özellikle neden-sonuç ilişkileri hakkında daha derinlemesine anlayışlar sağlamaya yardımcı olur.

Gizil değişkenler arasındaki ilişkileri değerlendirmek ve teorik modelleri test etmek amacıyla eğitim, psikoloji, sosyoloji, pazarlama alanlarında YEM'den yararlanır. YEM ölçme aracının yapı geçerliğine dair kanıt oluşturmak için aracın psikometrik özelliklerini incelerken, ölçme değişmezliğini test ederken, gözlenen değişkenlerdeki hataları dikkate alırken, çok

değişkenli modelleri analiz ederken, doğrudan ve dolaylı etkileri tahmin ederken kullanılır (Şen, 2020).

YEM kullanmanın korelasyon, çoklu regresyon, varyans analizi gibi sık kullanılan nicel yöntemlere göre bazı avantajları vardır. YEM'in değişkenler arasındaki ilişkiyi keşfetme kapasitesi diğer nicel yöntemlere göre avantajlarından biridir (Weston & Gore, 2006). Çok sayıda hipotezi aynı anda bir bütün olarak içeren karmaşık ilişkileri modellemeye ve test etmeye izin verir. YEM, yapıları temsil etmek amacıyla birden fazla ölçüm yapılmasına imkân tanır ve ölçmeye özgü hatayı giderir. Ayrıca ölme hataları ile gözlenen değişkenlerdeki hatalar arasındaki ilişkiyi ele alma kapasitesi YEM'in bir diğer avantajıdır (Şen, 2020). Diğer analiz yöntemlerinin kullanılması ise genellikle daha az geçerli sonuçlar sağlar veya birkaç ayrı analiz gerektirir.

Öte yandan YEM'in karmaşıklığı, istatistiksel ve yorumsal zorlukları, potansiyel sorunları beraberinde getirir. Örneğin genel koşula rağmen modelin belirli bölümleri tanımlanmayabilir. Bu durumda modeli yeniden formüle etmek, ek değişkenleri dâhil etmek veya tanımlanmış model parçalarını ayrı ayrı test etmek gerekebilir. Kestirim yönteminin varsayımlarının ihlal edildiği durumlarda veya analiz edilen modelin yanlış varsayımlara veya hipotezlere (yanlış tanımlanmış model) dayandığı durumlarda kestirme problemleri ortaya çıkabilir. Bu durumda ise farklı bir kestirme yöntemi (örneğin, daha az katı varsayımlara sahip bir yöntemin) kullanılabilir, model basitleştirilebilir, modelin parçalarının ayrı analizleri yapılabilir veya daha büyük bir örneklem alınabilir. Ayrıca küçük örneklem büyüklüklerini, normal olmayan verileri ve değişkenler arasındaki zayıf ampirik ilişkileri birleştirmek, kestirim sorunlarına ve güvenilir olmayan sonuçlara yol açabilir. Olası çözümler arasında belirli düzeltme faktörleri veya farklı bir kestirim yönteminin kullanımı sayılabilir. Bir diğer sorun ise sonuçların yorumlanmasına ilişkindir. Değişkenlerin nedensel ilişkileri hakkında temelde farklı varsayımlar yapan, ancak yine de tam olarak aynı model uyumuna yol açan ve yalnızca istatistiksel kriterlere dayalı bir karar vermeyi imkânsız kılan iki veya daha fazla alternatif modelin olduğu durumlarda karar sorunları ortaya çıkabilir (Bollen, 1989; Kline, 2016).

Kline (2016) YEM çalışmalarında izlenmesi gereken basamakları aşağıdaki gibi belirtmiştir:

- 1- Modelin belirlenmesi
- 2- Modelin tanımlanması
- 3- Model uyumunun kestirilmesi
- 4- Uyum istatistiklerinin değerlendirilmesi
- 5- Gerekli durumlarda modelin yeniden betimlenmesi
- 6- Sonuçların raporlanması

Modelin belirlenmesi basamağında araştırmada ele alınan değişkenlerin birbirleri ile ilişkileri tasvir edilir. Gözlenen ve gizil değişkenler arasında var olduğu veya olmadığı varsayılan ilişkiler bu aşamada belirtilir. Değişkenler arasındaki herhangi bir belirtilmemiş ilişkinin sifıra eşit olduğu varsayıldığından bu durum önemlidir (Weston & Gore, 2006).

Modelin tanımlanması basamağında bilinmeyen veya kestirilen parametreler incelenerek parametre sayıları belirlenir. Model tanımlamada sağlanması gereken bazı koşullar vardır. Bunlardan biri serbestlik derecesidir. Sıfır serbestlik derecesine sahip bir model her zaman tam olarak tanımlanır ve verilere mükemmel şekilde uyar. Sıfırdan büyük serbestlik derecesi olduğunda, model aşırı tanımlanır. Model verilere iyi uyduğunda, gözlenen ve gizil değişkenler arasındaki ilişkilerin önemli olduğu belirtilir. Serbestlik derecesi negatif olduğunda ise model eksik tanımlanır ve kestirilemez. Özetle modelin tanımlanabilmesi için serbestlik derecesi sıfır veya sıfırdan büyük olmalıdır. Ayrıca her bir gizil değişkenin ölçeklendirilmesi ve modelin boyut sayısına göre gözlenen değişken sayısı koşullarına dikkat edilmesi gereklidir (Kline, 2016).

Model uyumunun kestirilmesi basamağında bilinmeyen parametrelerin değeri ve kestirilen değerle ilişkili hata belirlenir. Model kestirimi için En Çok Olabilirlik (Maximum Likelihood), En Küçük Kareler (Least Squares), Ağırlıklandırılmamış En Küçük Kareler (Unweighted Least Squares), Genelleştirilmiş En Küçük Kareler (Generalized Least Squares) ve Asimptotik Olarak Dağılımdan Bağımsız (asymptotic distribution free) yöntem gibi çeşitli yöntemler vardır (Weston & Gore, 2006). Her kestirim yönteminin artıları ve eksileri

olduğundan analizleri gerçekleştirmeden önce hangi kestirim yönteminin kullanılacağı seçilmelidir.

Uyum istatistiklerinin değerlendirilmesi basamağında kestirilen modeldeki gözlenen ve gizil değişkenler arasındaki ilişkilerin, verilerde gözlenen ilişkileri yeterince yansıtır yansıtmadığı incelenir. İstatistikçiler, kestirilen parametrelerin önemi ve gücü, gözlenen ve gizil değişkenlerde açıklanan varyans ve çeşitli uyum indekslerinin gösterdiği gibi modelin gözlenen verilere ne kadar iyi uyduğu açısından değerlendirilmesi gerektiği konusunda hemfikirdir (Hoyle, 1995).

Uyum istatistiklerinin değerlendirilmesinin sonucunda model ile gözlenen veriler arasında kabul edilebilir düzeyde bir uyum saptanamamışsa başka bir model betimlenerek test edilir. Modelin yeniden betimlenmesinde modifikasyon indeksleri kullanılır. Basamakların tamamlanmasının ardından elde edilen sonuçlar raporlanır (Kline, 2016).

Araştırmada ölçme değişmezliği YEM'e dayalı yürütüleceğinden doğrulayıcı faktör analizi kullanılacaktır. Bu nedenle doğrulayıcı faktör analizi aşağıda açıklanmıştır.

Doğrulayıcı Faktör Analizi

Değişmezlik çalışmalarında YEM uygulamalarından biri olan Doğrulayıcı Faktör Analizi (DFA) yönteminden yararlanır. DFA gözlenen değişkenler (test maddeleri, puanlar gibi) ile gizil yapı(lar) arasındaki ilişkiyi tanımlamak ve doğrulamak için kullanılan bir yöntemdir. DFA'da faktör sayısı, gösterge-faktör yüklerinin modeli ve faktörlerin bağımsızlığı veya kovaryansı ve ayrıca göstergelere özgü varyanslar gibi parametreler önceden belirlenir (Brown & Moore, 2012). DFA önceden belirlenmiş bir modelin veriyle uyumunu inceler. Araştırmacı teorik bilgisini veya deneysel araştırmayı kullanır, hatta her ikisini bir araya getirebilir. İlişki modelini önceden varsayar, sonrasında hipotezleri istatistiksel yöntemlerle sınar.

Schumacker ve Lomax (1996) DFA kullanımının; test edilen araştırma hipotezi, örneklem büyüklüğü, ölçme araçları, çok değişkenli normallik, parametre tanımlamaları, uç

değerler, kayıp veriler ve model uyum indekslerinin yorumlanmasından etkilenebileceğini belirtmiştir.

DFA'nın temel yapısı, önceki ampirik ve teorik zeminlere dayalı olarak kurulduğunda kullanılır. DFA ile değişkenlerin tespit edilen faktörler ile yeterince iyi temsil edilip edilmedikleri ortaya koyulur (Thompson, 2004). DFA, faktör sayısı bilinen yapıya bağlı olarak önerilen modelin doğru olup olmadığını sınamak için uygulanır (Doğan, 2013). Test geliştirme ve uyarlama çalışmalarında sıklıkla kullanılan DFA, YEM çalışmalarında da sıkça tercih edilmektedir.

Araştırmacılar tarafından oluşturulan modellerin farklı gruplarda değişmezliğini ortaya koymak için DFA'dan yararlanır. Bu çalışmada da farklı ülkeler arasında değişmezlik çalışmalarını yapabilmek için Çok Gruplu Doğrulayıcı Faktör Analizi (ÇGDFA) kullanılmıştır.

Çok Gruplu Doğrulayıcı Faktör Analizi

Birden fazla örneklem grubunun faktör yapısını ve bu faktörlerin ölçümlerini karşılaştırmak için kullanılan istatistiksel bir tekniktir. ÇGDFA, DFA'ya dayanır ve DFA'dan farklı olarak birden fazla gruba çalışır. İki veya daha fazla grup arasındaki farklılıkların tespitinde kullanılan ÇGDFA, DFA'nın gruplarda aynı anda yapılmasını gerektirir. Bu analiz ile oluşturulan modelin gruplar için aynı olup olmadığı ortaya koyulur. Güçlü bir analiz olan (Gregorich, 2006) ve YEM çalışmalarında sıkça kullanılan (Tabachnick & Fidell, 2013) ÇGDFA ile farklı gruplardan sağlanan sonuçlar karşılaştırılır.

ÇGDFA ile ölçme değişmezliği birbirinin içine geçmiş dört aşamanın test edilmesi ile belirlenir. Ölçme değişmezliği "yapısal değişmezlik, metrik değişmezlik, ölçek değişmezliği ve katı değişmezlik" şeklinde sıralanan dört sıralı aşamada incelenir (Stark ve diğerleri, 2006). Model test edilirken en az koşul belirtilen aşamanın test edilmesi ile harekete geçilir. Ardından daha fazla koşulu olan aşamalar sırasıyla test edilir. Söz konusu aşamalar ölçme değişmezliği başlığı altında açıklanmıştır.

Alan yazında ölçme değişmezliği çalışmalarında çeşitli indekslerden yararlanılmaktadır. Mutlak uyum indeksi olarak kullanılan Ki-kare (Chi-square goodness of fit-

χ^2) indeksi bir modelin gözlenen verilere ne kadar iyi uyduğunu doğrudan değerlendirir ve alternatif hipotezleri (competing hypotheses) test ederken modelleri karşılaştırmada faydalıdır. Anlamlı bir χ^2 , modelin verilere uymadığını gösterir. Buna karşılık, anlamlı olmayan bir χ^2 , verilere iyi uyan bir modelin göstergesidir. En sık bildirilen mutlak uyum indeksi χ^2 olmasına rağmen, bu istatistiğin iki sınırlılığı vardır. İlk olarak bu istatistik, modelin verilere tam olarak uyup uymadığını test eder. Oysaki tam bir uyum bulmak zordur. İkincisi, çoğu istatistikte olduğu gibi, örneklem büyüklüğünün artması χ^2 'nin küçük etki büyüklükleriyle anlamlı olmasına yol açar (Henson, 2006). Sonuç olarak, model gözlenen verilere çok yakın olsa da anlamlı olmayan bir χ^2 olası olmayabilir. Bu sınırlılıklara rağmen, araştırmacılar χ^2 'yi neredeyse evrensel olarak tercih etmektedir (Martens, 2005). χ^2 istatistiğindeki örneklem büyüklüğü etkisini azaltmak amacıyla χ^2 'nin serbestlik derecesine oranı olan normlaştırılmış ki-kare (χ^2/sd) değeri kullanılmaktadır (West ve diğerleri, 2012). Model uyumunun kabul edilebilir olup olmadığını belirlemek için karşılaştırmalı uyum indeksi (Comparative Fit Index-CFI) gibi diğer uyum indeksleri de kullanılmaktadır. Bunlar arasında tahmini hataların ortalama karekökü (Root Mean Square Error of Approximation-RMSEA), standartlaştırılmış artık ortalama karekökü (Standardized Root Mean Square Residual-SRMR) indeksleri sayılabilir. Bir model uyumunun yeterliğini tek bir indeksin belirleyemeyeceği ve bu nedenle modellerin uyumunu değerlendirmek için birden fazla indeks kullanması gerektiği belirtilmektedir (Crowley & Fan, 1997). Araştırmada uyum indeksleri olarak χ^2 , χ^2/sd , RMSEA, CFI, TLI ve SRMR indeksleri raporlanmıştır. Duyuşsal özellik modelinde ve ölçeklerde örneklem büyüklüğü etkisi nedeniyle χ^2 ve χ^2/sd indeksleri değerlendirilmemiş RMSEA, CFI, TLI ve SRMR indeksleri yorumlanmıştır.

Bu araştırmada ÇGDFA ile farklı ülkelerdeki öğrenciler yani farklı gruplar arasındaki faktör yapısı ve faktörlerin ölçümlerindeki farklılıklar incelenmiştir. Böylelikle farklı grupların

benzer veya farklı faktör yapısına sahip olup olmadıkları ortaya koyularak faktör yapısının evrensel mi ya da kültürel olarak bağımlı hale mi geldiği belirlenmeye çalışılmıştır.

İlgili Araştırmalar

Bu bölümde geniş ölçekli uygulamalara yönelik hazırlanan ölçme değişmezliği ile ilgili uluslararası ve ulusal alan yazında yapılan araştırmalar sunulmuştur. Bu kapsamda başarı testleri (okuryazarlık testleri, beceri testleri vb.) ve duyuşsal özellikler modellerinin ya da ölçeklerinin farklı değişkenlere göre ölçme değişmezliğini test eden araştırmalardan örnekler verilmiştir.

Uluslararası Araştırmalar. He vd. (2019) araştırmasında PISA ve TIMSS 2015 uygulamalarında ortak yer alan 29 ülkenin Araçsal Motivasyon, Bilimden Zevk Alma ve Okula Aidiyet Duygusu Ölçekleri'ne ilişkin öğrenci verileri kullanılmıştır. Araştırmada ölçüm karşılaştırılabilirlik düzeyleri gösterilmiş, uygulamalar içindeki farklı ölçeklendirme yöntemleri ve uygulamalardaki öğrenci başarısı ile bu yapıların yakınsaması kontrol edilmiştir. Araştırma sonuçlarından biri her iki uygulamada yer alan üç ölçeğin (PISA'daki Okula Aidiyet Duygusu hariç) en azından metrik değişmezliğe ulaştığı şeklindedir. Karşılaştırmalı çıkarımlar yapmadan önce ölçme değişmezliğini kontrol etmenin önemi vurgulanmıştır.

Kankaras ve Moors (2014) araştırmasında PISA 2009 uygulamasında yer alan fen ve matematik okuryazarlık testleri ile okuma becerileri alt testlerinin kültürler arası karşılaştırılabilirliği test edilmiştir. Araştırmada MTK kullanılarak veri setindeki ölçüm eşdeğerliği incelenmiş ve farklı madde işleyişi Değişen Madde Fonksiyonu (DMF) analizleri ile test edilmiştir. Değişimlerin bölge ülkeleri gruplarında aynı yönde meydana gelme eğiliminde ve en çok etkilenen Güneydoğu Asya ülkeleri olduğu belirlenmiştir.

Hansson ve Gustafsson (2013) araştırmasında TIMSS 2003 uygulamasından elde edilen 8. sınıf İsveç verisinden yararlanarak öğrencilerin sosyoekonomik statülerinin öğrencilerin göçmenlik durumuna göre ölçme değişmezliği test edilmiştir. Araştırma öğrenci anketinde yer alan ve sosyoekonomik statü ile ilişkili olan dört madde üzerinden model tasarlanmıştır. Doğrulanan modelin öğrencilerin göçmenlik durumuna göre ölçme

değişmezliği incelenmiştir. ÇGDFA sonucunda öğrencilerin sosyoekonomik statülerinin göçmenlik durumlarına göre oluşturulan gruplar arasında sadece şekil ve metrik değişmezlik aşamalarını sağladığı tespit edilmiştir.

Wu vd. (2007) TIMSS 1999 uygulamasında 8. sınıf matematik başarı testinin ülkeler (Amerika Birleşik Devletleri, Kanada, Avustralya, Yeni Zelanda, Tayvan, Kore ve Japonya) arasında ölçme değişmezliği test edilmiştir. Araştırma seçilen yedi ülkenin ikili karşılaştırmaları yapılarak gerçekleştirilmiştir. Araştırma sonucunda şekil değişmezliğini ikili karşılaştırmaların tamamının sağladığı, metrik değişmezliği on dokuz ikili karşılaştırmaların sağladığı, ölçek değişmezliği ve katı değişmezliğini ise yedi ikili karşılaştırmaların karşıladığı sonucuna ulaşılmıştır.

Ulusal Araştırmalar. Bağdu Söyler vd. (2021) araştırmasında PISA 2015 uygulamasında yer alan okuma becerileri alt testi kapsamında ana dili İngilizce olan ve olmayan ülkeler arasında ölçme değişmezliği incelenmiştir. PISA uygulamalarında yer alan maddeler İngilizce olarak hazırlandığı ve birçok ülkede öğrenciler çeviri maddeleri yanıtladığından çevirilerden kaynaklanan ölçme değişmezliği sorunu olup olmadığı araştırılmıştır. Kanada, ABD ve İngiltere ana dili İngilizce olan; Japonya, Tayland ve Türkiye ise ana dili İngilizce olmayan ülke örneklemi olarak seçilmiştir. Geliştirilen modelin ölçme değişmezliği ÇGDFA ile test edilmiştir. Araştırmada okuma becerileri alt testinin dil değişkeni açısından ölçme değişmezliğini tam olarak sağlamadığı tespit edilmiştir.

Gönen (2021) araştırması kapsamında yapılan çalışmalardan biri PISA 2018 okuma becerileri testinin ülkeler (Türkiye, Singapur ve Amerika Birleşik Devletleri) arasında ölçme değişmezliğinin ÇGDFA ile incelenmesidir. Beş ünite üzerinden gerçekleştirilen araştırmada beş ünite yapısal değişmezlik aşamasını sağlamış, metrik değişmezlik aşaması ise dört ünite sağlanmıştır. Dört ünite iki tanesinde ise ölçek değişmezlik aşamasının sağlandığı sonucuna ulaşılmıştır.

Asar (2019) araştırmasında PISA 2015 matematik okuryazarlığı testinin farklı dilleri konuşan ülkeler arasında ölçme değişmezliği sınınanmıştır. İlişkisel tarama modelinin kullanıldığı araştırma 54 no'lu kitapçığı cevaplayan OECD üyesi ülkelerin verileri üzerinden

gerçekleştirilmiştir. Araştırma sonucunda farklı dilleri konuşan 12 OECD üyesi ülkesi için ölçme değişmezliğinin sağlanmadığı sonucu tespit edilmiştir.

Öncü (2019) araştırması kapsamında yapılan çalışmalardan biri TIMSS 2015 uygulamasında yer alan 8.sınıf matematik başarı testinin ülkeler arasında ölçme değişmezliğinin test edilmesidir. İlişkisel tarama modelinin kullanıldığı araştırma 11 no'lu kitapçığı alan OECD üyesi ülkelerinin verileri ile yürütülmüştür. Araştırma sonucunda OECD üyesi ülkeler için yapısal değişmezlik modelinin sağlandığı belirlenmiştir.

Polat (2019) araştırması kapsamında yapılan çalışmalardan biri TIMSS 2015 uygulamasında yer alan 8. sınıf öğrenci anketlerinden hazırlanan matematik ve fen duyuşsal özellik modellerinin kültürlere (Türkiye, Suudi Arabistan ve Singapur) göre ölçme değişmezliğinin test edilmesidir. ÇGDFA ile yapılan ölçme değişmezliği sonucunda matematik ve fen duyuşsal özellik modellerinin kültürlere göre ölçek değişmezliğini sağladığı sonucu elde edilmiştir.

Tavlıca (2019) araştırmasında TIMSS 2015 uygulamasında kullanılan 4.sınıf matematik başarı testinin OECD üyesi ülkeleri arasında ölçme değişmezliği incelenmiştir. Araştırmaya 24 ülke dâhil edilmiştir. İlişkisel tarama modelinde yürütülen araştırma 7 no'lu kitapçığı alan toplam 9.641 öğrenci verisi ile yürütülmüştür. ÇGDFA sonucunda yapısal değişmezlik modelinin sağlandığı belirtilmiştir. Buradan hareketle TIMSS 2015 uygulamasında 4.sınıf matematik başarı testi puanlarının seçilen OECD üyesi ülkeleri arasında ölçme değişmezliğini sağlamadığı tespit edilmiştir. Dolayısıyla adı geçen ülkeler arasında 4.sınıf matematik başarı testi puanlarının karşılaştırılamayacağı sonucuna varılmıştır.

Karakoç Alatl (2016) araştırması kapsamında yapılan çalışmalardan biri PISA 2012 matematik ve fen okuryazarlık testleri ile okuma becerileri testlerinin ülkeler arasında (Avustralya, Şangay-Çin, Türkiye ve Fransa) ölçme değişmezliğini farklı teknikler ile test edilmesidir. Betimsel olan araştırma uygulamaya seçilen ülkelere katılan, 3 no'lu kitapçığı alan öğrencilerin verileri ile gerçekleştirilmiştir. ÇGDFA sonucunda PISA 2012 kapsamında ele aldığı testlerin seçtiği ülke örneklemi arasında dil değişkenine göre yapısal değişmezlik

sağlanırken metrik değişmezliğin olmadığı belirlenmiştir. Dolayısıyla dil değişkenine göre ölçme değişmezliğinin sağlanamadığı bulunmuştur.

Başusta ve Gelbal (2015) araştırmasında PISA 2006 uygulamasında öğrenci anketinde yer alan fen bilgisi ve teknolojileri ile ilgili maddelere dayalı geliştirilen modelin cinsiyete göre ölçme değişmezliği test edilmiştir. Türkiye'den uygulamaya katılan 4942 öğrencinin verilerinin YEM'e dayalı analizleri sonucunda modelin biçimsel, metrik, ölçek ve katı değişmezliği sağladığı belirlenmiştir. Buradan hareketle modelden ulaşılan ölçümlerin psikometrik özelliklerinin cinsiyet grupları arası genellenebileceği ortaya koyulmuştur.

Kıbrısloğlu (2015) araştırması kapsamında yapılan çalışmalardan biri PISA 2012 uygulamasında yer alan anketlerden hazırlanan matematik öğrenme modelinin ülkeler (Türkiye, Çin (Şangay), Endonezya) arasında ölçme değişmezliğinin test edilmesidir. Oluşturulan modelin doğrulanmasının ardından ÇGDFA ile ölçme değişmezliği aşamaları sınanmıştır. Araştırmada matematik öğrenme algısı modelinin ülkeler arasında sadece şekil değişmezliğini sağladığı sonucuna ulaşılmıştır.

Alan yazın incelendiğinde ölçme değişmezliğini test eden araştırmalarda ya başarı testleri ya da duyuşsal özellikler ile ilgili modellerin ele alındığı görülmektedir. Başarı testi ile duyuşsal özellikler ile ilgili oluşturulan modelin aynı ülke örneklemi için test edildiği araştırmaya rastlanmamıştır. Bu araştırmada ise TIMSS 2019 uygulamasında yer alan hem 8.sınıf matematik başarı testi hem de duyuşsal özelliklerle ilgili oluşturulan modelin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği test edilmiştir. Yani seçilen üç ülke örnekleme için başarı testi veya duyuşsal özelliklerle ilgili oluşturulan modelin aynı biçimde algılanıp algılanmadığı ortaya koyulmuştur. Böylelikle başarı testi veya duyuşsal özelliklerle ilgili oluşturulan modelin örneklem arasında kıyaslama yapılıp yapılmayacağı saptanmıştır. Diğer bir ifadeyle üç ülke örnekleminin benzer veya farklı faktör yapısına sahip olup olmadıkları tespit edilerek faktör yapısının evrensel mi yoksa kültürel olarak bağımlı mı olduğunu karakterize etmek için çalışılmıştır. Öte yandan alan yazındaki araştırmalarda duyuşsal özellik bölümünde yer alan ölçekler ya ayrı ayrı ele alınıp ölçme değişmezliği test edilmiş ya da birkaç ölçek bir arada ele alınarak tasarlanan modelin ölçme değişmezliği

incelenmiştir. TIMSS raporlarında ise ölçeklerin her biri için ayrı analizler yapıldığı görülmektedir. Bu nedenle araştırma kapsamında ölçeklerin bir arada bulunduğu modelin yanı sıra modelde yer alan her bir ölçeğin de ölçme değişmezliği incelenmiştir. Böylelikle her iki durumdan elde edilen değerlerin karşılaştırılmasına fırsat sunulmuştur.

Bölüm 3

Yöntem

Araştırmanın bu bölümünde sırasıyla araştırmanın türü, çalışma grubu, veri toplama süreci, veri toplama araçları ve verilerin analizleri yer almaktadır.

Araştırmanın Türü

Bu araştırmada, TIMSS 2019 uygulamasında kullanılan 8.sınıf matematik başarı testinin, duyuşsal özelliklerle ilgili oluşturulacak modelin ve modelde yer alan her bir ölçeğin ülkelere (Türkiye, Güney Kore ve Finlandiya) göre ölçme değışmezliğini sağlayıp sağlamadığını ortaya koymak amacıyla nicel araştırma yöntemlerinden ilişkisel tarama modeli kullanılmıştır. Karasar (2000:79) tarafından ilişkisel tarama modeli “İki ve daha çok sayıdaki değışken arasında birlikte değışim varlığını ve/veya derecesini belirlemeyi amaçlayan araştırma modeli” şeklinde tanımlanmıştır. Araştırmada gözlenen değışkenlerin gizil değışkenler üzerindeki etkisi incelendiğinden diğeri bir deyişle matematik başarı testinin ve duyuşsal özelliklerle ilgili oluşturulacak modelin ve ölçeklerin geçerlik düzeyi belirlenmeye çalışıldığından araştırma ilişkisel tarama modelindedir.

Çalışma Grubu

TIMSS 2019 uygulamasında 8.sınıf düzeyinde 39 ülke ve 7 karşılaştırma katılımcısı yer almıştır. Bu araştırmanın evrenini Türkiye, Güney Kore ve Finlandiya ülkelerinde 8.sınıfta öğrenim gören tüm öğrenciler oluşturmaktadır. Bu ülkelerin seçiminde farklı kültürleri olması, kişi başına düşen milli gelirlerinin farklı olması ve başarı sıralamalarının anlamlı derecede farklılaşması gibi özellikler etkili olmuştur. Öte yandan her üç ülke de eTIMSS değerlendirmesinde yer almıştır. Her ne kadar kâğıt-kalem tabanlı ve bilgisayar tabanlı değerlendirme türü içerik olarak mümkün olduğunca benzer olacak şekilde geliştirilmiş olsa da, iki farklı değerlendirme tarzının bir sonucu olarak aralarında kaçınılmaz olarak bazı farklılıklar oluşabileceğinden aynı değerlendirme türüne katılan ülkelere seçim yapılmıştır.

Araştırmada 8.sınıf matematik başarı testinin ülkelere göre ölçme değişmezliğinin test edilmesinde yer alan çalışma grubu Türkiye, Güney Kore ve Finlandiya için 14 nolu kitapçığı cevaplayan öğrencilerdir. Duyuşsal özelliklerle ilgili oluşturulan modelin ve ölçeklerin ülkelere göre ölçme değişmezliğinin incelenmesinde yer alan çalışma grubu ise Türkiye, Güney Kore ve Finlandiya'dan 8.sınıf düzeyinde katılan tüm öğrencilerdir. Ölçeklere ilişkin verilerin alındığı BSG...Z7 dosyalarında 1-16 nolu kitapçıkları yani 16 adet kitapçığı cevaplayan tüm öğrencilerin verileri vardır. Dolayısıyla eTIMSS dosyaları içerisinde yer alan PSI madde bloklarının yer aldığı 15-16 nolu kitapçıkları cevaplayan öğrencilerin ölçeklere ilişkin verileri de kapsama alınmıştır. Ana raporlarda yapılan karşılaştırmalarda PSI madde blokları sadece eTIMSS değerlendirmesinde yer aldığından TIMSS ve eTIMSS için ortak olarak 14 adet kitapçığın kullanıldığı unutulmamalıdır. Bu kapsamda başarı testi için 14 nolu kitapçığı cevaplayan öğrenci sayısı ile duyuşsal özelliklerle ilgili oluşturulan model için 8.sınıf düzeyinde katılan tüm öğrenci sayıları Tablo 2'de sunulmuştur:

Tablo 2

Ülkelere Göre Öğrenci Sayıları

Ülkeler	Başarı Testi 14 No'lu Kitapçık (N)	Duyuşsal Özellik Modeli ve Ölçekler (N)
Türkiye	295	4662
Güney Kore	273	4409
Finlandiya	354	5570
Toplam	922	14641

Veri Toplama Süreci

Bu araştırmada TIMSS 2019 uygulamasında yer alan 8. sınıf matematik başarı testi olarak 14 nolu kitapçıkta yer alan ve duyuşsal özellikler ile ilgili olarak "Matematik Öğrenmeyi Sevme", "Matematikte Kendine Güven" ve "Matematiğe Değer Verme" ölçeklerinde bulunan maddelerin Türkiye, Güney Kore ve Finlandiya ülkelerine ilişkin verileri kullanılmıştır. Veriler TIMSS&PIRLS Uluslararası Araştırma Merkezi'nin internet sitesinden indirilmiştir (<https://timss2019.org/international-database/>). Ülkelere göre 8.sınıf düzeyinde başarı testi

için BSA...Z7 dosyalarından, ölçekler için ise BSG...Z7 dosyalarından yararlanılmıştır. "...” yerine ISO 3166 alfa kodlama şemasına dayalı 3 karakterli ülke kısaltmaları gelmektedir (Fishbein ve diğerleri, 2021, 64).

Veri Toplama Araçları

Araştırma kapsamında veri toplama araçları olarak öğrenci başarı testi ve duyuşsal özelliklerle ilgili ölçekler kullanılmıştır. Veri toplama araçlarının araştırma problemlerine göre dağılımı Tablo 3'te verilmiştir:

Tablo 3

Veri Toplama Araçları

Araştırma Problemleri	Veri Toplama Aracı
Soru 1	Başarı Testi
Soru 2-3	Duyuşsal Özellik Modelinde Yer Alan Ölçekler

Başarı Testi

TIMSS 2019 uygulamasında kâğıt-kalem ve bilgisayar tabanlı değerlendirmelerde hem matematik hem de fen alanındaki başarı testleri 14 farklı kitapçık olarak tasarlanmıştır. Her ülkedeki her bir kitapçığı cevaplayan öğrenciler, öğrenci becerileri açısından yaklaşık olarak eşdeğerdir. Diğer yandan her bir maddeyi yeterli sayıda öğrencinin yanıtladığından emin olmak için her ülkede en az 4.000 öğrenciden oluşan bir örneklem hedeflenmiştir. Her öğrenci bir kitapçık doldurmuştur. 14 adet öğrenci kitapçığı, katılımcı sınıflardaki öğrencilere önceden belirlenmiş bir sıraya göre dağıtılmış, böylece öğrencilerin her kitapçığa yaklaşık olarak eşit oranda yanıt vermesi sağlanmıştır (Mullis & Martin, 2017).

Öğrenci kitapçıklarında yer alan bloklar 8. sınıf düzeyinde 12-18 arası maddeden oluşmaktadır. TIMSS 2019 uygulamasında önceki uygulamalarda olduğu gibi iki madde türü kullanılmıştır. Bunlar çoktan seçmeli (örneğin öğrencilerin bir dizi seçenek arasından yanıt(lar)ını seçtikleri) madde ve açık uçlu (örneğin öğrencilerin kendi yanıt(lar)ını

oluşturdukları) maddelerdir. Çoktan seçmeli maddeler, tekli seçim ve çoklu seçim olacak şekilde yapılandırılmıştır.

TIMSS 2019 uygulamasında 8.sınıf matematik değerlendirmesinde “sayı, cebir, geometri, veri ve olasılık” olmak üzere dört öğrenme alanı değerlendirilmiştir. Sayı öğrenme alanında öğrencilerden tamsayılar, kesirler ve ondalık sayıların yanı sıra daha gelişmiş tam sayı kavramları ile işlemlerini içeren problemleri hesaplamaları ve çözmeleri istenmiştir. Cebir öğrenme alanında öğrencilerden cebirsel modelleri kullanarak gerçek hayat problemlerini çözmeleri ve cebirsel kavramları içeren ilişkileri açıklamaları gerekli görülmüştür. Örneğin iki niceliği içeren bir formülde bir nicelik verildiğinde diğer niceliği bulmaları beklenmiştir. Ayrıca doğrusal denklemler ve fonksiyonları içeren problemler yer almıştır. Geometri öğrenme alanında öğrencilerden çeşitli iki ve üç boyutlu şekillerin özelliklerini analiz etmeleri ile çevre, alan ve hacim hesaplamaları yapmaları istenmiştir. Eşlik, benzerlik ve Pisagor teoremi gibi geometrik ilişkilere dayalı problemler çözmeleri ve açıklamalar yapmaları beklenmiştir. Veri ve olasılık öğrenme alanında ise öğrencilerden çeşitli görselleri okumaları ve bunlardan önemli anlamlar çıkarmaları, veri dağılımlarının altında yatan istatistikleri tanımaları, verileri düzenlemeleri ve temsil etmeleri istenmiştir. Temel olasılık kavramlarıyla ilgili bazı sorular da yer almıştır (Mullis ve diğerleri, 2020). TIMSS 2019 uygulamasında 8.sınıf matematik değerlendirmesinde yer alan öğrenme ve konu alanlarının dağılımı Tablo 4’te verilmiştir:

Tablo 4

Öğrenme ve Konu Alanlarının Dağılımı

Öğrenme Alanı	Konu Alanı	Yüzde	Toplam
Sayılar	Tam Sayılar	%10	%30
	Kesirler ve Ondalık Sayılar	%10	
	Oran, Orantı ve Yüzde	%10	
Cebir	İfadeler, İşlemler ve Denklemler	%20	%30
	İlişkiler ve Fonksiyonlar	%10	
Geometri	Geometrik Şekiller ve Ölçümler	%20	%20
Veri ve Olasılık	Veri	%15	%20
	Olasılık	%5	

TIMSS 2019 uygulamasında 8.sınıf matematik değerlendirmesinde öğrencilerin öğrenme alanlarında bir dizi bilişsel beceriden yararlanmaları beklenmiştir. Bu beceriler “bilme, uygulama ve akıl yürütme” olmak üzere üç bilişsel alanda kategorize edilmiştir. Bilme alanı öğrencilerin bilmesi gereken gerçekleri, kavramları ve prosedürleri kapsarken, ikinci alan olan uygulama alanı öğrencilerin sorunları çözmek veya soruları cevaplamak için bilgiyi ve kavramsal anlayışı uygulama becerisine aktarmaya odaklanmıştır. Akıl yürütme alanı ise alışılmadık durumları, karmaşık bağlamları ve çok adımlı sorunları kapsayacak şekilde rutin problemlerin çözümünün ötesine geçmiştir (Mullis ve diğerleri, 2020). TIMSS 2019 uygulamasında 8.sınıf düzeyinde matematik değerlendirmesinin bilişsel alanlarının dağılımları Tablo 5’te verilmiştir:

Tablo 5

Bilişsel Alanların Dağılımı

Bilişsel Alan	Yüzde
Bilme	%35
Uygulama	%40
Akıl Yürütme	%25

Kaynak. Cotter, Centurino ve Mullis, 2020, 1.25

Araştırma kapsamında 8.sınıf düzeyinde matematik alanında yer alan öğrenci kitapçıkları içerdiği madde türleri bakımından incelenmiştir. Aynı kitapçıkta yer alan tek bir madde türü alınmamış, hem çoktan seçmeli maddeler hem de açık uçlu maddeler araştırmaya dâhil edilmiştir. Çoktan seçmeli (4 seçenekli) madde sayısının ve açık uçlu madde sayısının yakın olmasına dikkat edilmiştir. İstenen durumu sağlayan kitapçıklardan biri olan “14 nolu öğrenci kitapçığı” seçilmiştir. Öte yandan her bir bloktaki öğrenme ve bilişsel alanlardaki maddelerin dağılımının, mümkün olduğunca, madde havuzundaki dağılımla eşleştiği belirtilmektedir (Mullis & Martin, 2017). 14 nolu kitapçıkta yer alan, araştırma kapsamında seçilen 33 madde ile bu maddelere ilişkin detaylı bilgiler Tablo 6’da sunulmuştur:

Tablo 6*Başarı Testindeki Maddeler ve Maddelere İlişkin Bilgiler*

Madde Kodu	Blok – Blok Dizisi	Öğrenme Alanı	Konu Alanı	Bilişsel Alan	Madde Türü
ME52024	ME01_01	Sayılar	Kesirler ve Ondalık Sayılar	Bilme	ÇSM
ME52058A	ME01_02A	Sayılar	Tam Sayılar	Uygulama	AUM
ME52058B	ME01_02B	Sayılar	Oran, Orantı ve Yüzde	Uygulama	AUM
ME52125	ME01_03	Sayılar	Tam Sayılar	Akıl Yürütme	ÇSM
ME52229	ME01_04	Sayılar	Kesirler ve Ondalık Sayılar	Bilme	AUM
ME52063	ME01_05	Cebir	İfadeler, İşlemler ve Denklemler	Uygulama	ÇSM
ME52072	ME01_06	Cebir	İfadeler, İşlemler ve Denklemler	Bilme	ÇSM
ME52146A	ME01_07A	Cebir	İlişkiler ve Fonksiyonlar	Akıl Yürütme	AUM
ME52146B	ME01_07B	Cebir	İlişkiler ve Fonksiyonlar	Akıl Yürütme	AUM
ME52092	ME01_08	Cebir	İlişkiler ve Fonksiyonlar	Uygulama	ÇSM
ME52046	ME01_09	Geometri	Geometrik Şekiller ve Ölçümler	Akıl Yürütme	ÇSM
ME52083	ME01_10	Geometri	Geometrik Şekiller ve Ölçümler	Uygulama	ÇSM
ME52082	ME01_11	Geometri	Geometrik Şekiller ve Ölçümler	Uygulama	ÇSM
ME52161	ME01_12	Veri ve Olasılık	Olasılık	Uygulama	ÇSM
ME52418A	ME01_13A	Veri ve Olasılık	Veri	Uygulama	ÇSM
ME52418B	ME01_13B	Veri ve Olasılık	Veri	Uygulama	ÇSM
ME72005	ME14_01	Sayılar	Tam Sayılar	Bilme	ÇSM
ME72021	ME14_02	Sayılar	Kesirler ve Ondalık Sayılar	Bilme	AUM
ME72026	ME14_03	Sayılar	Kesirler ve Ondalık Sayılar	Akıl Yürütme	AUM
ME72041A	ME14_04A	Sayılar	Oran, Orantı ve Yüzde	Uygulama	AUM
ME72041B	ME14_04B	Sayılar	Oran, Orantı ve Yüzde	Uygulama	AUM
ME72223	ME14_05	Cebir	İfadeler, İşlemler ve Denklemler	Bilme	ÇSM

Tablo 6 (devam)

ME72094	ME14_06	Cebir	İfadeler, İşlemler ve Denklemler	Bilme	AUM
ME72059	ME14_07	Cebir	İfadeler, İşlemler ve Denklemler	Uygulama	AUM
ME72080	ME14_08	Cebir	İfadeler, İşlemler ve Denklemler	Bilme	ÇSM
ME72081	ME14_09	Cebir	İfadeler, İşlemler ve Denklemler	Uygulama	AUM
ME72140	ME14_10	Geometri	Geometrik Şekiller ve Ölçümler	Bilme	AUM
ME72120	ME14_11	Geometri	Geometrik Şekiller ve Ölçümler	Akıl Yürütme	AUM
ME72131	ME14_12	Geometri	Geometrik Şekiller ve Ölçümler	Akıl Yürütme	AUM
ME72147	ME14_13	Geometri	Geometrik Şekiller ve Ölçümler	Uygulama	AUM
ME72154	ME14_14	Veri ve Olasılık	Veri	Uygulama	ÇSM
ME72192	ME14_15	Veri ve Olasılık	Veri	Uygulama	ÇSM
ME72161	ME14_16	Veri ve Olasılık	Veri	Akıl Yürütme	AUM

Veriler <https://timss2019.org/international-database/> sayfasından indirilmiştir.

ÇSM: Çoktan Seçmeli Madde

AUM: Açık Uçlu Madde

Araştırma kapsamında açık uçlu maddelerde “doğru cevap (correct response)” olarak belirlenmiş yanıtlar “1”; “yanlış cevap (incorrect response)”, “erişilememiş”, “atlanmış veya geçersiz” olarak belirlenmiş yanıtlar ise “0” olarak dönüştürülmüştür. Böylelikle 14 nolu kitapçıktan seçilen 16 tanesi çoktan seçmeli, 17 tanesi açık uçlu olmak üzere toplam 33 maddeye (alt madde) ilişkin yanıtlar (veri seti) “1,0” olacak şekilde iki kategorili biçimde düzenlenmiştir.

Duyuşsal Özelliklerle İlgili Oluşturulan Modelde Yer Alan Ölçekler

TIMSS uygulamalarında 1995 yılından bu yana öğrencilerin matematik ve fen bilimleri başarılarına yönelik tutumları değerlendirilmektedir. TIMSS 2019 uygulamasında da öğrencilerin tutumlarını değerlendirmek için çeşitli ölçekler yer almıştır. TIMSS 2019 uygulamasında yer alan “Matematik Öğrenmeyi Sevme”, “Matematikte Kendine Güven” ve

“Matematiğe Değer Verme” ölçekleri araştırmada duyuşsal özelliklerle ilgili oluşturulan modelde kullanılmıştır. “Matematik Öğrenmeyi Sevme” ölçeğinde öğrencilerin konuları öğrenmeye yönelik içsel motivasyonunu kapsayan dokuz madde; “Matematikte Kendine Güven” ölçeğinde öğrencilerin konuya özgü benlik kavramlarına ilişkin dokuz madde; “Matematiğe Değer Verme” ölçeğinde ise öğrencilerin dışsal motivasyonunu anlamaya yönelik dokuz madde yer almıştır (Mullis & Martin, 2017). Öğrenciler bu ölçeklerdeki maddelere “çok katılıyorum, biraz katılıyorum, biraz katılmıyorum, çok katılmıyorum” şeklinde verdikleri yanıtlara göre puanlanmıştır. Çok katılıyorum 1, biraz katılıyorum 2, biraz katılmıyorum 3, çok katılmıyorum 4 ile kayıp veya geçersiz veriler ise 9 ile kodlanmıştır. “Matematik Öğrenmeyi Sevme” alt ölçeğinde 2 madde (BSBM16B, BSBM16C); “Matematikte Kendine Güven” alt ölçeğinde ise 5 madde (BSBM19B, BSBM19C, BSBM19E, BSBM19H ve BSBM19I) ters kodlanmıştır. “Matematiğe Değer Verme” alt ölçeğinde ise ters kodlanan madde bulunmamaktadır. Ölçeklerdeki maddelerin kodları ve ifadeleri Tablo 7’de verilmiştir:

Tablo 7*Duyuşsal Özelliklerle İlgili Oluşturulan Modelde Yer Alan Ölçeklerin Maddeleri*

Ölçek	Madde Kodu	Madde İfadesi
Matematik Öğrenmeyi Sevme	BSBM16A	Matematik öğrenmekten zevk alırım.
	BSBM16B*	Keşke matematik çalışmak zorunda olmasam.
	BSBM16C*	Matematik sıkıcıdır.
	BSBM16D	Matematikte birçok ilginç şey öğrenirim.
	BSBM16E	Matematiği severim.
	BSBM16F	Sayıları içeren herhangi bir okul çalışmasını severim.
	BSBM16G	Matematik problemlerini çözmeyi severim.
	BSBM16H	Matematik derslerini sabırsızlıkla beklerim.
	BSBM16I	Matematik en sevdiğim derslerden biridir.
Matematikte Kendine Güven	BSBM19A	Matematikte genelde iyiyimdir.
	BSBM19B*	Matematik benim için birçok sınıf arkadaşımın göre daha zordur.
	BSBM19C*	Matematik güçlü yanlarımdan biri değildir.
	BSBM19D	Matematikte her şeyi çabuk öğrenirim.
	BSBM19E*	Matematik beni sinirlendirir.
	BSBM19F	Zor matematik problemlerini çözmekte iyiyimdir.
	BSBM19G	Öğretmenim bana matematikte iyi olduğumu söyler.
	BSBM19H*	Matematik benim için diğer derslerden daha zordur.
	BSBM19I*	Matematik kafamı karıştırır.
Matematiğe Değer Verme	BSBM20A	Matematik öğrenmenin günlük hayatımda bana yardımcı olacağını düşünürüm.
	BSBM20B	Diğer dersleri öğrenmek için matematiğe ihtiyacım vardır.
	BSBM20C	İstediğim üniversiteye girebilmek için matematikte başarılı olmam gerekir.
	BSBM20D	İstediğim işe girebilmek için matematikte başarılı olmam gerekir.
	BSBM20E	Matematiği kullanmayı içeren bir işte çalışmak isterim.
	BSBM20F	Hayatta başarılı olmak için matematik öğrenmek önemlidir.
	BSBM20G	Matematik öğrenmek yetişkin olduğumda bana daha fazla iş fırsatı sunacaktır.
	BSBM20H	Ailem matematikte iyi olmamın önemli olduğunu düşünür.
	BSBM20I	Matematikte başarılı olmak önemlidir.

* Maddeler ters kodlanmıştır.

Verilerin Analizi

Verilerin analizi sürecinde öncelikle sayıtların sağlanıp sağlanmadığı araştırılmış ve sağlanmadığı takdirde gerekli analiz yöntemleri kullanılarak veriler analize uygun hale getirilmiştir. TIMSS 2019 8. sınıf matematik başarı testindeki ve duyuşsal özelliklerle ilgili maddeler betimsel olarak incelenerek açımlayıcı faktör analizleri yapılmıştır. Ardından doğrulayıcı faktör analizleri yapılarak doğrulama işlemleri gerçekleştirilmiştir. Sonrasında 8. sınıf matematik başarı testi ve duyuşsal özelliklerle ilgili oluşturulan modelin ve ölçeklerin ülkelere göre ölçme değişmezliğini test etmek için çok gruplu doğrulayıcı faktör analizi (ÇGDFA) yapılmıştır. Verilerin analizi sürecinde Mplus 8.3, IBM SPSS Statistics 22 ve Factor (12.03.02.WIN64) programları kullanılmıştır. Bu kapsamda gerçekleştirilen tüm analizler aşağıda detaylı bir şekilde sunulmuştur:

Başarı Testine İlişkin Analizler

Kayıp Veri. TIMSS 2019 başarı veri dosyalarında çoktan seçmeli ve açık uçlu maddelerde “doğru cevap” olarak belirlenmiş yanıtlar “1”; “yanlış cevap”, “erişilememiş”, “atlanmış veya geçersiz” olarak belirlenmiş yanıtlar ise “0” olarak atanmıştır. Böylelikle 14 nolu kitapçıkta bulunan tüm yanıtlar iki kategorili şekilde belirtilmiştir. Türkiye, Güney Kore ve Finlandiya veri dosyaları incelendiğinde sistem hatası (system missing) olarak belirtilen yanıtların ise sadece Finlandiya dosyasında olduğu tespit edilmiştir. Finlandiya veri dosyasında 3 öğrencinin 33 maddeye ilişkin hiçbir yanıtının olmaması nedeniyle verileri silinmiştir. Yapılan atama ve silme işlemlerinin ardından başarı veri dosyalarında 919 öğrenci verisi yer almıştır.

Uç Değer. Başarı testlerinde yer alan veriler iki kategorili biçime dönüştürüldüğünden verilerde uç değerler incelenmemiştir.

Normallik. Araştırmada normallik incelemesi için basıklık ve çarpıklık değerleri kullanılmıştır. Tablo 8 incelendiğinde çarpıklık değerlerinin -1.076 ile 1.761; basıklık değerlerinin ise -1.998 ile 1.100 olduğu belirlenmiştir. George ve Mallery (2010) basıklık

aralıkların ± 2 ; Tabachnick ve Fidell (2013) ise çarpıklık aralıklarının $\pm 1,5$ olduğunu kabul etmiştir. Bu değerlerden hareketle normallikte sapmanın önemli olmadığı yorumlanmıştır.

Elde edilen basıklık ve çarpıklık değerleri Tablo 8'de verilmiştir:

Tablo 8

Başarı Testindeki Maddelerin Çarpıklık ve Basıklık Değerleri

Madde	Ortalama	Varyans	Çarpıklık	Basıklık
ME52024	0.634	0.232	-0.559	-1.687
ME52058A	0.721	0.201	-0.989	-1.022
ME52058B	0.305	0.212	0.850	-1.278
ME52125	0.508	0.250	-0.033	-1.998
ME52229	0.536	0.249	-0.146	-1.978
ME52063	0.522	0.250	-0.089	-1.991
ME52072	0.642	0.230	-0.593	-1.648
ME52146A	0.437	0.246	0.253	-1.935
ME52146B	0.170	0.141	1.761	1.100
ME52092	0.305	0.212	0.850	-1.278
ME52046	0.297	0.209	0.889	-1.209
ME52083	0.436	0.246	0.257	-1.933
ME52082	0.640	0.230	-0.583	-1.659
ME52161	0.737	0.194	-1.076	-0.843
ME52418A	0.410	0.242	0.365	-1.866
ME52418B	0.486	0.250	0.054	1.996
ME72005	0.518	0.250	-0.072	-1.994
ME72021	0.506	0.250	-0.024	-1.998
ME72026	0.510	0.250	-0.041	-1.997
ME72041A	0.490	0.250	0.041	-1.997
ME72041B	0.411	0.242	0.361	-1.869
ME72223	0.553	0.247	-0.213	-1.954
ME72094	0.669	0.221	-0.720	-1.481
ME72059	0.328	0.220	0.736	-1.458
ME72080	0.362	0.231	0.573	-1.671

Tablo 8 (devam)

ME72081	0.309	0.214	0.827	-1.315
ME72140	0.556	0.247	-0.226	-1.948
ME72120	0.311	0.214	0.816	-1.333
ME72131	0.280	0.201	0.983	-1.034
ME72147	0.257	0.191	1.115	-0.758
ME72154	0.712	0.205	-0.935	-1.125
ME72192	0.609	0.238	-0.449	-1.798
ME72161	0.420	0.244	0.324	-1.894

Bağlantılılık. Maddeler arasında 0.90 üzerinde korelasyon olması çoklu bağlantı sorunu oluşturmaktadır. Maddelere ilişkin yanıtlar kategorik olduğundan maddeler arasındaki korelasyonlar tetrakorik korelasyon matrisindeki değerler incelenerek belirlenmiştir. ME72041A ile ME72041B maddeleri arasındaki korelasyon değeri 0.933 olarak hesaplandığından iki maddenin benzer özellikler ölçtüğü anlaşılmıştır (Tabachnick ve Fidell, 2013). Söz konusu iki madde aynı öğrenme, konu ve bilişsel alandan ayrıca aynı madde türünden olduğu için ME72041B maddesi çıkarılmıştır. 32 madde ile analizlere devam edilmiştir.

Tüm sayıtlılar kontrol edilmiş olup gerekli işlemler yapılmıştır. Böylelikle veriler ÇGDFA için hazır hale getirilerek model oluşturma adımlarına geçilmiştir.

Modelin oluşturulması. Öncelikle 32 maddelik veri seti üzerinden AFA yapılmıştır. Verilerin faktörlenebilir olup olmadıkları tespit edilmiştir. Açıklanan toplam varyans değeri dikkate alınarak başarı testinin faktör sayısına karar verilmiş ve maddelerin faktör yükleri incelenerek maddelerin uygunluğu kontrol edilmiştir. Ardından oluşturulan başarı testi DFA ile doğrulanmış ve yol diyagramı hazırlanmıştır.

Açımlayıcı Faktör Analizi. 32 maddelik veri seti ile AFA yapılmıştır. Veriler kategorik olduğundan AFA, tetrakorik korelasyon matrislerine göre Factor (12.03.02.WIN64) programında yapılmıştır. 32 maddelik başarı testinin KMO = 0.96975 ve Bartlett değeri [$\chi^2=10436.5$; $p < .05$] olarak bulunmuştur. KMO değeri 0.60'dan yüksek ve Bartlett değeri

anamlı bulunduğundan elde edilen değerler ile verilerin faktörlenebilir olduğu ortaya koyulmuştur (Çokluk ve diğerleri, 2010).

Özdeğeri 1'den büyük 4 faktör vardır. Birinci faktörün açıkladığı toplam varyans %51, ikinci faktörün açıkladığı toplam varyans %04, üçüncü faktörün açıkladığı toplam varyans %03, dördüncü faktörün açıkladığı toplam varyans da %03'tür. Birinci faktörün diğer faktörlerle arasında yaklaşık %50'lik yük farkı olduğu görülmektedir. Tek faktörün açıkladığı toplam varyans %30 ve üzeri olduğunda veri toplama aracı tek boyutlu kabul edilebileceğinden (Büyüköztürk, 2002) başarı testi tek boyutlu alınmıştır. Alan yazın incelendiğinde başarı testlerinin tek bir boyutta toplanması beklenen bir durumdur. Maddelerin faktör yükleri Tablo 9'da verilmiştir:

Tablo 9

Başarı Testindeki Maddelerin Faktör Yükleri

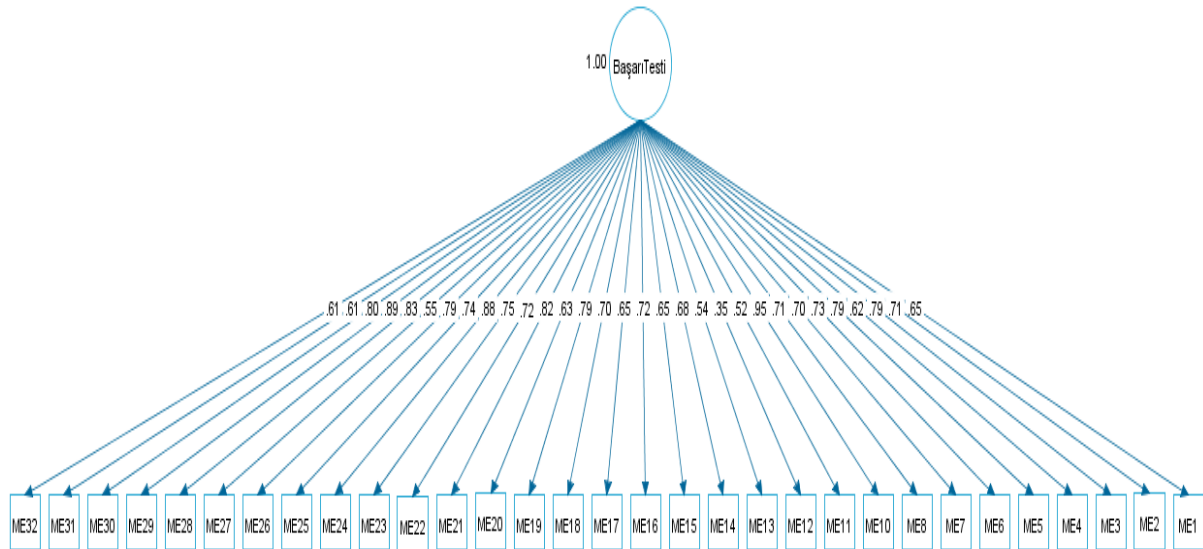
Maddeler	Faktör Yükleri	Maddeler	Faktör Yükleri
ME1 ME52024	0.652	ME17 ME72005	0.703
ME2 ME52058A	0.712	ME18 ME72021	0.791
ME3 ME52058B	0.792	ME19 ME72026	0.628
ME4 ME52125	0.622	ME20 ME72041A	0.818
ME5 ME52229	0.789	ME21 ME72223	0.722
ME6 ME52063	0.728	ME22 ME72094	0.745
ME7 ME52072	0.703	ME23 ME72059	0.874
ME8 ME52146A	0.718	ME24 ME72080	0.746
ME9 ME52146B	0.948	ME25 ME72081	0.796
ME10 ME52092	0.524	ME26 ME72140	0.548
ME11 ME52046	0.351	ME27 ME72120	0.831
ME12 ME52083	0.539	ME28 ME72131	0.892
ME13 ME52082	0.673	ME29 ME72147	0.797
ME14 ME52161	0.652	ME30 ME72154	0.616
ME15 ME52418A	0.720	ME31 ME72192	0.608
ME16 ME52418B	0.645	ME32 ME72161	0.702

Başarı testindeki maddelerin faktör yüklerinin 0.351 ile 0.948 arasında değiştiği görülmektedir. Maddelerin faktör yükleri .30'dan büyük olduğundan (Şencan, 2005) 32 maddelik başarı testinden madde çıkarılmamıştır. Elde edilen değerlerden hareketle maddelerin uygunluğu belirlenmiştir.

Doğrulayıcı Faktör Analizi. AFA ile ortaya koyulan değerlerin doğrulanması amacıyla DFA yapılmıştır. 32 madde tek faktörlü başarı testi için yapılan analiz sonucunda RMSEA=0.033, CFI=0.983, TLI =0.982 ve SRMR=0.048 değerlerine ulaşılmıştır. DFA çıktısı incelendiğinde ME52146B (M9) maddesinin eşik (threshold) değerinin diğer maddelere kıyasla çok yüksek olduğu tespit edilmiştir. Söz konusu madde, ME52146A maddesi ile aynı öğrenme, konu, bilişsel alana sahip ve aynı tip madde türü olduğundan ME52146B (ME9) maddesi çıkarılmış ve analizlere 31 madde ile devam edilmiştir. 31 maddelik başarı testi için çizilen yol diyagramı Şekil 1'de sunulmuştur:

Şekil 1

Başarı Testi Yol Diyagramı



Şekil 1'de görüldüğü üzere 31 madde tek boyutta yer almıştır. Maddelere ait faktör yükleri 0.35 ile 0.95 arasında değişmektedir. Elde edilen faktör yüklerinin 0.05 düzeyinde anlamlı olduğu belirlenmiştir. Başarı testinin veriye uyumunu kontrol etmek için uyum değerleri hesaplanmış ve Tablo 10'da sunulmuştur:

Tablo 10*Başarı Testi Uyum Değerleri*

İndeks	İyi Uyum	Kabul Edilebilir Uyum	Araştırma Değeri
χ^2	$p > 0.05$	$p > 0.05$	896.083 ($p < .05$)
χ^2 / sd	$0 \leq \chi^2 / sd \leq 3$	$3 < \chi^2 / sd \leq 5$	2.06 (896.083 / 434)
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 < RMSEA \leq 0.08$	0.034
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI < 0.95$	0.983
TLI	$0.95 \leq TLI \leq 1.00$	$0.90 \leq TLI < 0.95$	0.981
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 < SRMR \leq 0.08$	0.049

*Hu & Bentler (1999)

Başarı testinin uyum değerleri incelendiğinde tüm indekslerin (RMSEA, CFI, TLI ve SRMR) iyi uyum aralığında olduğu görülmektedir. Bu değerlerden hareketle başarı testinin veriye uyumunun iyi düzeyde olduğu ifade edilebilir.

Çok Gruplu Doğrulamalı Faktör Analizi. Başarı testinin doğrulanmasının ardından testin seçilen ülkelere (Türkiye, Güney Kore, Finlandiya) göre ölçme değişmezliğini test etmek için Mplus 8.3 programı ile ÇGDFA yapılmıştır. Veriler kategorik olduğunda Mplus 8.3 programı varsayılan olarak WLSMV kestirim yöntemini kullanmaktadır. WLSMV kestirim yöntemi kategorik veya sıralı verileri modellemek amacıyla kullanılır ve bu modellerde verilerin normal dağılıma sahip olma zorunluluğu yoktur (Brown, 2006). Buradan hareketle araştırmada başarı testi verileri kategorik olduğundan ÇGDFA için WLSMV kestirim yöntemi ile THETA parametrelendirmesi tercih edilmiştir. Mplus programı veriler kategorik olduğunda ve WLSMV kestirim yöntemi kullanıldığında yapısal ve ölçek değişmezliği aşamalarını test etmeyi önermektedir (Muthén & Asparouhov, 2002, Muthén & Muthén, 1998-2017). Ölçek faktörlerinin veya hata varyanslarının gruplar arasında değişmesine izin verdiğinden metrik değişmezlik aşamasını test etmemektedir (Muthén & Muthén, 1998-2017, s.542). Putnick ve Bornstein (2016) araştırmasında çoğu ölçme değişmezliği testlerinde yapısal, metrik ve ölçek değişmezlik aşamalarının yer aldığını fakat araştırmalarda metrik değişmezlik aşamasının test edilme yüzdesinin ölçek değişmezliği aşamasının test edilme yüzdesinden daha düşük olduğunu ve bu durumun bazı analizlerin metrik değişmezlik testlerine izin vermemesi veya

bunları önermemesinden kaynaklandığını belirtmiştir. Örneğin Simsek vd. (2021) WLSMV kestirim yöntemi kullandığı araştırmalarında önerildiği üzere metrik değişmezlik aşamasını test etmemiştir. Yapılan açıklamalardan hareketle araştırma kapsamında başarı testinin yapısal ve ölçek değişmezliği aşamaları test edilmiştir.

Kategorik modeller test edilirken WLSMV kestirimi ve Theta parametresi kullanıldığında yapısal değişmezlik aşamasında faktör yükleri ve eşikleri (thresholds) gruplar arasında serbest bırakılmıştır. Hata varyansları tüm gruplarda bire sabitlenmiş ve faktör ortalamaları tüm gruplarda sıfıra sabitlenmiştir. Ölçek değişmezliği modelinde faktör yükleri ve eşikleri gruplar arasında eşit olacak şekilde kısıtlanmıştır. Hata varyansları bir grupta bire sabitlenmiş ve diğer gruplarda serbest bırakılmıştır. Faktör ortalamaları bir grupta sıfıra sabitlenmiş ve diğer gruplarda serbest bırakılmıştır (Muthén & Muthén, 1998-2017, s.542-543).

Duyuşsal Özelliklerle İlgili Oluşturulan Modele İlişkin Analizler

Kayıp Veri. Türkiye, Güney Kore ve Finlandiya veri dosyaları incelendiğinde 80 öğrencinin duyuşsal özellik modelinde yer alan 3 ölçekteki maddelerden ($9 \times 3 = 27$) hiçbirini cevaplamadığı tespit edilmiştir. Sistem hatası olarak belirtilen 80 öğrencinin verisi silinmiş ve veri dosyalarındaki sistem hatası giderilmiştir. Bu durumda 14.561 öğrenci verisi kalmıştır. Silme sonrası kayıp veri (Omitted or invalid yani 9 kodlu cevaplar) oranları 27 madde için 0.7 ile 1.6 aralığında değişmektedir. Kayıp veri oranları oldukça düşük olduğundan kayıp veri analiz yöntemlerinden beklenti maksimizasyonu kullanılarak atama yapılmıştır. Beklenti maksimizasyonu veri setindeki değerlere en yakın değerler üreten, diğer yöntemlere kıyasla performansı daha yüksek olan, en iyi çalışan yöntemlerden biridir (Bernaards & Sijtsma, 2000; Enders, 2013; Şahin Kürşat & Nartgün, 2015).

Uç Değer. Maddelere verilen puanlar toplanarak elde edilen toplam puanlar z puanlarına dönüştürülmüştür. Z puan değerleri -2,09 ile 2,7 aralığında yer almıştır. Geniş örnekleme ($n > 100$) çalışıldığında z puanların ± 4 aralığında olması beklenmektedir (Harrington, 2009). Örneklem büyüklüğü ve elde edilen z puanları belirtilen aralıkta olduğundan uç değer saptanmamıştır.

Normallik. Normallik, basıklık ve çarpıklık değerlerinin hesaplanması ile incelenmiştir.

Tablo 11 incelendiğinde çarpıklık değerlerinin -0.4 ile 1 arasında, basıklık değerlerinin ise -1.3 ile 0.27 arasında değiştiği görülmektedir. Bu durumda normallik sağlanmaktadır (George & Mallery, 2010).

Tablo 11

Duyuşsal Özellik Modelindeki Maddelerin Çarpıklık ve Basıklık Değerleri

	N	Ortalama	Mod	Std. Sapma	Çarpıklık	Basıklık
BSBM16A	14561	2	2	0,96	0,367	-0,822
BSBM16B	14561	2,48	3	1,062	0,002	-1,227
BSBM16C	14561	3	3	1,005	-0,047	-1,078
BSBM16D	14561	2	2	0,94	0,329	-0,803
BSBM16E	14561	2	2	1,035	0,188	-1,124
BSBM16F	14561	3	3	0,991	-0,179	-1
BSBM16G	14561	2	2	1,024	0,032	-1,127
BSBM16H	14561	3	3	0,946	-0,436	-0,737
BSBM16I	14561	3	4	1,113	-0,263	-1,294
BSBM19A	14561	2	2	0,943	0,283	-0,806
BSBM19B	14561	2	2	0,987	0,248	-1,003
BSBM19C	14561	3	3	1,044	-0,108	-1,169
BSBM19D	14561	2	2	0,902	0,224	-0,731
BSBM19E	14561	2	3	0,991	0,006	-1,043
BSBM19F	14561	3	3	0,932	-0,17	-0,863
BSBM19G	14561	3	2	0,96	-0,023	-0,953
BSBM19H	14561	2	1	1,091	0,166	-1,277
BSBM19I	14561	3	3	0,997	-0,106	-1,042
BSBM20A	14561	2	2	0,948	0,562	-0,606
BSBM20B	14561	2	2	0,879	0,53	-0,362
BSBM20C	14561	2	1	0,892	0,89	-0,031
BSBM20D	14561	2	1	0,925	0,671	-0,482

Tablo 11 (devam)

BSBM20E	14561	3	3	1,027	-0,267	-1,069
BSBM20F	14561	2	2	0,932	0,583	-0,588
BSBM20G	14561	2	1	0,858	0,957	0,266
BSBM20H	14561	2	1	0,858	0,901	0,055
BSBM20I	14561	2	1	0,856	0,955	0,19

Çoklu Bağlantı. Maddelere ilişkin yanıtlar sıralı olduğundan maddeler arasındaki korelasyonlar polikorik korelasyon matrisindeki değerler incelenerek belirlenmiştir. Yapılan incelemede 0.90 üzerinde korelasyon saptanmamıştır. Doğrusal ilişkiye sahip olması beklenen madde ikililerinin arasında yüksek ilişki bulunmamış, çoklu bağlantı sorunu tespit edilmemiştir. Bu nedenle 27 madde ile analizlere devam edilmiştir.

Tüm sayıtlar incelenmiş ve gerekli işlemler yapılmıştır. Böylelikle veriler ÇGDFA için hazır hale getirilerek model oluşturma adımlarına geçilmiştir.

Modelin oluşturulması. 27 maddelik duyuşsal özellik model için öncelikle AFA yapılmış, ardından duyuşsal özellik modeli DFA ile doğrulanmıştır.

Açımlayıcı Faktör Analizi. 27 maddelik veri seti ile AFA yapılmıştır. 27 maddelik duyuşsal özellik modelinin KMO = 0.961 ve Bartlett değeri [$\chi^2=269334,149$; $p < .05$] olarak bulunmuştur. KMO değeri 0.60'dan yüksek ve Barlett değeri anlamlı bulunduğundan elde edilen değerler ile verilerin faktörlenebilir olduğu belirlenmiştir.

Özdeğeri 1'den büyük 4 faktör vardır. 1. faktörün açıkladığı varyans %43.038, 2.faktörün açıkladığı varyans %12.013, 3. faktörün açıkladığı varyans %6.280 ve 4. faktörün açıkladığı varyans %4.431'dir. Bu durumda 4 faktörün açıkladığı toplam varyans %65,762'dir. Mplus 8.3 programı ile yapılan 4 faktörlü AFA sonuçlarına göre model uyum değerleri (RMSEA = 0.063, CFI = 0.946, TLI= 0.924, SRMR = 0.024) kabul edilebilir uyum ya da iyi uyum aralığındadır. 4 faktörlü modelde maddelerin faktör yükleri Tablo 12'de verilmiştir:

Tablo 12*Duyuşsal Özellik Modelindeki Maddelerin Faktör Yükleri*

Maddeler	Faktör Yükleri			
	MÖS	MKG1	MDV	MKG2
BSBM16A	0.788*	0.072*	0.045*	-0.003
BSBM16B	0.439*	-0.084*	0.098*	0.403*
BSBM16C	0.679*	-0.089*	0.026*	0.279*
BSBM16D	0.655*	0.038*	0.126*	-0.064*
BSBM16E	0.863*	0.077*	-0.003	0.022*
BSBM16F	0.784*	0.055*	0.016*	-0.091*
BSBM16G	0.723*	0.194*	-0.010*	-0.018*
BSBM16H	0.821*	0.001	-0.032*	0.004
BSBM16I	0.781*	0.135*	-0.044*	0.051*
BSBM19A	0.055*	0.751*	0.076*	0.075*
BSBM19B	-0.162*	0.524*	-0.003	0.417*
BSBM19C	0.080*	0.507*	-0.013*	0.373*
BSBM19D	0.115*	0.691*	0.051*	0.009
BSBM19E	0.043*	0.113*	-0.057*	0.589*
BSBM19F	0.058*	0.774*	0.007	0.001
BSBM19G	0.058*	0.707*	0.050*	-0.044*
BSBM19H	-0.048*	0.321*	0.002	0.663*
BSBM19I	0.055*	0.195*	-0.003	0.668*
BSBM20A	0.281*	-0.051*	0.469*	0.044*
BSBM20B	0.187*	-0.033*	0.502*	0.013
BSBM20C	-0.044*	0.073*	0.757*	-0.014*
BSBM20D	-0.019*	0.039*	0.781*	0.001
BSBM20E	0.303*	0.126*	0.387*	0.053*
BSBM20F	0.103*	-0.061*	0.770*	0.015*
BSBM20G	-0.032*	0.000	0.818*	0.036*
BSBM20H	-0.020*	0.133*	0.599*	-0.128*
BSBM20I	0.072*	0.026*	0.736*	-0.063*

*0.05 düzeyinde anlamlı

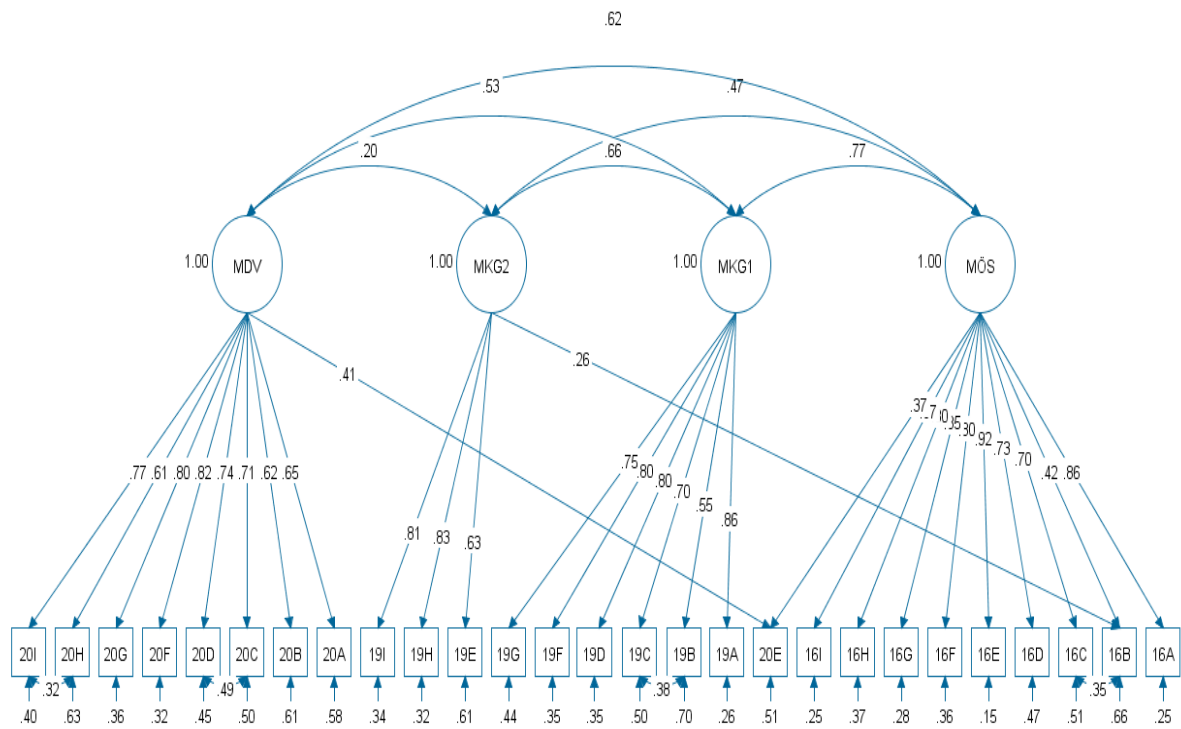
Duyuşsal özellik modelindeki maddelerin faktör yükleri 0.303 ile 0.863 arasında değişmektedir. Duyuşsal özellik modeli 4 faktörlü olarak alınmıştır. Değerler incelendiğinde “Matematik Öğrenmeyi Sevme” ölçeğindeki BSBM16A, BSBM16B, BSBM16C, BSBM16D, BSBM16E, BSBM16F, BSBM16G, BSBM16H, BSBM16I maddelerin bir boyutta (MÖS) ve “Matematiğe Değer Verme” ölçeğindeki BSBM20A, BSBM20B, BSBM20C, BSBM20D, BSBM20E, BSBM20F, BSBM20G, BSBM20H ve BSBM20I maddelerin de bir boyutta (MDV) toplandığı; “Matematikte Kendine Güven” ölçeğindeki maddelerin ise iki boyuta dağıldığı görülmektedir. BSBM19A, BSBM19B, BSBM19C, BSBM19D, BSBM19F ve BSBM19G maddelerinin bir boyutta (MKG1); BSBM19E, BSBM19H ve BSBM19I maddelerinin başka bir boyutta (MKG2) yer aldığı belirlenmiştir. Öğrencilerin benlik kavramını ölçen “Matematikte Kendine Güven” ölçeğindeki MKG1 boyutundaki maddeler öğrencilerin matematik dersinde kendini nasıl algıladığı, kendisine yönelik farkındalığı ile ilgili iken, MKG2 boyutundaki maddelerin ise öğrencilerin matematik dersinde ne hissettiği daha çok duygularına yönelik olduğu düşünülmektedir. Bu nedenle “Matematikte Kendine Güven” ölçeğindeki maddelerin iki faktöre dağılması maddelerin anlamları açısından uygun bulunmuştur. Öte yandan BSBM16B ile BSBM19E maddelerinin iki faktöre yüklenme oranları arasındaki fark 0.10’dan düşüktür. Fakat her iki faktöre de 0.30’dan daha yüksek bir oranla yüklenmişlerdir. Ölçeklerin TIMSS 2019 uygulamasındaki orijinal hallerinden farklılaşmasını önlemek amacıyla BSBM16B ile BSBM19E maddeleri ölçeklerden çıkarılmamış, maddelerin her iki faktöre de yüklenmeleri tercih edilmiştir. Böylelikle 27 madde ile analize devam edilmiştir.

Doğrulayıcı Faktör Analizi. AFA ile ortaya koyulan duyuşsal özellik modelinin doğrulanması amacıyla DFA yapılmıştır. 27 madde 4 faktörlü modelde yapılan analiz sonucunda RMSEA=0.073, CFI=0.909, TLI =0.899 ve SRMR=0.053 değerleri elde edilmiştir. İlgili değerlerde kabul edilebilir aralıklarda olmayan değerler olmasından dolayı model modifikasyon indeksleri incelenmiştir. Mplus programının modifikasyon önerileri sonucu “BSBM16B Keşke matematik çalışmak zorunda olmasam.” ile “BSBM16C Matematik sıkıcıdır”; “BSBM19B Matematik benim için birçok sınıf arkadaşşıma göre daha zordur.” ile “BSBM19C Matematik güçlü yanlarımdan biri değildir.”; “BSBM20C İstedğim üniversiteye

girebilmek için matematikte başarılı olmam gerekir.” ile “BSBM20D İstedğim işe girebilmek için matematikte başarılı olmam gerekir.”; “BSBM20H Ailem matematikte iyi olmamın önemli olduğunu düşünür.” ile “BSBM20I Matematikte başarılı olmak önemlidir.” maddeleri arasında modifikasyon uygulanmıştır. Belirtilen maddeler benzer anlamlarda olduğundan maddelerin benzer şekilde algılanarak cevaplandığı düşünülmüştür. DFA sonuçlarına göre çizilen duyuşsal özelliklerle ilgili oluşturulan modelin yol diyagramı Şekil 2’de sunulmuştur.

Şekil 2

Duyuşsal Özelliklerle İlgili Oluşturulan Modelin Yol Diyagramı



Şekil 2’de görüldüğü üzere 27 madde 4 boyutta toplanmıştır. Maddelere ait faktör yükleri 0.26 ile 0.92 arasında değişmektedir. Elde edilen faktör yüklerinin 0.05 düzeyinde anlamlı olduğu belirlenmiştir. Hata varyansları ise 0.25 ile 0.70 arasındadır. Duyuşsal özellik modelinin veriye uyumunu kontrol etmek için uyum değerleri hesaplanmış ve Tablo 13’te sunulmuştur:

Tablo 13*Duyuşsal Özellik Modeli Uyum Deęerleri*

İndeks	İyi Uyum	Kabul Edilebilir Uyum	Araştırma Deęeri
χ^2	$p > 0.05$	$p > 0.05$	16496.577 ($p < .05$)
χ^2 /sd	$0 \leq \chi^2 /sd / \leq 3$	$3 < \chi^2 /sd \leq 5$	52.87 (16496.577/312)
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 < RMSEA \leq 0.08$	0.060
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI < 0.95$	0.940
TLI	$0.95 \leq TLI \leq 1.00$	$0.90 \leq TLI < 0.95$	0.932
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 < SRMR \leq 0.08$	0.047

* Hu & Bentler (1999)

Modifikasyon sonrası duyuşsal özellik modelinin uyum indeksleri incelendięinde RMSEA, CFI ve TLI deęerlerinin kabul edilebilir, SRMR deęerinin ise iyi uyum aralıęında olduęu görölmektedir. Bu deęerlerden hareketle duyuşsal özellik modelinin veriye uyumunun kabul edilebilir düzeyde olduęu ifade edilebilir.

Çok Gruplu Doğrulatory Faktör Analizi. Duyuşsal özellik modelinin doğrulanmasının ardından modelin seęilen ölkelere (Türkiye, Güney Kore, Finlandiya) göre ölçme deęişmezlięini test etmek için Mplus 8.3 programı ile ÇGDFA yapılmıştır. Veriler Sıralı (kategorik) olduęundan ve normal daęılım saęlandığından Mplus 8.3 programı kullanılarak ML (maksimum olabilirlik) kestirim yöntemi tercih edilmiştir. Bu durumda modeller test edilirken yapısal, metrik ve ölçek deęişmezlięi aşamaları incelenmiştir.

Sıralı (kategorik) deęişkenler için ML kestirim yöntemi kullanıldığında yapısal deęişmezlik aşamasında faktör yükleri ve eşikleri gruplar arasında serbesttir ve faktör ortalamaları tüm gruplarda sifıra sabitlenmiştir. Metrik deęişmezlik aşamasında faktör yükleri gruplar arasında eşit olacak şekilde kısıtlanmış, eşikler gruplar arasında serbest bırakılmış ve faktör ortalamaları tüm gruplarda sifıra sabitlenmiştir. Ölçek deęişmezlik aşamasında faktör yükleri ve eşikler gruplar arasında eşit olacak şekilde kısıtlanmış ve faktör ortalamaları bir grupta sifıra sabitlenirken dięer gruplarda serbest bırakılmıştır (Muthén & Muthén, 1998-2017, s.545-546).

Ölçeklere İlişkin Analizler

Matematik Öğrenmeyi Sevme Ölçeğinin Açıklayıcı Faktör Analizi. 9 maddelik veri seti ile AFA yapılmıştır. Matematik Öğrenmeyi Sevme ölçeğinin KMO = 0.943 ve Bartlett değeri [$\chi^2=102185,326$; $p < .05$] olarak bulunmuştur. KMO değeri 0.60'tan yüksek ve Bartlett testi sonuçları anlamlı bulunduğundan elde edilen verilerin faktör analizi için uygun olduğu sonucuna varılmıştır.

Özdeğeri 1'den büyük tek faktör belirlenmiştir. Tek faktörün açıkladığı varyans %66.845'tir. Matematik Öğrenmeyi Sevme ölçeğinde yer alan maddelerin faktör yükleri Tablo 14'te verilmiştir:

Tablo 14

Matematik Öğrenmeyi Sevme Ölçeğindeki Maddelerin Faktör Yükleri

Maddeler	Faktör Yükleri
BSBM16A	0.863*
BSBM16B	0.547*
BSBM16C	0.711*
BSBM16D	0.728*
BSBM16E	0.924*
BSBM16F	0.798*
BSBM16G	0.846*
BSBM16H	0.799*
BSBM16I	0.865*

*0.05 düzeyinde anlamlı

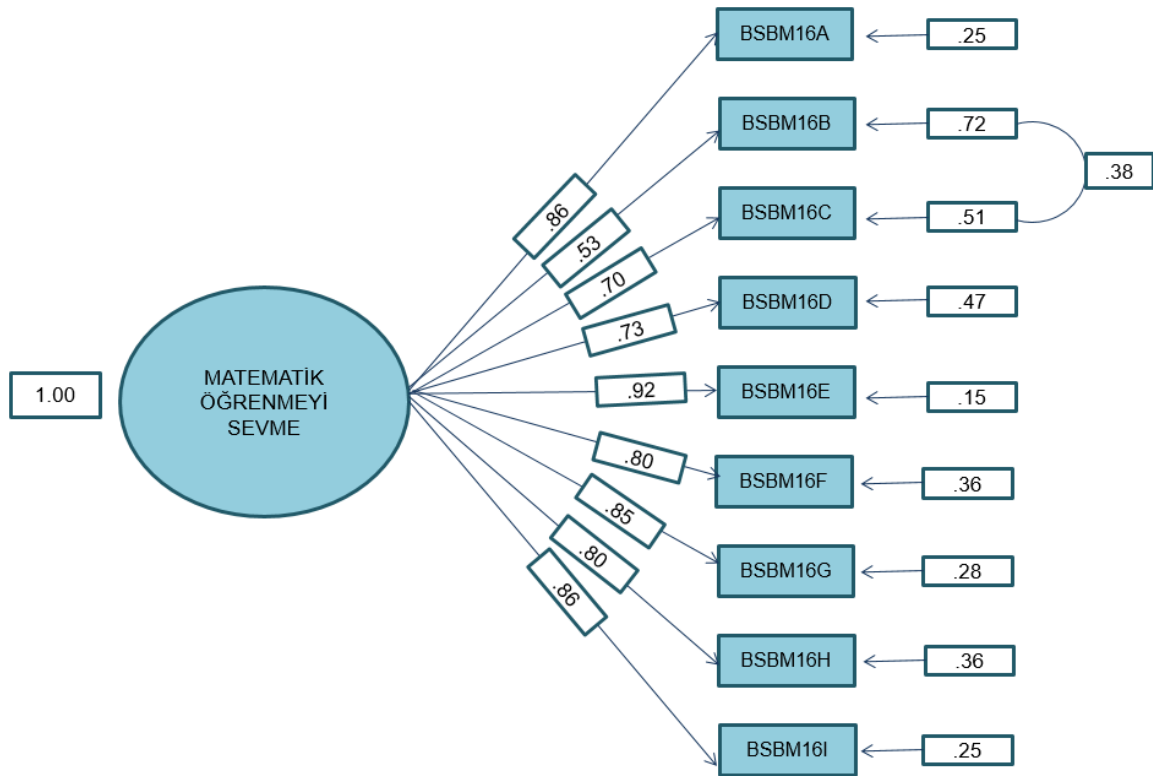
Matematik Öğrenmeyi Sevme ölçeğindeki maddelerin faktör yüklerinin 0.547 ile 0.924 arasında değiştiği ve her bir madde faktör yükünün 0.50'den yüksek olduğu görülmektedir.

Matematik Öğrenmeyi Sevme Ölçeğinin Doğrulayıcı Faktör Analizi. AFA ile ulaşılan değerlerin doğruluğu DFA yapılarak sağlanmıştır. 9 madde tek faktörlü ölçek için yapılan DFA sonucunda RMSEA= 0.107, CFI= 0.956, TLI = 0.941 ve SRMR= 0.035 değerlerine ulaşılmıştır. Söz konusu değerlerde kabul edilebilir aralıkta olmayan değerler

elde edildiğinden modifikasyon önerileri dikkate alınmıştır. Mplus programının modifikasyon önerilerine göre “BSBM16B Keşke matematik çalışmak zorunda olmasam.” ile “BSBM16C Matematik sıkıcıdır” maddeleri benzer anlamlarda olduğundan benzer şekilde algılandığı düşünülmüş ve maddeler arasında modifikasyon uygulanmıştır. Modifikasyon sonrası elde edilen DFA sonuçlarına göre çizilen Matematik Öğrenmeyi Sevme ölçeğine ait yol diyagramı Şekil 3’te verilmiştir:

Şekil 3

Matematik Öğrenmeyi Sevme Ölçeği Yol Diyagramı



Şekil 3’te görüldüğü üzere 9 madde tek boyutta toplanmıştır. Maddelere ait faktör yükleri 0.53 ile 0.92 arasında değişmektedir. Elde edilen faktör yüklerinin 0.05 düzeyinde anlamlı olduğu belirlenmiştir. Hata varyansları ise 0.15 ile 0.72 arasındadır. Matematik Öğrenmeyi Sevme ölçeğine ilişkin uyum değerleri elde edilmiş ve Tablo 15’te verilmiştir:

Tablo 15*Matematik Öğrenmeyi Sevme Ölçeği Uyum Değerleri*

İndeks	İyi Uyum	Kabul Edilebilir Uyum	Araştırma Değeri
χ^2	$p > 0.05$	$p > 0.05$	2395.102 ($p < .05$)
χ^2 /sd	$0 \leq \chi^2 /sd/ \leq 3$	$3 < \chi^2 /sd \leq 5$	92.119 (2395.102/26)
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 < RMSEA \leq 0.08$	0.079
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI < 0.95$	0.977
TLI	$0.95 \leq TLI \leq 1.00$	$0.90 \leq TLI < 0.95$	0.968
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 < SRMR \leq 0.08$	0.019

* Hu & Bentler (1999)

Modifikasyon sonrası Matematik Öğrenmeyi Sevme ölçeğinin uyum indeksleri incelendiğinde RMSEA değerlerinin kabul edilebilir; CFI, TLI ve SRMR değerinin ise iyi uyum aralığında olduğu görülmektedir. Bu değerlerden hareketle Matematik Öğrenmeyi Sevme ölçeğinin veriye uyumunun iyi düzeyde olduğu söylenebilir.

Matematik Öğrenmeyi Sevme Ölçeği'nin Çok Gruplu Doğrulayıcı Faktör Analizi. Mplus 8.3 programı kullanılarak Matematik Öğrenmeyi Sevme Ölçeği'nin seçilen ülkelere (Türkiye, Güney Kore, Finlandiya) göre ölçme değişmezliği ÇG DFA yapılarak belirlenmiştir. Veriler sıralı (kategorik) olduğundan ve normal dağılım sağlandığından kestirim yöntemi olarak ML (maksimum olabilirlik) seçilmiş ve ölçme değişmezliği modellerinden yapısal, metrik ve ölçek değişmezliği aşamaları incelenmiştir.

Matematikte Kendine Güven Ölçeğinin Açıklayıcı Faktör Analizi. 9 maddelik veri seti ile AFA yapılmıştır. Matematikte Kendine Güven ölçeğinin KMO = 0.907 ve Bartlett değeri [$\chi^2 = 69443,261$; $p < .05$] olarak bulunmuştur. KMO değeri 0.60'tan yüksek ve Bartlett testi sonuçları anlamlı çıkmıştır. Böylelikle verilerin faktör analizi için uygunluğu ortaya koyulmuştur.

Özdeğeri 1'den büyük 2 faktör tespit edilmiştir. 1 faktörün açıkladığı varyans %54.415 iken 2 faktörün açıkladığı toplam varyans %68.756'dır. Matematikte Kendine Güven ölçeğinde bulunan maddelerin faktör yükleri Tablo 16'da verilmiştir:

Tablo 16*Matematikte Kendine Güven Ölçeğindeki Maddelerin Faktör Yükleri*

Maddeler	Faktör Yükleri	
	Matematikte Kendine Güven 1	Matematikte Kendine Güven 2
BSBM19A	0.819*	0.064*
BSBM19B	0.245*	0.505*
BSBM19C	0.410*	0.449*
BSBM19D	0.806*	-0.004
BSBM19E	-0.096*	0.706*
BSBM19F	0.808*	0.007
BSBM19G	0.807*	-0.073*
BSBM19H	0.060*	0.775*
BSBM19I	-0.004*	0.793*

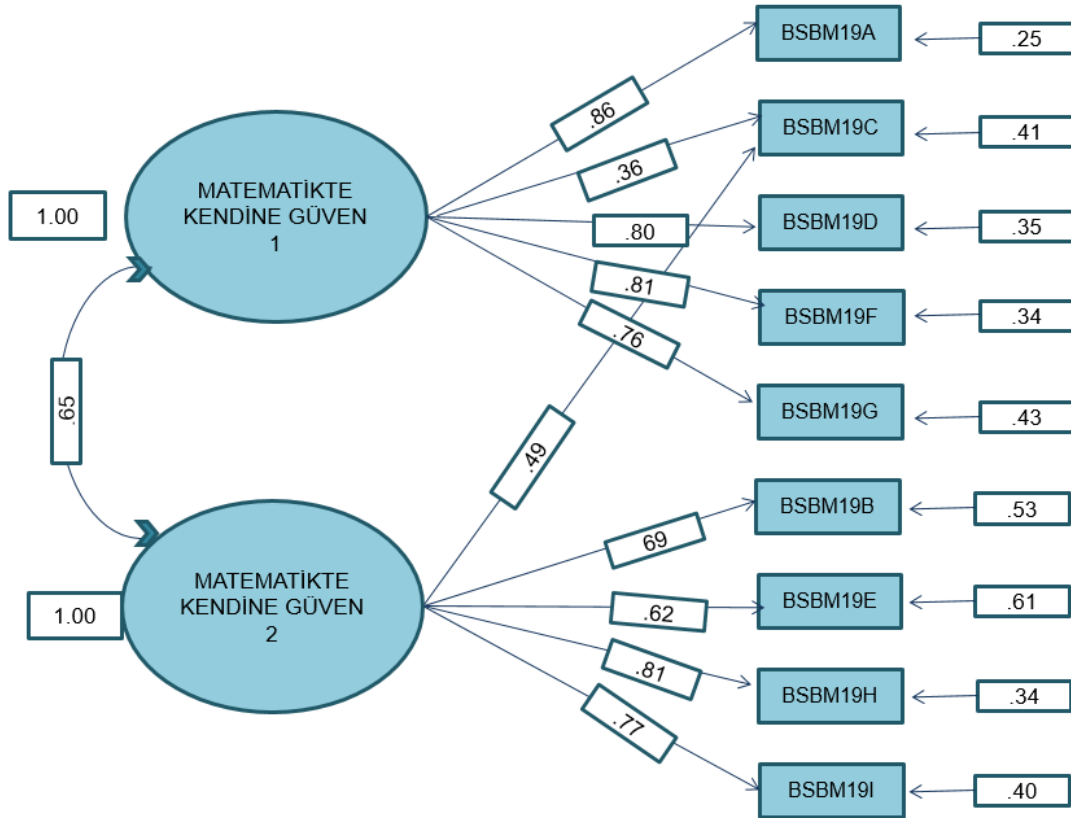
*0.05 düzeyinde anlamlı

Matematikte Kendine Güven ölçeğindeki maddelerin faktör yüklerinin 0.410 ile 0.819 arasında değiştiği ve her bir madde faktör yükünün 0.40'dan yüksek olduğu görülmektedir. Matematikte Kendine Güven ölçeğindeki BSBM19A, BSBM19D, BSBM19F ve BSBM19G maddeleri bir boyuta (MKG1); BSBM19B, BSBM19E, BSBM19H ve BSBM19I maddeleri ise diğer boyuta (MKG2) yüklenmiştir. MKG1 boyutundaki maddeler matematik dersinde öğrencinin kendisini algılama biçimine yönelik iken MKG2 boyutundaki maddeler ise öğrencilerin matematik dersi için ne hissettiği ile ilgilidir. Burada duyuşsal özellik modelinden farklı olarak BSBM19B maddesinin MKG2 faktörüne yüklendiği görülmektedir. Diğer yandan BSBM19C maddesi ise her iki boyuta da 0.30'dan yüksek bir oranla yüklenmiştir. Matematikte Kendine Güven ölçeğinin TIMSS 2019 uygulamasındaki özgün şeklini koruması için BSBM19C maddesi ölçekten çıkarılmamış, her iki faktöre de yüklenmiştir. Bu sayede Matematikte Kendine Güven ölçeği 9 madde olarak analizlere tabi tutulmuştur.

Matematikte Kendine Güven Ölçeğinin Doğrulayıcı Faktör Analizi. AFA ile hesaplanan değerlerin doğruluğunu test etmek için DFA yapılmıştır. DFA sonuçlarına göre çizilen Matematikte Kendine Güven ölçeğinin yol diyagramı Şekil 4'te sunulmuştur:

Şekil 4

Matematikte Kendine Güven Ölçeği Yol Diyagramı



Şekil 4'te görüldüğü üzere 9 madde 2 boyutta yer almıştır. Maddeye ait faktör yükleri 0.36 ile 0.81 arasında değişmektedir. Elde edilen faktör yüklerinin 0.05 düzeyinde anlamlı olduğu belirlenmiştir. Hata varyansları ise 0.25 ile 0.61 arasındadır. Matematikte Kendine Güven ölçeğine ilişkin uyum değerleri elde edilmiş ve Tablo 17'de sunulmuştur:

Tablo 17*Matematikte Kendine Güven Ölçeği Uyum Değerleri*

İndeks	İyi Uyum	Kabul Edilebilir Uyum	Araştırma Değeri
χ^2	$p > 0.05$	$p > 0.05$	2226.502 ($p < .05$)
χ^2 /sd	$0 \leq \chi^2 /sd/ \leq 3$	$3 < \chi^2 /sd \leq 5$	89.060 (2226.502/25)
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 < RMSEA \leq 0.08$	0.078
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI < 0.95$	0.968
TLI	$0.95 \leq TLI \leq 1.00$	$0.90 \leq TLI < 0.95$	0.954
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 < SRMR \leq 0.08$	0.036

* Hu & Bentler (1999)

9 madde 2 faktörlü ölçek için elde edilen DFA değerleri RMSEA= 0.078, CFI= 0.968, TLI = 0.954 ve SRMR= 0.036 şeklindedir. Bu değerlerde RMSEA kabul edilebilir aralıkta iken CFI, TLI ve SRMR değerleri iyi uyum aralığındadır. Bu değerlerden hareketle Matematikte Kendine Güven ölçeğinin veriye uyumunun iyi düzeyde olduğu söylenebilir.

Matematikte Kendine Güven Ölçeğinin Çok Gruplu Doğrulayıcı Faktör Analizi.

Mplus 8.3 programı kullanılarak Matematikte Kendine Güven ölçeğinin seçilen ülkelere (Türkiye, Güney Kore, Finlandiya) göre ölçme değişmezliğini incelemek için ÇGDFA kullanılmıştır. Veriler sıralıdır (kategorik) ve normal dağılım özelliği sağlamaktadır. Bu nedenle ML (maksimum olabilirlik) kestirim yöntemi kullanılmıştır. Yapısal, metrik ve ölçek değişmezliği aşamaları incelenerek ölçeğin ölçme değişmezliği test edilmiştir.

Matematiğe Değer Verme Ölçeğinin Açımlayıcı Faktör Analizi. 9 maddelik veri seti ile AFA yapılmıştır. Matematiğe Değer Verme ölçeğinin KMO = 0.910 ve Bartlett değeri [$\chi^2 = 70521,209$; $p < .05$] olarak bulunmuştur. KMO değeri 0.60'tan yüksek ve Bartlett testi sonuçları anlamlı bulunduğundan verilerin faktör analizi için uygunluğu tespit edilmiştir.

Özdeğeri 1'den büyük 1 faktör belirlenmiştir. 1 faktörün açıkladığı varyans %56.712'dir. Matematiğe Değer Verme ölçeğindeki maddelerin faktör yükleri Tablo 18'de yer almaktadır:

Tablo 18*Matematiğe Değer Verme Ölçeğindeki Maddelerin Faktör Yükleri*

	Maddeler	Faktör Yükleri
A19	BSBM20A	0.616*
A20	BSBM20B	0.601*
A21	BSBM20C	0.761*
A22	BSBM20D	0.787*
A23	BSBM20E	0.631*
A24	BSBM20F	0.803*
A25	BSBM20G	0.796*
A26	BSBM20H	0.638*
A27	BSBM20I	0.784*

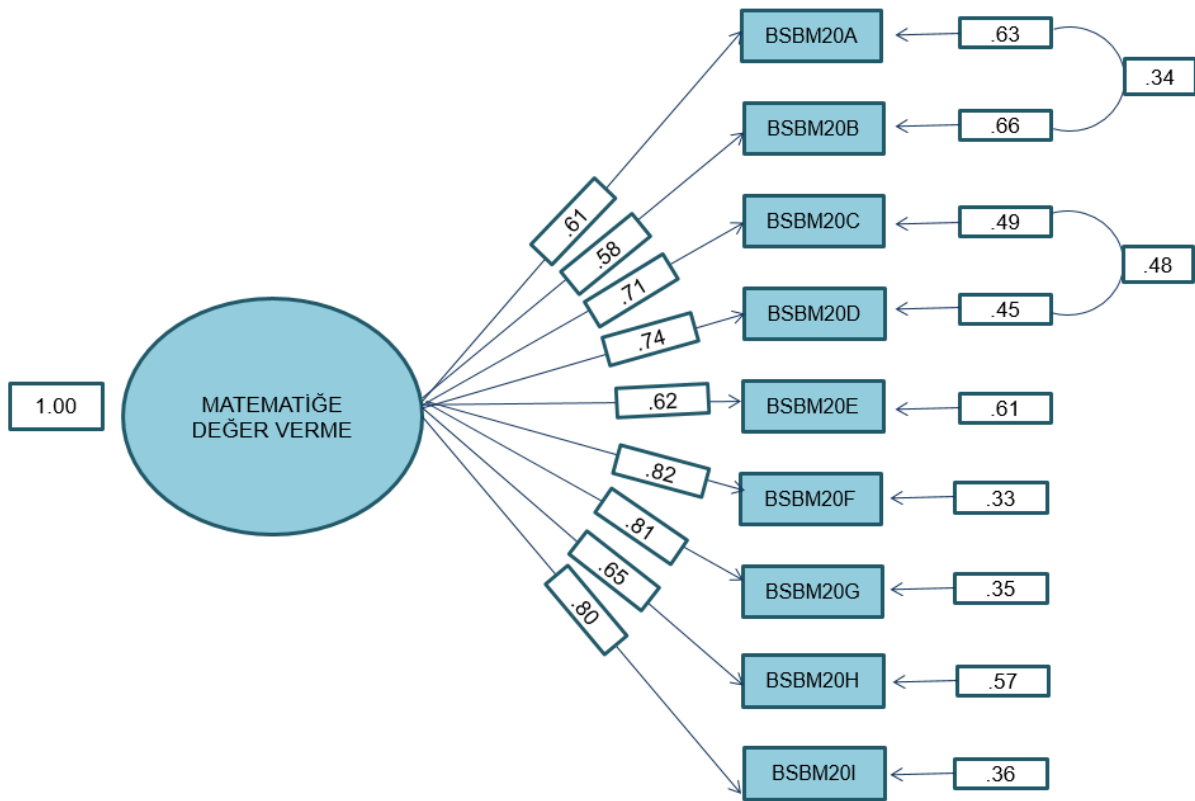
*0.05 düzeyinde anlamlı

Matematiğe Değer Verme ölçeğindeki maddelerin faktör yüklerinin 0.601 ile 0.803 arasında değiştiği ve her bir madde faktör yükünün 0.60'dan yüksek olduğu görülmektedir.

Matematiğe Değer Verme Ölçeğinin Doğrulayıcı Faktör Analizi. AFA ile elde edilen değerler DFA yapılarak doğrulanmıştır. 9 madde tek faktörlü modelde yapılan analiz sonucunda RMSEA=0.138, CFI=0.893, TLI =0.858 ve SRMR=0.050 değerlerine ulaşılmıştır. Elde edilen değerlerin kabul edilebilir aralıklarda olmadığı görülmüştür. Bu nedenle Mplus programının modifikasyon önerileri incelenmiştir. Bunun sonucunda “BSBM20A Matematik öğrenmenin günlük hayatımda bana yardımcı olacağını düşünürüm.” ile “BSBM20B Diğer dersleri öğrenmek için matematiğe ihtiyacım vardır.”; “BSBM20C İstedğim üniversiteye girebilmek için matematikte başarılı olmam gerekir.” ile “BSBM20D İstedğim işe girebilmek için matematikte başarılı olmam gerekir.” maddeleri arasında modifikasyon uygulanmıştır. Belirtilen maddeler benzer anlamlarda olduğundan maddelerin benzer şekilde algılanarak cevaplandığı düşünülmüştür. Ulaşılan sonuçlara göre çizilen Matematiğe Değer Verme ölçeğinin yol diyagramı Şekil 5'te sunulmuştur:

Şekil 5

Matematiğe Değer Verme Ölçeği Yol Diyagramı



Şekil 5'te görüldüğü üzere tek boyutta 9 madde vardır. Maddelere ait faktör yükleri 0.58 ile 0.82 arasında yer almıştır. Elde edilen faktör yüklerinin 0.05 düzeyinde anlamlı olduğu belirlenmiştir. Hata varyansları ise 0.33 ile 0.66 arasındadır. Matematiğe Değer Verme ölçeği için hesaplanan uyum değerleri Tablo 19'da sunulmuştur:

Tablo 19

Matematiğe Değer Verme Ölçeği Uyum Değerleri

İndeks	İyi Uyum	Kabul Edilebilir Uyum	Araştırma Değeri
χ^2	$p > 0.05$	$p > 0.05$	2716.227 ($p < .05$)
χ^2/sd	$0 \leq \chi^2/sd \leq 3$	$3 < \chi^2/sd \leq 5$	108.649 (2716.227 / 25)
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 < RMSEA \leq 0.08$	0.086
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI < 0.95$	0.962
TLI	$0.95 \leq TLI \leq 1.00$	$0.90 \leq TLI < 0.95$	0.945
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 < SRMR \leq 0.08$	0.032

* Hu & Bentler (1999)

9 madde tek faktörlü ölçek için RMSEA=0.086, CFI=0.962, TLI=0.945 ve SRMR=0.032 şeklindedir. Alan yazında RMSEA değeri için net bir standart olmadığı ve bazı araştırmalarda $0.08 \leq RMSEA \leq 0.10$ değerlerinin zayıf uyum olarak belirtildiği görülmüştür (Tabacknick & Fidell, 2013). Bu nedenle söz konusu aralıktaki RMSEA değerleri elde edildiğinde zayıf uyum kabul edilerek analizlere devam edilmiştir. Bu açıklamadan hareketle RMSEA değeri zayıf uyum, TLI değeri kabul edilebilir uyum iken CFI ve SRMR değerleri iyi uyum göstermektedir. Bu değerlerden hareketle Matematikte Kendine Güven ölçeğinin veriye uyumunun kabul edilebilir düzeyde olduğu söylenebilir.

Matematiğe Değer Verme Ölçeğinin Çok Gruplu Doğrulayıcı Faktör Analizi.

Mplus 8.3 programı kullanılarak Matematiğe Değer Verme ölçeğinin seçilen ülkelere (Türkiye, Güney Kore, Finlandiya) göre ölçme değişmezliğini test etmek amacıyla ÇGDFA yapılmıştır. Veriler sıralıdır (kategorik) ve normal dağılım özelliği sağlamaktadır. Bu nedenle ML (maksimum olabilirlik) kestirim yöntemi kullanılmıştır. Yapısal, metrik ve ölçek değişmezliği aşamaları incelenerek ölçeğin ölçme değişmezliği test edilmiştir.

Ölçme değişmezliği ÇGDFA ile test edilirken değişmezlik aşamalarının uyum indeksleri hesaplanır ve modeller arasındaki uyum indeksleri değişimi incelenir. Başarı testinin ölçme değişmezliği modellerinin uyumunun belirlenmesinde değişmezlik aşamalarına ilişkin χ^2 , RMSEA, CFI, TLI, SRMR uyum indeksleri verilmiş ve χ^2 indeksinin değişimi yorumlanmıştır.

Örneklemler büyük olduğunda χ^2 indeksi anlamlı çıktığından model uyumları incelenirken χ^2 uygun bir test olmamaktadır (Kline, 2016). Bu durumlarda χ^2/Sd değeri ve alternatif uyum indekslerinin kullanılması önerilmektedir. Örneklem büyüklüğü etkisinden dolayı duyuşsal özellik modelinde ve ölçeklerde χ^2 ve χ^2/Sd değerleri raporlanmış fakat değerleri yorumlanmamıştır. Öte yandan RMSEA, CFI, TLI gibi uyum indeksleri en sık kullanılan karşılaştırmalı indeksler arasında yer almaktadır. Araştırmalarda modellerin uyumu değerlendirilirken birden fazla uyum indeksinin raporlanması gerektiği önerilmektedir (Crowley & Fan, 1997). Daha önceden belirtildiği gibi bu araştırmada uyum indeksleri olarak

RMSEA, CFI, TLI ve SRMR indeksleri incelenmiş ve bu değerlerin kabul edilebilir aralıkta olup olmadıkları değerlendirilmiştir. Alan yazında iç içe geçmiş modellerin arasındaki uyum indeksleri değişimleri incelenirken ΔCFI ve ΔTLI değerlerinin kullanılması tavsiye edilmektedir (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000). Bu araştırmada da ölçme değişmezliğinin tespiti için ölçme değişmezliği modellerinin arasındaki farklılıklar ΔCFI , ΔTLI değerleri sayesinde belirlenmiştir. Söz konusu uyum indeksleri farkları $-0.01 \leq \Delta CFI \leq 0.01$ ve $-0.01 \leq \Delta TLI \leq 0.01$ aralığında olduğunda ilgili değişmezlik aşamasında koşulların sağlandığı kabul edilmiş (0.014 değeri 0.01 değerine yuvarlanmıştır), elde edilen fark belirtilen aralığın dışında kaldığında ise ilgili değişmezlik aşamasında değişmezlik koşulunun sağlanmadığına karar verilmiştir (Cheung & Resvold, 2002).

Bölüm 4

Bulgular, Yorumlar ve Tartışma

Bu bölümde araştırma kapsamında ele alınan başarı testinin, duyuşsal özelliklerle ilgili oluşturulan üç ölçeğin bir arada bulunduğu modelin ve ölçeklerin her birinin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği analizi bulgularına ve yorumlarına yer verilmiştir. ÇGDFA yöntemi ile yapılan analizlerde duyuşsal özelliklerin incelenmesinde yapısal, metrik ve ölçek değişmezlik aşamaları, başarı testinde ise yapısal ve ölçek değişmezlik aşamaları dikkate alınmıştır. Ölçme değişmezliğinin test edilme süreci aşamalı olduğundan değişmezliğin sağlanmadığı aşamada test analizleri sonlandırılmıştır.

Araştırmanın 1. Alt Problemine İlişkin Bulgular

Araştırmanın birinci alt problemi “TIMSS 2019 uygulamasında yer alan 8.sınıf matematik başarı testinin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği sağlanmakta mıdır?” şeklindedir. Bu alt problemi yanıtlamak amacıyla öncelikle başarı testinde Türkiye, Güney Kore ve Finlandiya örneklemi için ayrı ayrı DFA yapılmış, böylelikle her bir ülke için doğrulama işlemi gerçekleştirilmiştir. Ardından üç ülkeye ait veriler birlikte kullanılarak ÇGDFA yapılmıştır. DFA yapılarak elde edilen ülkelere göre başarı testinin uyum değerleri Tablo 20’de verilmiştir:

Tablo 20

Ülkelere Göre Başarı Testinin Uyum Değerleri

Ülkeler	χ^2	sd	χ^2/sd	RMSEA	CFI	TLI	SRMR
Türkiye	580.634 p< .05	434	1.34	0.034	0.980	0.978	0.075
Güney Kore	506.233 p< .05	434	1.17	0.025	0.989	0.989	0.069
Finlandiya	493.443 p< .05	434	1.14	0.020	0.980	0.978	0.086

Tablo 20’ye göre tüm ülke örneklemine ilişkin RMSEA, TLI ve CFI değerleri iyi uyum aralığında iken Türkiye ve Güney Kore örneklemi için SRMR değeri kabul edilebilir

uyum aralığında fakat Finlandiya örnekleme için SRMR değeri kabul edilebilir aralığa yakın değerdedir. Elde edilen değerlerden hareketle her bir ülke için başarı testinin ülke verileri ile uyumu belirlenmiş ve başarı testi doğrulanmıştır.

Başarı testinin doğrulanmasının ardından ülkelere göre ÇGDFA yapılmıştır. Ülkelere göre ÇGDFA sonucunda hesaplanan uyum değerleri Tablo 21’de verilmiştir.

Tablo 21

Ülkelere Göre Başarı Testinin ÇGDFA Uyum Değerleri

Ölçme Değişmezliği Aşamaları	χ^2	<i>Sd</i>	χ^2/Sd	RMSEA	CFI	Δ CFI	TLI	Δ TLI	SRMR
Yapısal Değişmezlik	1556.408 p< .05	1302	1.20	0.025	0.985	-	0.984	-	0.075
Ölçek Değişmezliği	2091.026 p< .05	1360	1.54	0.042	0.956	0.029	0.955	0.029	0.080
Ölçek Değişmezliği – Yapısal Değişmezlik	581.487 p< .05	58							

Ölçme değişmezliğinin ilk aşaması olan yapısal değişmezlik aşamasında başarı testinin uyumunun belirlenmesinde ulaşılan RMSEA, CFI ve TLI değerlerinin iyi uyum aralığında ve SRMR değerinin kabul edilebilir aralıkta olduğu görülmüştür (RMSEA= .025, CFI= .985, TLI= .984 SRMR= .075). Bu durum üç ülke örnekleme için faktör yapısının yani faktör sayısının ve her bir faktördeki madde sayısının aynı olduğu anlamındadır. Her bir ülke örneklemindeki öğrencilerin modelde yer alan maddeleri aynı kavramsal çerçeveden hareketle cevapladıkları belirtilebilir.

Daha sonra ölçek değişmezliği aşaması test edilmiştir. Ölçek değişmezlik aşamasının test edilmesinde ülke örneklemeleri arasında faktör yüklerinin ve madde sabitlerinin eşit olma sınırlılığı vardır. Ulaşılan değerlerden SRMR değerinin kabul edilebilir uyum aralığında diğerlerinin ise iyi uyum aralığında olduğu anlaşılmıştır (RMSEA= .042, CFI= .956, TLI= .955, SRMR= .080). Ölçek değişmezliği aşaması için elde edilen uyum değerleri önerilen sınırlar içinde kalsa da, yapısal değişmezlik aşamasına göre uyumda önemli bir düşüş meydana gelmiştir ($\Delta \chi^2(58) = 581.487, p < .05$). Bu durum ölçek değişmezlik aşamasının, verilere

yapısal değişmezlik aşamasından daha kötü uyum sağlaması nedeniyle ölçek değişmezliğinin kabul edilemeyeceği anlamına gelmektedir.

Başarı testinin ülke örneklerinde ölçek değişmezliğini sağladığına dair bir kanıt ulaşılamadığından ve değişmezlik aşamaları arasında bir hiyerarşi olduğundan katı değişmezlik aşaması test edilmemiştir.

Uluslararası çapta yapılan geniş ölçekli araştırmalara yönelik ülkelere göre ölçme değişmezliği incelemeleri yapılmıştır. Alan yazında metrik değişmezlik aşamasını sağlayan fakat ölçek değişmezliği aşamasını sağlamayan araştırmalar vardır. PISA 2015 matematik okuryazarlığı testinin farklı dilleri konuşan ülkelere göre ölçme değişmezliğini inceleyen Asar'ın (2019) araştırması bu duruma örnektir. Öncü (2019) araştırmasında ise TIMSS 2015 uygulamasına katılan sekizinci sınıf öğrencilerinin matematik başarılarının ülkelere göre ölçme değişmezliğinin en iyi güçlü faktöriyel değişmezlik modelinde çalıştığını belirlemiştir. Benzer sonucu Tavlıca (2019) araştırmasında TIMSS 2015 uygulamasında kullanılan dördüncü sınıf matematik başarı testinin OECD üyesi ülkelere göre test edilmesinde bulmuştur. Ayvalı (2016) araştırmasında ise benzer sonuca PISA 2012 uygulamasında yer alan matematik okuryazarlığı testinin OECD üyesi ülkelere göre ölçme değişmezliğini test edilmesinde ulaşmıştır. Gönen (2021) araştırmasında PISA 2018 uygulamasındaki okuma becerileri testinin belli üniteleri kapsayan ülkelere göre ölçme değişmezliği incelemesinde bir ünitenin yapısal değişmezlik, iki ünitenin metrik değişmezlik ve iki ünitenin de ise ölçek değişmezliği aşamasını sağladığını belirlemiştir. Öte yandan Kıbrıslıođlu (2015) PISA 2012 matematik öğrenme modelinin ülkelere göre ölçme değişmezliğini incelediđi araştırmasında ölçme değişmezliğinin sadece yapısal değişmezlik aşamasında sağlandığı sonucuna ulaşmıştır. Karakoç Alatlđ (2016) ise PISA 2012 uygulamasındaki okuryazarlık testlerinin farklı dil gruplarına göre ölçme değişmezliğini incelediđi araştırmada matematik, fen okuryazarlığı ile okuma becerileri alt testlerinin yapısal değişmezlik aşamasını sağladığını fakat metrik değişmezliği aşamasını sağlamadığını tespit etmiştir. Yapılan alan yazın taramasında TIMSS, PISA gibi uygulamalarda yer alan başarı testlerinin ülkelere göre ölçme değişmezliğinin incelendiđi araştırmalarda katı değişmezlik aşamasının sağlandığına dair

kanıtlara ise rastlanmamıştır. Başarı testlerinin farklı değişkenlere (cinsiyet, bölge vb.) göre ölçme değişmezliğinin incelendiği araştırmalarda ise değişmezlik aşamalarının farklı aşamalarına kanıt sağlandığı görülmektedir. Örneğin Gönen (2021) araştırmasında PISA 2018 uygulamasındaki okuma becerileri testinin belli üniteleri kapsayan Türkiye örneklemindeki cinsiyete göre ölçme değişmezliği incelemesinde üç ünitenin katı değişmezlik, iki ünitenin ise ölçek değişmezliği aşamasını sağladığını belirlemiştir. Gönen'in (2021) araştırmasında ülkelere göre elde ettiği ölçme değişmezliği aşamalarının ise daha alt aşamalarda kaldığı dikkat çekmektedir. Bu durum herhangi bir ölçme aracı için ülkelere göre elde edilen ölçme değişmezliği aşamasının, belli bir ülke örnekleminde yapılan cinsiyete ve bölgeye göre ölçme değişmezliği aşamasından daha düşük olduğu şeklinde yorumlanabilir. Bu yorumun ölçülen yapının doğası, ölçme aracındaki maddeler, çalışılan örneklemelerin özellikleri gibi çeşitli faktörlerin sınırlılığı dâhilinde yapılabileceği söylenebilir.

Araştırmanın 2. Alt Problemine İlişkin Bulgular

Araştırmanın ikinci alt problemi "TIMSS 2019 uygulamasında yer alan 8.sınıf öğrenci anketi kapsamında matematik alanına yönelik hazırlanan duyuşsal özelliklerle ilgili oluşturulan üç ölçeğin bir arada bulunduğu modelin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği sağlanmakta mıdır?" şeklindedir. Bu alt problemi yanıtlamak amacıyla öncelikle duyuşsal özellik modelinde Türkiye, Güney Kore ve Finlandiya örneklemi için ayrı ayrı DFA yapılmış, ilgili modelin ülkeler için doğrulanması kontrol edilmiştir. Ardından üç ülkeye ait veriler bir arada alınarak ÇGDFA yapılmıştır. DFA yapılarak elde edilen ülkelere göre duyuşsal özellik modelinin uyum değerleri Tablo 22'de verilmiştir:

Tablo 22*Ülkelere Göre Duyuşsal Özellik Modelinin Uyum Değerleri*

Ülkeler	χ^2	sd	χ^2 / Sd	RMSEA	CFI	TLI	SRMR
Türkiye	5485.753 p< .05	312	17.583	0.060	0.930	0.922	0.052
Güney Kore	6292.996 p< .05	312	20.170	0.066	0.931	0.922	0.058
Finlandiya	7371.555 p< .05	312	23.627	0.064	0.933	0.925	0.042

Tablo 22'e göre tüm ülke örneklemlerine ilişkin RMSEA, TLI ve CFI değerleri kabul edilebilir uyum aralığında iken SRMR değeri Türkiye ve Güney Kore örneklemleri için kabul edilebilir uyum aralığında, Finlandiya örneklemleri için SRMR değeri ise iyi uyum aralığındadır. Elde edilen bulgular ışığında her bir ülke örneklemleri için modelin ülke verileri ile uyumu belirlenmiş ve duyuşsal özellik modeli doğrulanmıştır.

Duyuşsal özellik modelinin doğrulanmasının ardından ülkelere göre ÇGDFA yapılarak tüm ülke verileri aynı anda kullanılarak ölçme değişmezliği sınanmıştır. Ülkelere göre ÇGDFA sonucunda hesaplanan uyum değerleri Tablo 23'te verilmiştir.

Tablo 23*Ülkelere Göre Duyuşsal Özellik Modelinin ÇGDFA Uyum Değerleri*

Ölçme Değişmezliği Aşamaları	χ^2	Sd	χ^2 / Sd	RMSEA	CFI	ΔCFI	TLI	ΔTLI	SRMR
Yapısal Değişmezlik	19150.305 p< .05	936	20.460	0.063	0.932	-	0.923	-	0.051
Metrik Değişmezlik	21625.118 p< .05	986	21.932	0.066	0.923	0,009	0.917	0,006	0.065
Ölçek değişmezliği	27911.273 p< .05	1032	27.046	0.073	0.899	0,024	0.897	0,020	0.071

Ölçme değişmezliğinin ilk aşaması olan yapısal değişmezlik aşamasında modele ilişkin faktör yükleri, faktörler arası korelasyonlar ve hata varyansları her bir ülke örneklemleri için serbest bırakılmaktadır. Model uyumunun belirlenmesinde ulaşılan değerlerin kabul

edilebilir uyum aralığında olduğu görülmüştür (RMSEA= .063, CFI= .932, TLI= .923 SRMR= .051). Bulgulardan hareketle modelin her bir ülke örnekleminde aynı faktör yapısında olduğu söylenebilir. Bu durum üç ülke örneklemini için faktör sayısının ve her bir faktördeki madde sayısının aynı olduğu anlamındadır. Her bir ülke örneklemindeki öğrencilerin modelde yer alan maddeleri aynı kavramsal çerçeveden hareketle cevapladıkları belirtilebilir.

Yapısal değişmezlik aşamasının sağlanmasının ardından ölçme değişmezliğinin bir sonraki aşaması olan metrik değişmezlik aşaması test edilmiştir. Bu değişmezlik aşamasında ülke örneklemleri arasında faktör yüklerinin eşit olma sınırlılığı vardır. Elde edilen değerlerin kabul edilebilir uyum aralığında olduğu belirlenmiştir (RMSEA= .066, CFI= .923, TLI= .917, SRMR= .065). Değişmezlik aşamalarının karşılaştırılması için yapısal değişmezlik ile metrik değişmezlik aşamalarında elde edilen CFI değerleri arasındaki fark $\Delta CFI= .009$ ve TLI değerleri arasındaki fark $\Delta TLI= .006$ şeklinde hesaplanmış ve ulaşılan farkın kabul edilebilir aralıkta olduğu görülmüştür ($-.01 < \Delta CFI < .01$, $-.01 < \Delta TLI < .01$) Modelin ülke örneklemleri arasında metrik değişmezliğini sağladığı tespit edilmiştir. Bu durum maddelerin faktör yüklerinin ülke örneklemleri arasında eşdeğer olduğunu göstermektedir. Buradan hareketle üç ülke örneklemindeki öğrencilerin de modeldeki maddeleri aynı şekilde yorumlandığı söylenebilir.

Ölçek değişmezliği aşamasında ise madde faktör yüklerine ek olarak madde sabitlerinin de ülke örneklemleri arasında eşit olma sınırlılığı vardır. Bu aşamanın test edilmesinde elde edilen RMSEA= .073 ve SRMR= .071 değerlerinin kabul edilebilir aralıkta, CFI= .899 ve TLI= .897 değerlerinin kabul edilebilir aralığa yakın olduğu tespit edilmiştir. Metrik değişmezlik ile ölçek değişmezliği aşamalarında elde edilen CFI değerleri arasındaki fark $\Delta CFI= .024$ ve TLI değerleri arasındaki fark $\Delta TLI= .020$ şeklinde bulunmuştur. Elde edilen bu farklar kabul edilebilir aralıkta ($-.01 < \Delta CFI < .01$, $-.01 < \Delta TLI < .01$) olmadığından model ülke örneklemleri arasında ölçek değişmezliğini sağlamamaktadır. Maddeler için oluşturulan regresyon denklemlerindeki madde sabitlerinin değişmez olduğu doğrulanamamıştır. Bu durumda maddelere ilişkin faktör yüklerinin ülkeler arasında değişmez olduğu fakat faktörler arası korelasyonların değişmezlik göstermediği ifade

edilebilir. Dolayısıyla ülkelerden biri(leri) lehine yanlılık olabileceği söylenebilir. Sonuç olarak veri setinde metrik değişmezlik sağlandığından ülke örneklemeleri sadece faktör yükleri açısından karşılaştırılabilir, madde sabitleri açısından karşılaştırılmaz. Dolayısıyla ülke örneklemelerinin ortalama puanlarını karşılaştırmanın anlamlı olmayacağı söylenebilir.

Duyuşsal özellik modelinin ülke örneklerimde ölçek değişmezliğini sağladığına dair bir kanıt ulaşılamadığından ve değişmezlik aşamaları arasında bir hiyerarşi olduğundan katı değişmezlik aşaması test edilmemiştir.

Alan yazında duyuşsal özelliklere ilişkin ölçeklerin bir araya getirilerek oluşturulduğu modellerin ölçme değişmezliğinin bu araştırmada olduğu gibi metrik değişmezlik aşamasında kaldığı yani ölçek değişmezliği aşamasının sağlanmadığı araştırmalar vardır. Örneğin Marsh vd. (2013) araştırmasında TIMSS 2017 uygulamasındaki ölçeklerle matematik ve fen motivasyon ölçeği (kendine güven, olumlu etki ve değer) tasarlamış ve ölçeğin Arapça ve İngilizce konuşan ülkelere göre ölçme değişmezliğinin metrik değişmezlik aşamasında kaldığını belirlemiştir. Benzer şekilde TIMSS 2011 verilerini kullanan Bofah ve Hannula (2015) araştırmasında motivasyonel yapılarla ilişkili kurduğu modelin beş Afrika ülkesine göre ölçme değişmezliği testinde metrik değişmezlik aşamasının sağladığını tespit etmiştir. Öte yandan Polat (2019) araştırmasında TIMSS 2015 uygulamasında yer alan ölçekleri kullanarak matematik ve fen duyuşsal özellik modelleri oluşturmuş ve her iki modelin ülkelere göre ölçme değişmezliği ölçek değişmezliği aşamasını sağlamıştır. Bunun yanı sıra Polat (2019) araştırmasında oluşturduğu modelin cinsiyete ve bölgelere göre de ölçme değişmezliğini test etmiş ve bölgelere göre ölçek değişmezliği, cinsiyete göre ise katı değişmezliğin sağlandığı sonucuna ulaşmıştır. Benzer şekilde Demir (2020) araştırmasında TIMSS 2015 uygulamasındaki fen duyuşsal özelliklere ilişkin kurduğu modelin Türkiye örneğinde cinsiyete ve bölgelere göre ölçme değişmezliğini incelemiş ve modelin cinsiyete göre katı değişmezlik aşamasını sağladığını, bölgelere göre ölçek değişmezliği aşamasını sağladığını belirlemiştir. Uyar (2021) araştırmasında ise TIMSS 2015 uygulamasına yönelik olan matematik tutum anketinin (matematiğe ilgisi, öğretime ilişkin görüşler, öğrencilerin matematiğe ilişkin özgüveni ve öğrencilerin matematiğe verdiği değer)

Türkiye örnekleminde cinsiyete göre ölçme değişmezliği testinde yapısal, metrik, ölçek, katı, varyans-kovaryans ve örtük ortalama değişmezliği olmak üzere altı değişmezlik aşamasının da sağlandığı sonucuna varmıştır. Söz konusu araştırmalarda duyuşsal özelliklere ilişkin ölçeklerin bir araya getirilerek oluşturulduğu modellerin ülkelere göre ölçme değişmezliğinin, bir ülke örnekleminde cinsiyet, bölge vb. değişkenlere göre yapılan ölçme değişmezliğine kıyasla daha alt aşamalarda kaldığı söylenebilir. Özellikle kültürel ve bağlamsal farklılıkların etkili olduğu durumlarda, ülkeler arasında ölçme değişmezliği sağlamak daha zor olabilir. Bu yoruma benzer olarak Asil ve Gelbal (2012) ülkeler arasında yaptığı ölçme değişmezliği araştırmasında ÇGDFA sonucunda maddelerde DMF varlığını belirleyerek dil ve kültür çeşitliliği arttıkça DMF içeren maddelerin sayısında artış olduğunu ortaya koymuştur. Ülke, cinsiyet veya bölgeler gibi farklı gruplar arasında ölçme değişmezliğinin sağlanması bağlama bağlı olup ölçülen yapının doğası, ölçme aracındaki maddeler, çalışılan örneklemelerin özellikleri gibi çeşitli faktörlerden etkilenebileceği ifade edilebilir.

Araştırmanın 3. Alt Problemine İlişkin Bulgular

Araştırmanın üçüncü alt problemi “TIMSS 2019 uygulamasında yer alan ve 8.sınıf öğrenci anketi kapsamında öğrencilerin matematik alanındaki duyuşsal özelliklerine yönelik hazırlanan ölçeklerin her birinin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliği sağlanmakta mıdır?” şeklindedir. Bu alt problemi yanıtlamak maksadıyla her bir ölçek için ayrı ayrı analizler yapılmıştır. Bu kapsamda öncelikle Türkiye, Güney Kore ve Finlandiya örneklemi için ölçek bazında DFA yapılmış, ilgili ölçeğin ülkeler için doğrulanması sağlanmıştır. Sonrasında üç ülkeye dair veriler birlikte kullanılarak ÇGDFA yapılmıştır.

Matematik Öğrenmeyi Sevme Ölçeği

DFA yapılarak elde edilen Matematik Öğrenmeyi Sevme ölçeğinin ülkelere göre uyum değerleri Tablo 24’te verilmiştir:

Tablo 24*Ülkelere Göre Matematik Öğrenmeyi Sevme Ölçeğinin Uyum Değerleri*

Ülkeler	χ^2	sd	χ^2 / sd	RMSEA	CFI	TLI	SRMR
Türkiye	946.610 p< .05	26	33.028	0.087	0.966	0.953	0.026
Güney Kore	728.209 p< .05	26	28.008	0.078	0.980	0.972	0.016
Finlandiya	901.390 p< .05	26	34.67	0.078	0.977	0.968	0.020

Tablo 24'e göre tüm ülke örneklemine ilişkin TLI, CFI ve SRMR değerleri iyi uyum aralığında iken Güney Kore ve Finlandiya örneklemeleri için RMSEA değerleri kabul edilebilir aralıkta, Türkiye örnekleme için RMSEA değeri zayıf uyum aralığındadır. Hesaplanan bu değerlerden ölçeğin ülke verileri ile uyumu ortaya koyulmuş ve ölçek doğrulanmıştır. Matematik Öğrenmeyi Sevme ölçeğinin doğrulanmasından sonra ülkelere göre ÇGDFA yapılmıştır. Ülkelere göre ÇGDFA yapılarak elde edilen uyum değerleri Tablo 25'te verilmiştir.

Tablo 25*Ülkelere Göre Matematik Öğrenmeyi Sevme Ölçeğinin ÇGDFA Uyum Değerleri*

Ölçme Değişmezliği Aşamaları	χ^2	Sd	χ^2 / Sd	RMSEA	CFI	ΔCFI	TLI	ΔTLI	SRMR
Yapısal Değişmezlik	2576.209 p< .05	78	33.03	0.081	0.975	-	0.965	-	0.021
Metrik Değişmezlik	3730.280 p< .05	94	39.68	0.089	0.963	0.012	0.958	0.007	0.074
Ölçek değişmezliği	6819.220 p< .05	110	61.99	0.112	0.933	0.030	0.934	0.024	0.089

Yapısal değişmezlik aşamasında ölçeğin uyumunun tespitinde elde edilen CFI, TLI ve SRMR değerlerinin iyi uyum aralığında ve RMSEA değerinin kabul edilebilir aralıkta olduğu görülmüştür (RMSEA= .081, CFI= .975, TLI= .965 SRMR= .021). Bu bulgular, her bir ülke örnekleme için ölçeğin aynı faktör yapısına uygun olduğunu göstermektedir. Bu durum faktör sayısının ve her bir faktördeki madde sayısının üç ülke örnekleme için de aynı olduğu

anlamına gelir. Dolayısıyla, her ülkedeki öğrencilerin, ölçekte yer alan maddeleri aynı kavramsal çerçeveden yola çıkarak yanıtladıkları söylenebilir.

Yapısal değişmezlik aşamasının başarıyla tamamlanmasının ardından ölçme değişmezliğinin bir sonraki aşaması olan metrik değişmezlik test edilmiştir. Bu aşamada, ülkeler arasında faktör yüklerinin eşit olma sınırlılığı dikkate alınmıştır. Elde edilen CFI ve TLI değerleri iyi uyum gösterirken RMSEA ve SRMR değerleri de kabul edilebilir uyum aralığında bulunmuştur (RMSEA= .089, CFI= .963, TLI= .958, SRMR= .074). Değişmezlik aşamalarının karşılaştırılması için yapısal değişmezlik ile metrik değişmezlik aşamalarında elde edilen CFI değerleri arasındaki fark $\Delta CFI = .012$ ve TLI değerleri arasındaki fark $\Delta TLI = .007$ şeklinde hesaplanmış, bu farkın kabul edilebilir aralıkta olduğu görülmüştür ($-.01 < \Delta CFI < .01$, $-.01 < \Delta TLI < .01$). Bu bulgu, ölçeğin ülkeler arasında metrik değişmezliğini başarıyla sağladığını ortaya koymaktadır. Bu durumda, ölçeğin ülke örneklemelerinde faktör yüklerinin eşit olduğu ve ülkelerdeki öğrencilerin maddeleri benzer şekilde anladığı ifade edilebilir.

Ölçek değişmezliği aşamasında elde edilen değerlendirmeler RMSEA= .0112 ve SRMR= .089 değerlerinin kabul edilebilir aralıkta olmadığını, ancak CFI= .933 ve TLI= .934 değerlerinin kabul edilebilir aralıkta olduğunu göstermiştir. Ayrıca, metrik değişmezlik ile ölçek değişmezliği aşamalarında elde edilen CFI değerleri arasındaki fark $\Delta CFI = .030$ ve TLI değerleri arasındaki fark $\Delta TLI = .024$ olarak bulunmuştur. Bu farklar, kabul edilebilir aralıkta değildir ($-.01 < \Delta CFI < .01$, $-.01 < \Delta TLI < .01$). Bu nedenle ölçek değişmezliği ülkeler arasında sağlanmamaktadır. Bu durumda, maddelere ilişkin faktör yüklerinin ülkeler arasında değişmez olduğu, ancak faktörler arası korelasyonların değişmezlik göstermediği ifade edilebilir. Bu bulgular, veri setinde metrik değişmezliğin sağlandığını bu nedenle ülkelerin sadece faktör yükleri açısından karşılaştırılabileceğini, ancak madde sabitleri açısından karşılaştırılamayacağını göstermektedir. Dolayısıyla ülkelerden biri(leri) lehine yanlılık olabileceği, ülke örneklemeleri ortalama puanlarının karşılaştırılmasının anlamlı olmadığı belirtilebilir. Matematik Öğrenmeyi Sevme ölçeğinin ülke örneklerimde ölçek değişmezliğini sağladığına dair bir kanıtı ulaşılamadığından ve değişmezlik aşamaları arasında bir hiyerarşi olduğundan katı değişmezlik aşaması test edilmemiştir.

Matematikte Kendine Güven Ölçeği

DFA yapılarak elde edilen Matematikte Kendine Güven ölçeğinin ülkelere göre uyum değerleri Tablo 26'da verilmiştir:

Tablo 26

Ülkelere Göre Matematikte Kendine Güven Ölçeğinin Uyum Değerleri

Ülkeler	χ^2	<i>Sd</i>	χ^2 / Sd	RMSEA	CFI	TLI	SRMR
Türkiye	408.733 p< .05	25	16.35	0.058	0.982	0.974	0.024
Güney Kore	1208.705 p< .05	25	48.32	0.104	0.948	0.925	0.057
Finlandiya	839.491 p< .05	25	33.58	0.077	0.972	0.959	0.028

Tablo 26'ya göre Türkiye ve Finlandiya örneklemi için CFI, TLI ve SRMR değerleri iyi uyum aralığında iken, Güney Kore örneğine ait RMSEA değeri zayıf uyum aralığındadır. Üç ülke örneği için diğer uyum değerleri kabul edilebilir aralıktadır. Hesaplanan bu değerlerden ölçeğin ülke örneklemi verileri ile uyumu ortaya koyulmuş ve ölçek doğrulanmıştır. Matematikte Kendine Güven ölçeğinin doğrulanmasından sonra ülkelere göre ÇGDFA yapılmıştır. Ülkelere göre ÇGDFA yapılarak elde edilen uyum değerleri Tablo 27'de verilmiştir.

Tablo 27

Ülkelere Göre Matematikte Kendine Güven Ölçeğinin ÇGDFA Uyum Değerleri

Ölçme Değişmezliği Aşamaları	χ^2	<i>Sd</i>	χ^2 / Sd	RMSEA	CFI	Δ CFI	TLI	Δ TLI	SRMR
Yapısal Değişmezlik	2456.92 9 p< .05	75	32.76	0.081	0.967	-	0.953	-	0.038
Metrik Değişmezlik	3469.06 1 p< .05	91	38.12	0.087	0.953	0.014	0.945	0.008	0.071
Ölçek değişmezliği	5879.72 4 p< .05	105	56.00	0.106	0.920	0.033	0.918	0.027	0.076

Yapısal değişmezlik aşamasında ölçek için elde edilen uyum değerleri şu şekildedir: RMSEA= .081, CFI= .967, TLI= .953 ve SRMR= .038. Bu bulgular, her bir ülke örneklemini için ölçeğin aynı faktör yapısında olduğunu ortaya koymaktadır, yani faktör sayısı ve her bir faktördeki madde sayısı üç ülke örnekleminde aynıdır. Öyleyse her bir ülke örneklemindeki öğrencilerin ölçekte yer alan maddeleri aynı kavramsal çerçeveden hareketle yanıtladıkları belirtilebilir.

Yapısal değişmezlik aşamasının başarılı bir şekilde tamamlanmasının ardından ölçme değişmezliğinin bir sonraki aşaması olan metrik değişmezlik incelenmiştir. Bu aşamada, ülke örneklemleri arasında faktör yüklerinin eşit olma sınırlılığına odaklanılmıştır. CFI değeri iyi uyum aralığında, TLI ve SRMR değerleri kabul edilebilir iyi uyum aralığında ancak RMSEA değeri zayıf uyum aralığındadır (RMSEA= .087, CFI= .953, TLI= .945, SRMR= .071). Değişmezlik aşamalarının karşılaştırılması için yapısal değişmezlik ile metrik değişmezlik aşamalarında elde edilen CFI değerleri arasındaki fark $\Delta CFI = .014$ ve TLI değerleri arasındaki fark $\Delta TLI = .008$ olarak hesaplanmıştır. Bu farkların kabul edilebilir aralıkta olması ($-.01 < \Delta CFI < .01$, $-.01 < \Delta TLI < .01$) ölçeğin ülke örneklemleri arasında metrik değişmezliğini sağladığını göstermektedir. Bu durumda, ölçeğin ülke örneklemlerinde faktör yüklerinin eşit olduğu ve ülkelerdeki öğrencilerin maddeleri benzer şekilde anladığı ifade edilebilir.

Ölçek değişmezliği aşamasında ise RMSEA= .106 değerinin kabul edilebilir aralıkta olmadığı ancak CFI= .940, TLI= .939 ve SRMR=0.076 değerlerinin kabul edilebilir aralıkta olduğu görülmüştür. Ayrıca, metrik değişmezlik ile ölçek değişmezliği aşamalarında elde edilen CFI değerleri arasındaki farkın ($\Delta CFI = .033$) ve TLI değerleri arasındaki farkın ($\Delta TLI = .027$) kabul edilebilir aralıkta olmadığı tespit edilmiştir. Elde edilen değerlerden ölçek değişmezliğinin ülke örneklemleri arasında sağlanmadığı anlaşılmıştır. Bu noktada metrik değişmezlik sağlandığından maddelere ilişkin faktör yüklerinin ülke örneklemleri arasında değişmez olduğu fakat ölçek değişmezliği sağlanmadığından faktörler arası korelasyonların değişmezlik göstermediği söylenebilir. Dolayısıyla ülkeler arasında faktör yüklerinin karşılaştırılabileceği fakat madde sabitlerinin karşılaştırılamayacağı ifade edilebilir.

Dolayısıyla ülkelerden biri(leri) lehine yanlılık olabileceği, ülke örneklemi ortalama puanlarının karşılaştırılmasının anlamlı olmadığı belirtilebilir. Matematikte Kendine Güven ölçeğinin ülke örneklerimde ölçek değişmezliğini sağladığına dair bir kanıt ulaşılamadığından ve değişmezlik aşamaları arasında bir hiyerarşi olduğundan katı değişmezlik aşaması test edilmemiştir.

Matematiğe Değer Verme Ölçeği

DFA yapılarak elde edilen Matematiğe Değer Verme ölçeğinin ülkelere göre uyum değerleri Tablo 28'de verilmiştir:

Tablo 28

Ülkelere Göre Matematiğe Değer Verme Ölçeğinin Uyum Değerleri

Ülkeler	χ^2	sd	χ^2/sd	RMSEA	CFI	TLI	SRMR
Türkiye	597.962 p< .05	25	23.92	0.070	0.967	0.953	0.031
Güney Kore	1192.841 p< .05	25	47.71	0.103	0.945	0.921	0.046
Finlandiya	1390.346 p< .05	25	55.61	0.099	0.953	0.932	0.036

Tablo 28'e göre tüm ülke örneklemine ilişkin SRMR değerleri iyi uyum aralığındadır. Türkiye örneğine ait CFI ve TLI değerleri ile Finlandiya örneğine ait CFI değeri iyi uyum fakat belirtilen uyum indekslerinin diğer değerleri kabul edilebilir uyum aralığındadır. RMSEA değerleri ise Türkiye örneğinde kabul edilebilir uyum, Güney Kore ve Finlandiya örneğinde zayıf uyum göstermektedir. Hesaplanan bu değerlerden ölçeğin ülke örneklemi verileri ile uyumu belirlenmiş ve ölçek doğrulanmıştır. Ükelere göre ÇGDFA yapılarak elde edilen uyum değerleri Tablo 29'da verilmiştir:

Tablo 29*Ülkelere Göre Matematiğe Değer Verme Ölçeğinin ÇG DFA Uyum Değerleri*

Ölçme Değişmezliği Aşamaları	χ^2	<i>Sd</i>	χ^2 / Sd	RMSEA	CFI	ΔCFI	TLI	ΔTLI	SRMR
Yapısal Değişmezlik	3181.148 p< .05	75	42.42	0.092	0.954		0.934		0.038
Metrik Değişmezlik	3812.431 p< .05	91	41.89	0.092	0.945	0.009	0.935	0.009	0.061
Ölçek Değişmezliği	6632.580 p< .05	107	61.99	0.112	0.904	0.041	0.903	0.032	0.081

Yapısal değişmezlik aşamasında Matematiğe Değer Verme ölçeği için elde edilen RMSEA= .092, CFI= .954, TLI= .934 ve SRMR= .038 değerleri, her bir ülke örneklemini için ölçeğin aynı faktör yapısına sahip olduğunu göstermektedir. Faktör sayısı ve her bir faktördeki madde sayısı üç ülke örneklemini için de aynıdır. Bu durumda, her bir ülke örneklemindeki öğrencilerin, ölçekte yer alan maddelere aynı kavramsal çerçeveden hareketle yanıt verdikleri ifade edilebilir.

Metrik değişmezlik aşaması, yapısal değişmezlik aşamasının başarılı bir şekilde tamamlanmasının ardından gerçekleştirilmiştir. CFI, TLI ve SRMR değerleri genel olarak kabul edilebilir uyum aralığında, ancak RMSEA değeri zayıf uyum aralığındadır (RMSEA= .092, CFI= .945, TLI= .935, SRMR= .061). Yapısal değişmezlik ve metrik değişmezlik aşamalarındaki CFI ve TLI değerleri arasındaki farklar, sırasıyla $\Delta CFI = .009$ ve $\Delta TLI = .009$ olarak hesaplanmıştır. Ulaşılan farklar kabul edilebilir aralıkta ($-.01 < \Delta CFI < .01$, $-.01 < \Delta TLI < .01$) bulunmaktadır. Dolayısıyla ölçek ülke örneklemleri arasında metrik değişmezliğini sağlamaktadır. Bu durumda ölçeğin ülke örneklemlerinde faktör yüklerinin eşit olduğu ve ülkelerdeki öğrencilerin maddeleri benzer şekilde anladığı söylenebilir.

Ölçek değişmezliği aşamasında elde edilen değerlere göre RMSEA= .112 değeri kabul edilebilir aralıkta olmadığı halde, CFI= .904, TLI= .903 ve SRMR=0.081 değerleri kabul edilebilir aralıkta bulunmaktadır. Ayrıca, metrik değişmezlik ile ölçek değişmezliği aşamalarındaki CFI değerleri arasındaki farkın ($\Delta CFI = .041$) ve TLI değerleri arasındaki farkın ($\Delta TLI = .032$) kabul edilebilir aralıkta olmadığı gözlemlenmiştir. Elde edilen bu değerler,

ölçek değişmezliğinin ülke örneklemi arasında sağlanmadığını ortaya koymaktadır. Metrik değişmezlik aşaması sağlandığından faktörlere ait yüklerin ülkeler arasında değişmez olduğu görülmüştür. Ancak ölçek değişmezliğinin sağlanmadığı tespit edilmiştir. Bu durumda, faktörler arasındaki korelasyonların değişmezlik göstermediği belirtilebilir. Dolayısıyla ülkelerden biri(leri) lehine yanlılık olabileceği, ülke örneklemi ortalama puanlarının karşılaştırılmasının anlamlı olmadığı belirtilebilir. Matematiğe Değer Verme ölçeğinin ülke örneklerimde ölçek değişmezliğini sağladığına dair bir kanıt ulaşılamadığından ve değişmezlik aşamaları arasında bir hiyerarşi olduğundan katı değişmezlik aşaması test edilmemiştir.

Alan yazında bu araştırmaya benzer şekilde duyuşsal özelliklerle ilişkili ölçeklerin ölçme değişmezliğinin metrik değişmezlik aşamasında durduğu, ölçek ve katı değişmezlik aşamalarına ulaşamadığı gözlemlenen araştırmalar bulunmaktadır. Örneğin, He vd. (2018) farklı ülkelerdeki PISA 2015 ve TIMSS 2015 uygulamalarında fen dersiyle ilgili duyuşsal özellik ölçeklerinin (fen dersini sevme, fen dersine atfedilen değer ve öğrencilerin okula aidiyet duygusu) ölçme değişmezliğini araştırmıştır. Elde ettikleri bulgular sadece metrik değişmezlik aşamasının karşılandığını göstermiştir. Atılgan (2022) araştırmasında ise TIMSS 2019 uygulamasında yer alan matematiği sevme, matematikte kendine güven, matematiğe değer verme ve okula ait olma duygusu ölçeklerinin ölçme değişmezliğini Türkiye örneğinde cinsiyete göre incelemiştir. Araştırma sonucunda matematiği sevme ve matematiğe değer verme ölçeklerinin katı değişmezlik aşamasını sağladığını ve okula ait olma duygusu ölçeğinin ölçek değişmezliği aşamasını sağladığını bulmuştur. Atılgan (2022) matematikte kendine güven ölçeği DFA ile doğrulanamadığından ölçeği kapsam dışında bırakmıştır. Ertürk & Oyar (2021) araştırmasında ise TIMSS 2015 uygulamasında kullanılan matematiği sevme ölçeğinin aynı ve farklı dilleri konuşan ülkelere göre ölçme değişmezliğini farklı kuramsal temellere dayanan modellere göre incelemiştir. Araştırmasının sonucu farklı teorik temellere sahip modellerin farklı ölçme değişmezliği sonuçları verdiğini göstermektedir.

Bölüm 5

Sonuç ve Öneriler

Araştırmanın bu bölümünde alt problemlere ilişkin sonuçlar verilmiş ve bu sonuçlara dayalı geliştirilen öneriler sunulmuştur.

Sonuçlar

Bu araştırmada TIMSS 2019 uygulamasında yer alan 8.sınıf matematik başarı testinin, 8.sınıf öğrenci anketi kapsamında öğrencilerin matematik alanındaki duyuşsal özelliklerine yönelik hazırlanan üç ölçeğin bir arada bulunduğu modelin ve ölçeklerin her birinin Türkiye, Güney Kore ve Finlandiya örneklemi için ölçme değişmezliğini incelemek amaçlanmıştır. Bu kapsamda 14 no'lu kitapçıktaki matematik başarı testi ve duyuşsal özelliklerle ilgili "Matematik Öğrenmeyi Sevme", "Matematikte Kendine Güven" ve "Matematiğe Değer Verme" ölçekleri ele alınmıştır. Ülkeler arasındaki ölçme değişmezliğini incelemek için öncelikle her bir başarı testi, model ve ölçekler için ülkelere göre ayrı ayrı DFA yapılarak ilgili başarı testi, model ve ölçekler ülke örneklemi için doğrulanmıştır. Ardından başarı testi, model ve ölçeklerde ülke örneklemine göre ÇGDFA yapılmıştır. Ulaşılan sonuçlar şu şekildedir:

1. Başarı testinin tek faktörlü olduğu belirlenmiş ve elde edilen yapı DFA sonuçlarıyla doğrulanmıştır. Bu sebeple başarı testinin Türkiye, Güney Kore ve Finlandiya örneklemi için geçerli olduğu söylenebilir. Başarı testi için yapılan ÇGDFA sonucunda başarı testinin ölçek değişmezliği sağladığına ilişkin kanıta ulaşılammıştır. Sonuç olarak ülke örneklemelerinin ortalama puanlarının karşılaştırılmasının doğru olmayacağı, ülkelerden en az biri lehine yanlılık olabileceği söylenebilir.

2. "Matematik Öğrenmeyi Sevme", "Matematikte Kendine Güven" ve "Matematiğe Değer Verme" olmak üzere üç ölçek bir arada alınarak duyuşsal özellik modeli oluşturulmuştur. Söz konusu model, üç ölçekten oluşmasına rağmen modelin dört faktörlü olduğu belirlenmiştir. Bu durum DFA sonuçları ile doğrulanmıştır. Modelin ülke örneklemi

için geçerli olduğu ortaya koyulmuştur. Model için yapılan ÇGDFA sonucunda modelin metrik değişmezlik sağladığına ilişkin kanıtlara ulaşılabılırken ölçek değişmezliği sağladığına ilişkin kanıta ulaşılamamıştır. Sonuç olarak üç ülke örneklemindeki öğrencilerin de modeldeki maddeleri aynı şekilde anladığı, ülke örneklemlerinin sadece faktör yükleri açısından karşılaştırılabileceği, ülke örneklemlerinin ortalama puanlarını karşılaştırmanın doğru olmayacağı ifade edilebilir.

3. “Matematik Öğrenmeyi Sevme”, “Matematiğe Değer Verme” ölçekleri tek faktörlü “Matematikte Kendine Güven” ölçeği ise iki faktörlü belirlenmiş ve elde edilen yapılar DFA sonuçlarıyla doğrulanmıştır. Üç ölçeğin bir arada ele alındığı modelde de “Matematikte Kendine Güven” ölçeğinde yer alan maddeler iki faktörde toplanmıştır. Bu durumda ölçekler tek alınsa da birkaç ölçek bir arada alınsa da oluşacak faktör sayısının aynı olma eğilimine sahip olduğu söylenebilir. Fakat model ve tek ölçek incelemesinde iki faktör altında toplanan maddelerin birebir aynı olmadığı da görülmüştür. “Matematik benim için birçok sınıf arkadaşına göre daha zordur.” maddesi model ve tek ölçek incelemesinde aynı ölçek içinde farklı faktörlere yüklenmiştir. Diğer yandan model ve tek ölçek incelemesinde önerilen modifikasyonların bazılarının farklı maddeler için olduğu ifade edilebilir. Ölçekler için yapılan ÇGDFA sonuçlarında ölçeklerin metrik değişmezlik sağladıklarına ilişkin kanıtlara ulaşılabılırken ölçek değişmezliği sağladıklarına ilişkin kanıta ulaşılamamıştır. Sonuç olarak üç ülke örneklemindeki öğrencilerin de her bir ölçekteki maddeleri aynı şekilde anladığı, ülke örneklemlerinin sadece faktör yükleri açısından karşılaştırılabileceği, ülke örneklemlerinin ortalama puanlarını karşılaştırmanın doğru olmayacağı belirtilebilir.

Sonuç olarak duyuşsal özelliklerle ilgili “Matematik Öğrenmeyi Sevme”, “Matematiğe Değer Verme” ile “Matematikte Kendine Güven” ölçeklerini kapsayan modelde, her bir ölçeğin ayrı ayrı incelenmesinde ve başarı testinde ölçme değişmezliğinin 'metrik değişmezlik' aşamasını destekleyen kanıtlar elde edilmiştir. Ancak, 'ölçek değişmezliği' aşamasına ilişkin kanıt sağlanamamıştır. Bu sebeple ölçme değişmezliği ile ilgili yapılacak karşılaştırmalarda ortaya çıkacak farklılıkların kaynağının, ülkelerden mi yoksa ölçme

araçlarının kendisinden mi kaynaklandığını kesin bir şekilde belirtmek güçtür. Dolayısıyla ülkelerle ilgili yorumda bulunurken bu durumu ihmal etmemek gerekir.

Öneriler

Araştırma sonuçlarından hareketle elde edilen öneriler araştırmacılara ve uygulayıcılara yönelik olarak ayrı başlıklar altında aşağıda verilmiştir.

Araştırmacılara Yönelik Öneriler

1. Araştırmada örneklem büyüklüğünün etkisinden dolayı DFA yapılırken χ^2 uyum iyiliği indeksinin anlamlılığı ve ÇGDFA yapılırken değişmezlik aşamalarını karşılaştırmada kullanılan $\Delta \chi^2$ değerleri duyuşsal özelliklerle ilgili ölçeklerde ve modelde yorumlanmamıştır. Örneklem büyüklüğünün etkisini azaltmak için benzer araştırmalar daha küçük örneklem seçilerek tekrarlanabilir.
2. Araştırmada başarı testinin, duyuşsal özelliklerle ilgili oluşturulan üç ölçeğin bir arada bulunduğu modelin ve ölçeklerin her birinin ülkelere göre ölçme değişmezliğinin ölçek değişmezliği aşamasını sağlamadığı belirlendiğinden ülkelerden biri(leri) lehine yanlılık olabileceği belirtilmiştir. Bu nedenle yanlılıkların hangi madde(ler)den kaynaklandığının belirlenmesi amacıyla başarı testi, duyuşsal özellik modeli ve ölçeklerin MTK'ye dayalı DMF test edilebilir.
3. Araştırmada üç ölçeğin bir arada bulunduğu model ile ölçeklerin her birinin ölçme değişmezliğinin test edilmesinde aynı değişmezlik aşamasında sonlandığı ayrıca elde edilen sonuçların faktör sayısı bağlamında benzer fakat faktörler altında toplanan maddeler ve modifikasyon önerileri bağlamında farklılıklar olduğu ortaya koyulmuştur. Bu noktada TIMSS 2019 teknik raporunda bağlam anketleri kapsamında yer alan ölçeklerin oluşturulması ve yorumlanması bölümünde ölçeklerin ayrı ayrı analiz edildiği de göz önüne alınarak araştırmacılara ölçekler bazında analiz yapmaları tavsiye edilebilir.

4. Araştırmada Matematikte Kendine Güven ölçeğinin yapısı hem modelde hem ölçek bazında iki faktörlü olarak belirlenmiştir. Oysaki her bir ölçeğin tek faktörlü yapıda olması beklenmektedir. Elde edilen sonuçların beklentiye uygun olmadığı göz önüne alındığında Matematikte Kendine Güven ölçeğine dayalı yorumlar yapılırken dikkat edilmesi yararlı olabilir. Ayrıca söz konusu ölçeğin sekizinci ve dördüncü sınıf fen bilimleri ve dördüncü sınıf matematik verileri kullanılarak ölçme değişmezliği test edilebilir.
5. Araştırmada başarı testinin, duyuşsal özelliklerle ilgili oluşturulan üç ölçeğin bir arada bulunduğu modelin ve ölçeklerin her birinin ülkelere göre ölçme değişmezliği test edilmiştir. Araştırmacılar başarı testi, duyuşsal özellik modeli ve ölçeklerin başka değişkenlere göre ölçme değişmezliğini inceleyebilir.
6. Alan yazında genellikle başarı testi ya da ölçeklerin verileri kullanılarak en az bir değişkene göre ölçme değişmezlikleri incelenmektedir. Bu araştırmada ise hem başarı testi hem de ölçek verileri kullanılarak tek bir değişkene (ülke) göre ölçme değişmezliği incelenmiştir. Analiz sonuçlarında her bir ölçme değişmezliğinin aynı ölçme değişmezliği aşamasında kaldığı belirlenmiştir. Buradan hareketle farklı araçların tek değişkene göre ölçme değişmezliği yapılarak ulaşılan değişmezlik aşamaları karşılaştırılabilir.
7. Farklı teorik temellere sahip yöntemlerin farklı ölçme değişmezliği sonuçları verdiği alan yazında görülmektedir. Bu nedenle araştırmada kullanılan veri toplama araçlarının ölçme değişmezliği farklı kuramsal temellere dayanan yöntemlerle incelenebilir (Çok Gruplu LCA ve Karma Rasch Modeli gibi).
8. Araştırma kapsamında üç ülke örneklemini karşılaştırılarak ölçme değişmezliği test edilmiştir. Ülkeler ikili gruplar şeklinde karşılaştırılarak ülkeler arasında ölçme değişmezlikleri test edilebilir.

Uygulayıcılara Yönelik Öneriler

1. Uluslararası çapta düzenlenen araştırmalarda katılımcı ülkeler arasında karşılaştırmalar yapılırken dikkatli bir metodoloji benimsenmelidir. Bu tür

karşılaştırmalarda, farklılıklar üzerinden yapılan yorumlar bilimsel temellere dayandırılmalıdır. Bu temel, veri setlerinin doğru bir şekilde analiz edilmesi, katılımcı ülkelerin eğitim sistemlerindeki özgünlüklerin anlaşılması ve sonuçların objektif bir şekilde yorumlanmasıyla oluşturulabilir.

2. Karşılaştırma yapılmadan önce kullanılacak başarı testi veya ölçekler şeklindeki veri toplama araçlarının ölçme değişmezliği detaylı bir şekilde incelenmelidir. Bu süreç, veri toplama araçlarının güvenilirliği ve geçerliği üzerine odaklanmalıdır. Testin yapısından kaynaklanan farklılıkların belirlenmesi, testin doğru bir şekilde karşılaştırmaya uygun olup olmadığını değerlendirmeyi içermelidir.

Kaynaklar

- Asar, E. (2019). *PISA 2015 Matematik Okur-Yazarlığı Testinin Farklı Dilleri Konuşan Ülkeler Arasında Ölçme Değişmezliğinin İncelenmesi* (Yüksek Lisans Tezi). Akdeniz Üniversitesi, Antalya.
- Asil, M. & Gelbal, S. (2012). PISA öğrenci anketinin kültürler arası eşdeğerliği. *Eğitim ve Bilim*, 37(166).
- Atılğan, M. (2022). *TIMSS 2019 Matematik Başarısına İlişkin Duyuşsal Özelliklerin Cinsiyete Göre Ölçme Değişmezliğinin İncelenmesi* (Yüksek Lisans Tezi). Ankara Üniversitesi, Ankara.
- Aydın, M. (2020). Eğitim sisteminde uluslararası değerlendirmelerin yeri. *Eğitime Bakış*, 49, 1-10.
- Ayvallı, M. (2016). *PISA 2012 Matematik Okuryazarlığı Testinin Ölçme Değişmezliğinin İncelenmesi*. (Yüksek Lisans Tezi). Akdeniz Üniversitesi, Antalya.
- Bağdu Söyler, P., Aydın, B., & Atılğan, H. (2021). PISA 2015 reading test item parameters across language groups: A measurement invariance study with binary variables. *Journal of Measurement and Evaluation in Education and Psychology*, 12(2); 112-128.
- Başusta, N. B. & Gelbal, S. (2015). Gruplar arası karşılaştırmalarda ölçme değişmezliğinin test edilmesi: PISA öğrenci anketi örneği. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30(4), 80-90.
- Bernaards, C. A., & Sijtsma, K. (2000). Influence of imputation and EM methods on factor analysis when item nonresponse in questionnaire data is nonignorable. *Multivariate Behavioral Research*, 35(3), 321–364.
- Bloom, B. S. (2012). *İnsan Nitelikleri ve Okulda Öğrenme* (D. A. Özçelik, Çev.). Pegem Akademi.

- Bofah, E. A., & Hannula, M. S. (2015). TIMSS data in an African comparative perspective: Investigating the factors influencing achievement in 48 mathematics and their psychometric properties. *Large Scale Assessment in Education*, 3(4), 1-35.
- Bollen, K. A. (1989). *Structural equation modeling with latent variables* (2nd ed.). Wiley.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. The Guilford Press.
- Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. *Handbook of structural equation modeling*, 361, 379.
- Bryman, A., & Cramer, D. (2001). *Quantitative data analysis with SPSS release 10 for Windows*. Routledge.
- Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, 34(2), 155–175.
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Eğitim Yönetimi Dergisi*, 32, 470-483.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464-504. doi.org/10.1080/10705510701301834
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. doi.org/10.1207/S15328007sem0902_5
- Creswell, J. W. & Plano Clark, V. L. (2014). *Karma yöntem araştırmaları: Tasarımı ve yürütülmesi* (Çev. Ed. Y. Dede ve S. B. Demir). Anı Yayıncılık.
- Crowley, S. L., & Fan, X. (1997). Structural equation modeling: Basic concepts and applications in personality assessment research. *Journal of personality assessment*, 68(3), 508-531. https://doi.org/10.1207/s15327752jpa6803_4

- Çakıcı Eser, D. (2021). Investigation of measurement invariance according to home resources: TIMSS 2015 mathematical affective characteristics questionnaire. *International Journal of Assessment Tools in Education*, 8(3), 633–648.
- Çokluk, Ö., Şekercioğlu, G. & Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok değişkenli istatistik SPSS ve Lisrel uygulamaları*. Pegem Akademi.
- Demir, E. (2017). Testing measurement invariance of the students' affective characteristics model across gender sub-groups. *Educational Sciences: Theory & Practice*, 17, 47–62.
- Demir, M. C. (2020). *TIMSS 2015 Fen Duyuşsal Özelliklerinin Cinsiyet ve Bölgelere Göre İncelenmesi* (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Demir, S. (2022). Comparison of normality tests in terms of sample sizes under different skewness and kurtosis coefficients, *International Journal of Assessment Tools in Education*, 9(2), 397–409.
- Doğan, M. (2013). *Doğrulayıcı Faktör Analizinde Örneklem Hacmi, Tahmin Yöntemleri ve Normalliğin Uyum Ölçütlerine Etkisi* (Yüksek Lisans Tezi). Osmangazi Üniversitesi, Eskişehir.
- Doğan, N., & Barış, F. (2010). Tutum, değer ve özyeterlik değişkenlerinin TIMSS-1999 ve TIMSS-2007 sınavlarında öğrencilerin matematik başarılarını yordama düzeyleri. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1), 44-50.
- Enders, C.K. (2013). Dealing with missing data in developmental research. *Child Development Perspectives*, 7(1), 27- 31.
- Ertürk, Z., & Erdinç Akan, O. (2018). TIMSS 2015 matematik başarısını etkileyen değişkenlerin yapısal eşitlik modeli ile incelenmesi. *Ulusal Eğitim Akademisi Dergisi (UEAD)*, 2(2), 14-34.

- Ertürk, Z., & Oyar, E. (2021). Examining the measurement invariance of TIMSS 2015 mathematics liking scale through different methods. *International Journal of Assessment Tools in Education*, 8(1), 67-89.
- Fishbein, B., Foy, P., & Yin, L. (2021). *Timss 2019 User Guide For The International Database* (2nd Edition). Boston College, TIMSS & PIRLS International Study Center. <https://timss2019.org/international-database/downloads/TIMSS-2019-User-Guide-for-the-International-Database-2nd-Ed.pdf>
- Foy, P., Fishbein, B., von Davier, M., & Yin, L. (2020). Implementing the TIMSS 2019 Scaling Methodology. In Martin, M.O., von Davier, M., & Mullis, I.V.S. (Eds.), *Methods and procedures: TIMSS 2019 technical report* (pp. 12.1-12.61). Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/methods/chapter-12.html>
- George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Study Guide and Reference*, 17.0 update (10a ed.) Pearson.
- Gönen, E. (2021). PISA 2018 *Okuma Becerileri Testinin Ülkelere ve Cinsiyete Göre Ölçme Değişmezliğinin İncelenmesi* (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, 44(11), 78-94. doi.org/10.1097/01.mlr.0000245454.12228.8f
- Grek, S. (2009). Governing by Numbers: the PISA 'Effect' in Europe. *Journal of Education Policy*, 24(1), 23-37.
- Güngör, M., & Atalay Kabasakal, K. (2020). Investigation of measurement invariance of science motivation and self-efficacy model: PISA 2015 Turkey sample. *International Journal of Assessment Tools in Education*, 7(2), 207–222.

- Gürten, E., Demirkaya, A.S., & Doğan, N. (2019). Uzmanların PISA ve TIMMS sınavlarının eğitim politika ve programlarına etkisine ilişkin görüşleri. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 52, 287-319.
- Hansson, A., & Gustafsson, J-E. (2013). Measurement invariance of socioeconomic status across migrational background. *Scandinavian Journal of Educational Research*, 57(2), 148-166. DOI: 10.1080/00313831.2011.625570
- Harrington, D. (2009). *Confirmatory Factor Analysis*. Oxford University Press, Inc.
- He, J., Barrera-pedemonte, F., & Buchholz, J. (2018). Cross-cultural comparability of noncognitive constructs in TIMSS and PISA. *Assessment in Education: Principles, Policy & Practice*, 26(4), 369-385. <https://doi.org/10.1080/0969594X.2018.1469467>
- Henson, R. K. (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist*, 34, 601-629.
- Hohensinn, C. ve Kubinger, K.D. (2011). On the impact of missing values on the item fit and the model validness of the Rasch model. *Psychological Test and Assessment Modeling*, 53 (3), 380-393.
- Hong, S., Malik, M. L., & Lee, M. K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-western sample. *Educational and Psychological Measurement*, 63(4), 636-654. doi.org/10.1177/0013164403251332
- Hooper, M., Mullis, I.V.S., Martin, M. O. & Fishbein, B. (2017). TIMSS 2019 Context Questionnaire Framework. In Mullis, I. V. S., & Martin, M. O. (Eds.). *TIMSS 2019 Assessment Frameworks*. Boston College, TIMSS & PIRLS International Study Center.
- Hox, J., & Bechger, T. (1999). An introduction to structural equation modeling. *Family Science Review*, 11, 354-373.

- Hoyle, R.H. (1995). *Structural Equation Modeling: Concept, Issues and Application*. Sage Publications, Inc.
- Hu, L., & Bentler, P. M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Kankaras, M., & Moors, G. B. D. (2014). Analysis of cross-cultural comparability of PISA 2009 scores. *Journal of Cross-Cultural Psychology*, 45(3), 381-399.
- Karakoç Alatlı, B. (2016). *Uluslararası Öğrenci Değerlendirme Programı (PISA - 2012) Okuryazarlık Testlerinin Ölçme Değişmezliğinin İncelenmesi* (Yüksek Lisans Tezi). Ankara Üniversitesi, Ankara.
- Karasar, N. (2000). *Bilimsel Araştırma Yöntemi* (10.Baskı). Nobel Yayın Dağıtım.
- Kıbrıslıoğlu, N. (2015). *PISA 2012 Matematik Öğrenme Modelinin Kültürlere ve Cinsiyete Göre Ölçme Değişmezliğinin İncelenmesi: Türkiye-Çin(Şangay)-Endonezya Örneği* (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Kline, P. (1994). *An Easy Guide To Factor Analysis*. Routledge.
- Kline, R. B., (2016). *Principles and Practices of Structural Equation Modelling*. The Guilford Press.
- Kuhn, J. T., & Holling, H. (2009). Measurement invariance of divergent thinking across gender, age, and school forms. *European Journal of Psychological Assessment*, 25(1), 1-7. doi.org/10.1027/1015-5759.25.1.1
- Lay, Y. F., Ng, K. T., & Chong, P. S. (2013). Analyzing affective factors related to eighth grade learners' science and mathematics achievement in TIMSS 2007. *Asia-Pacific Education Researcher*, 24(1), 103–110. doi: 10.1007/s40299-013-0163-0
- Marsh, H. W., Abduljabbar, A.S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfettah, F., Leung, K. C., Xu, M. K., Nagengast, B., & 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), 108-128.
- Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The Counseling Psychologist*, 33, 269-298.
- Martin, M.O., von Davier, M., & Mullis, I.V.S. (Eds.) (2020). *Methods and Procedures: TIMSS 2019 Technical Report*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/methods/pdf/TIMSS-2019-MP-Technical-Report.pdf>
- MEB. (2020). *TIMSS 2019 Türkiye Ön Raporu*.
- Meredith, W. (1993). MI, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543.
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111-121.
- Mueller R.O., & Hancock G.R. (2018). *The Reviewer's Guide to Quantitative Methods In The Social Sciences*. Routledge. Structural equation modeling; 445–456.
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 Assessment Frameworks*. Boston College, TIMSS & PIRLS International Study Center.
- Mullis, I.V.S., Martin, M.O., & Hooper, M. (2017). Measuring changing educational contexts in a changing world: Evolution of the TIMSS and PIRLS questionnaires. In M. Rosén, K.Y. Hansen, & U. Wolff (Eds.), *Cognitive Abilities and Educational Outcomes* (pp. 207–222). Methodology of Educational Measurement and Assessment. Switzerland: Springer International Publishing.

- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. Boston College, TIMSS & PIRLS International Study Center. <https://timss2019.org/reports/>
- Muthén, B., & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus version 5. <http://www.statmodel.com/examples/webnote.shtml%20-%20web4>.
- Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus User's Guide*. Eighth Edition. Muthén & Muthén.
- OECD. 2023. *Education at a Glance 2023: OECD Indicators*. OECD Publishing.
- Ölçüoğlu, R., & Çetin, S. (2016). TIMSS 2011 sekizinci sınıf öğrencilerinin matematik başarısını etkileyen değişkenlerinin bölgelere göre incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 7(1), 202-220. doi: 10.21031/epod.34424
- Öncü, Ö. (2019). *TIMSS 2015 Sekizinci Sınıf Matematik Başarı Testinin OECD Ülkelerine Göre Ölçme Değişmezliğinin İncelenmesi* (Yüksek Lisans Tezi). Akdeniz Üniversitesi, Antalya.
- Polat, M. (2019). *TIMSS-2015 Matematik ve Fen Duyuşsal Özellik Modellerinin Kültürlere, Cinsiyete ve Bölgelere Göre Ölçme Değişmezliğinin İncelenmesi* (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Raju, S. N., Laffitte, J. L., & Byrne, M. B. (2002). Measurement equivalence: a comparison of methods based of confirmatory factor analysis and item response theory. *Journal of Applied Psychology*, 87 (3), 517-529.
- Ritzen J. (2013) International large-scale assessments as change agents. In: von Davier M, Gonzalez E, Kirsch I, et al. (Eds) *The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research*, Dordrecht, Netherlands: Springer.

- Schumacker, R.E., & Lomax, R.G. (1996). *A Beginners Guide to Structural Equation Modeling*. Erlbaum.
- Simşek, E., Jones, I., Hunter, J., & Xenidou-Dervou, I. (2021). Mathematical equivalence assessment: Measurement invariance across six countries. *Studies in Educational Evaluation, 70*: 51-64.
- Stark, S., Chernyshenko, O., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology, 91*(6): 1292-1306.
- Şahin Kürşad, M. & Nartgün, Z. (2015). Kayıp veri sorununun çözümünde kullanılan farklı yöntemlerin ölçeklerin geçerlik ve güvenirliği bağlamında karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 6*(2); 254-267.
- Şen, S. (2020). *Mplus ile Yapısal Eşitlik Modellemesi Uygulamaları*. Nobel Akademik Yayıncılık.
- Şencan, H. (2005). *Sosyal ve Davranışsal Ölçümlerde Güvenilirlik ve Geçerlilik*. Seçkin Yayıncılık.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Pearson.
- Tan, C. (2019). PISA and education reform in Shanghai. *Critical Studies in Education, 60*(3), 391-406.
- Tavlıca, A. (2019). *TIMSS 2015 Dördüncü Sınıf Matematik Testinin Ölçme Değişmezliğinin Ülkelere Göre İncelemesi* (Yüksek Lisans Tezi). Akdeniz Üniversitesi, Antalya.
- Tekin, Y. T., & Çobanoğlu Aktan, D. (2021). Investigation of measurement invariance of PISA 2015 collaborative problem solving skills: Turkey, Norway and Singapore, *International Journal of Assessment Tools in Education, 8*(1), 90–105.
- Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. American Psychological Association.

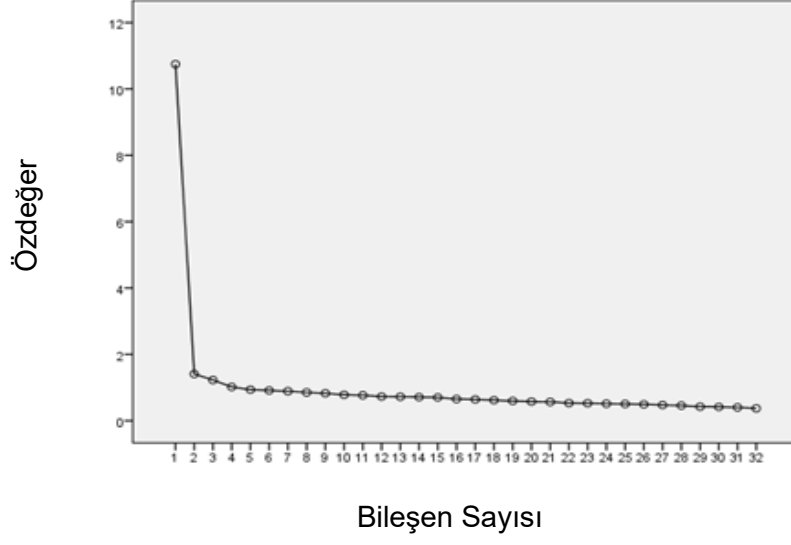
- Uyar, Ş. (2021). Factor structure and measurement invariance of the TIMSS 2015 mathematics attitude questionnaire: Exploratory structural equation modelling approach. *International Journal of Assessment Tools in Education*, 8(4), 855-871.
- Uyar, Ş., & Doğan, N. (2014). Türkiye örnekleminde öğrenme stratejileri modelinin farklı gruplarda ölçme değişmezliğinin incelenmesi [An investigation of measurement invariance of learning strategies model across different groups in PISA Turkey sample]. *International Journal of Turkish Education Sciences*, 3, 30-43.
- Ülkü, S. (2019). *ABİDE 2016 Türkçe ve Fen Bilimleri Alt-Testlerinin Öğretmen Özelliklerine Göre Ölçme Değişmezliğinin İncelenmesi* (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Ültanır, G. (2003). Eğitim ve kültür ilişkisi-eğitimde kültürün hangi boyutlarının genç kuşaklara aktarılacağı kaygısı. *Gazi Eğitim Fakültesi Dergisi*, 23(3), 291-309.
- van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492. doi: 10.1080/17405629.2012. 686740
- Vandenberg, R., & Lance, C. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 4-70.
- Varış, F. (1994). *Eğitimde Program Geliştirme Teori ve Teknikleri*. Alkım Kitapçılık Yayıncılık.
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling*, (s. 209-231). The Guilford Press.
- Weston, R., & Gore, P. A. (2006). A Brief Guide to Structural Equation Modeling. *The Counseling Psychologist*, 34(5), 719–751. <https://doi.org/10.1177/0011000006286345>

Widaman, K.F., & Reise, S.P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. *The Science of Prevention: Methodological Advances From Alcohol and Substance Abuse Research*. American Psychological Association.

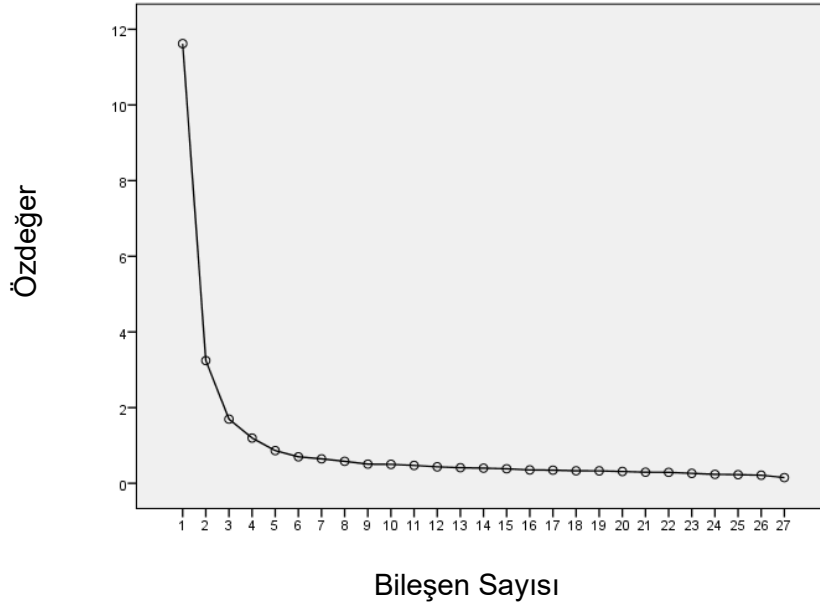
Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the Meaning of Factorial Invariance and Updating the Practice of Multi-group Confirmatory Factor Analysis: A Demonstration With TIMSS Data. *Practical Assessment, Research & Evaluation*, 12(3), 1-26.

EK-A: Yığılma Grafikleri

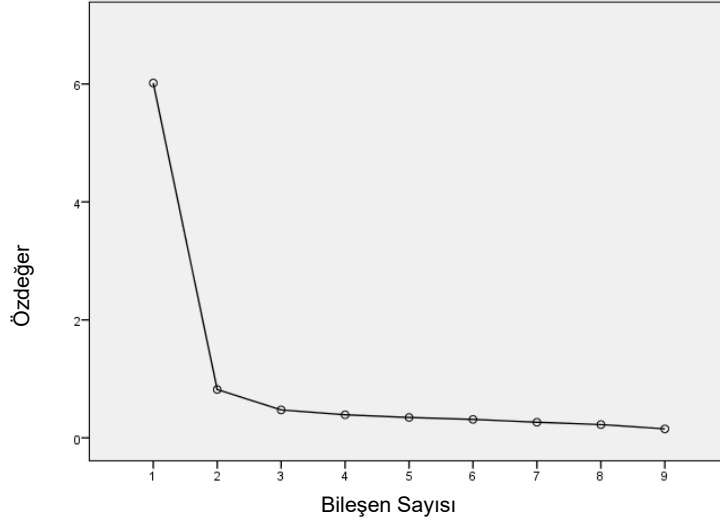
Başarı Testi Yığılma Grafiği



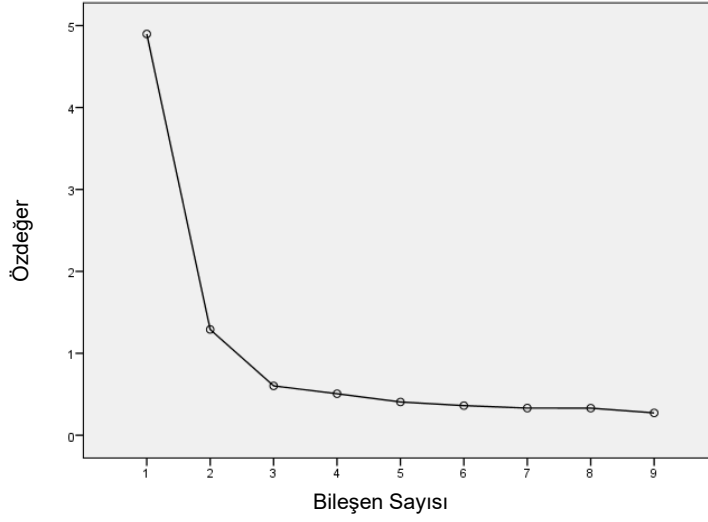
Duyuşsal Özellik Modeli Yığılma Grafiği



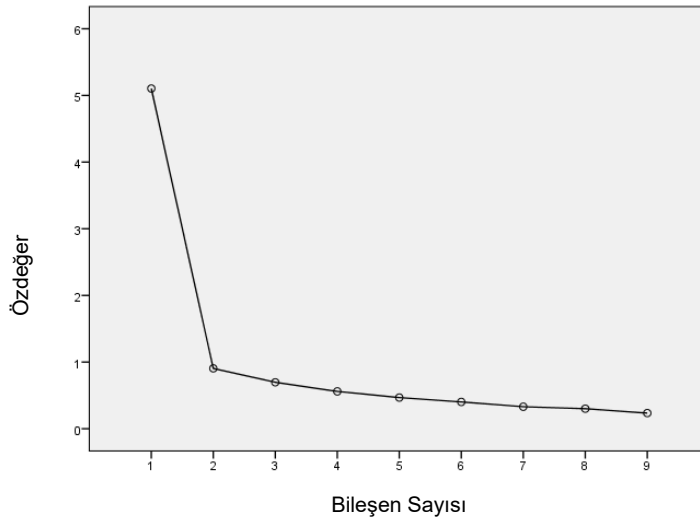
Matematik Öğrenmeyi Sevme Ölçeği Yığılma Grafiği



Matematikte Kendine Güven Ölçeği Yığılma Grafiği



Matematiğe Değer Verme Ölçeği Yığılma Grafiği



EK-B: Tetrakorik Korelasyon Tablosu

	ME1	ME2	ME3	ME4	ME5	ME6	ME7	ME8	ME9	ME10	ME11	ME12	ME13	ME14	ME15	ME16	ME17	ME18	ME19	ME20	ME21	ME22	ME23	ME24	ME25	ME26	ME27	ME28	ME29	ME30	ME31	ME32
ME1	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME2	0.446	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME3	0.554	0.722	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME4	0.340	0.437	0.461	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME5	0.463	0.531	0.591	0.489	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME6	0.392	0.417	0.544	0.468	0.599	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME7	0.444	0.490	0.524	0.460	0.574	0.489	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME8	0.482	0.525	0.583	0.477	0.580	0.481	0.423	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME9	0.599	0.554	0.731	0.608	0.774	0.755	0.705	0.823	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME10	0.281	0.299	0.490	0.306	0.439	0.391	0.377	0.319	0.574	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME11	0.219	0.243	0.314	0.214	0.265	0.185	0.212	0.209	0.399	0.215	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME12	0.404	0.386	0.472	0.289	0.400	0.278	0.284	0.391	0.521	0.292	0.188	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME13	0.412	0.521	0.537	0.419	0.537	0.495	0.411	0.463	0.665	0.304	0.216	0.376	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME14	0.427	0.573	0.635	0.425	0.475	0.458	0.412	0.399	0.638	0.255	0.334	0.346	0.468	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME15	0.512	0.481	0.643	0.412	0.469	0.549	0.475	0.481	0.709	0.311	0.328	0.382	0.451	0.438	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME16	0.421	0.474	0.529	0.360	0.434	0.416	0.468	0.438	0.644	0.409	0.342	0.260	0.386	0.450	0.634	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME17	0.362	0.348	0.487	0.431	0.600	0.676	0.414	0.429	0.718	0.377	0.123	0.316	0.537	0.395	0.447	0.332	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME18	0.482	0.595	0.632	0.489	0.688	0.556	0.607	0.507	0.717	0.429	0.257	0.392	0.557	0.507	0.565	0.543	0.562	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ME19	0.471	0.508	0.479	0.391	0.477	0.368	0.445	0.429	0.594	0.362	0.271	0.375	0.363	0.366	0.483	0.460	0.258	0.495	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-
ME20	0.538	0.562	0.624	0.540	0.638	0.611	0.552	0.613	0.708	0.388	0.245	0.363	0.582	0.511	0.547	0.509	0.648	0.665	0.493	1.000	-	-	-	-	-	-	-	-	-	-	-	-
ME21	0.485	0.444	0.551	0.439	0.540	0.499	0.527	0.509	0.778	0.363	0.248	0.379	0.483	0.405	0.522	0.428	0.512	0.542	0.426	0.585	1.000	-	-	-	-	-	-	-	-	-	-	-
ME22	0.469	0.582	0.591	0.523	0.666	0.549	0.539	0.559	0.704	0.341	0.215	0.324	0.468	0.522	0.527	0.454	0.515	0.670	0.460	0.600	0.484	1.000	-	-	-	-	-	-	-	-	-	-
ME23	0.561	0.574	0.665	0.456	0.707	0.677	0.673	0.566	0.770	0.477	0.216	0.410	0.601	0.549	0.618	0.577	0.677	0.712	0.561	0.752	0.671	0.671	1.000	-	-	-	-	-	-	-	-	-
ME24	0.478	0.454	0.568	0.383	0.577	0.546	0.501	0.529	0.703	0.467	0.253	0.462	0.467	0.479	0.489	0.489	0.477	0.619	0.549	0.546	0.547	0.489	0.720	1.000	-	-	-	-	-	-	-	
ME25	0.528	0.492	0.645	0.536	0.593	0.615	0.567	0.549	0.762	0.421	0.278	0.464	0.519	0.476	0.612	0.512	0.629	0.574	0.425	0.671	0.560	0.542	0.689	0.600	1.000	-	-	-	-	-	-	
ME26	0.375	0.472	0.425	0.366	0.409	0.307	0.419	0.432	0.538	0.286	0.241	0.284	0.331	0.403	0.325	0.329	0.295	0.424	0.366	0.357	0.331	0.458	0.420	0.412	0.458	1.000	-	-	-	-	-	
ME27	0.549	0.517	0.596	0.548	0.652	0.646	0.588	0.574	0.767	0.481	0.199	0.443	0.566	0.547	0.548	0.538	0.653	0.599	0.509	0.706	0.656	0.545	0.755	0.637	0.655	0.441	1.000	-	-	-	-	
ME28	0.658	0.547	0.690	0.568	0.689	0.692	0.616	0.638	0.786	0.461	0.298	0.490	0.661	0.548	0.643	0.545	0.726	0.685	0.527	0.778	0.668	0.570	0.769	0.676	0.727	0.498	0.756	1.000	-	-	-	
ME29	0.581	0.563	0.574	0.553	0.648	0.475	0.602	0.560	0.666	0.343	0.328	0.677	0.523	0.499	0.612	0.495	0.458	0.580	0.581	0.664	0.627	0.576	0.643	0.626	0.633	0.529	0.636	0.693	1.000	-	-	
ME30	0.408	0.581	0.471	0.359	0.542	0.416	0.417	0.515	0.545	0.261	0.365	0.322	0.499	0.389	0.445	0.355	0.332	0.518	0.445	0.489	0.444	0.521	0.490	0.429	0.388	0.363	0.462	0.448	0.417	1.000	-	
ME31	0.455	0.445	0.430	0.351	0.481	0.371	0.430	0.370	0.469	0.249	0.204	0.333	0.376	0.422	0.502	0.380	0.334	0.438	0.448	0.510	0.405	0.499	0.540	0.467	0.479	0.432	0.540	0.529	0.558	0.398	1.000	-
ME32	0.433	0.613	0.595	0.470	0.523	0.499	0.569	0.557	0.593	0.392	0.247	0.374	0.413	0.524	0.468	0.452	0.466	0.517	0.414	0.536	0.513	0.471	0.562	0.536	0.606	0.383	0.618	0.614	0.517	0.445	0.448	1.000

ME1-ME5, ME17-ME20: Sayılar; ME6-ME10, ME21-ME25: Cebir;
 ME11-ME13, ME26-ME29: Geometri; ME14-ME16, ME30-ME32: Veri ve Olasılık

EK-C: Polikorik Korelasyon Tablosu

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23	A24	A25	A26	A27
A1	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A2	0,533	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A3	0,693	0,679	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A4	0,738	0,438	0,558	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A5	0,889	0,576	0,735	0,754	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A6	0,761	0,456	0,593	0,691	0,804	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A7	0,802	0,521	0,654	0,704	0,845	0,794	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A8	0,763	0,527	0,649	0,662	0,808	0,753	0,76	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A9	0,825	0,538	0,719	0,685	0,887	0,757	0,814	0,828	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A10	0,645	0,47	0,518	0,523	0,681	0,556	0,651	0,554	0,666	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A11	0,341	0,346	0,372	0,244	0,365	0,255	0,381	0,265	0,397	0,565	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A12	0,527	0,468	0,537	0,402	0,582	0,449	0,561	0,463	0,614	0,683	0,696	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A13	0,625	0,427	0,502	0,54	0,658	0,558	0,644	0,558	0,648	0,774	0,495	0,608	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A14	0,256	0,387	0,34	0,173	0,293	0,209	0,269	0,265	0,301	0,357	0,459	0,508	0,315	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-
A15	0,579	0,408	0,465	0,487	0,618	0,537	0,664	0,532	0,635	0,766	0,503	0,633	0,747	0,324	1,000	-	-	-	-	-	-	-	-	-	-	-	-
A16	0,563	0,372	0,438	0,48	0,586	0,511	0,57	0,527	0,59	0,738	0,452	0,561	0,673	0,257	0,707	1,000	-	-	-	-	-	-	-	-	-	-	-
A17	0,369	0,494	0,433	0,264	0,424	0,269	0,405	0,33	0,448	0,575	0,619	0,638	0,491	0,577	0,502	0,429	1,000	-	-	-	-	-	-	-	-	-	-
A18	0,383	0,476	0,457	0,268	0,416	0,312	0,397	0,365	0,435	0,502	0,541	0,589	0,454	0,61	0,455	0,391	0,747	1,000	-	-	-	-	-	-	-	-	-
A19	0,539	0,428	0,421	0,525	0,537	0,498	0,499	0,497	0,471	0,422	0,15	0,269	0,414	0,122	0,366	0,4	0,212	0,215	1,000	-	-	-	-	-	-	-	-
A20	0,453	0,348	0,361	0,475	0,457	0,443	0,43	0,43	0,422	0,361	0,119	0,21	0,363	0,07	0,347	0,362	0,171	0,16	0,66	1,000	-	-	-	-	-	-	-
A21	0,469	0,274	0,362	0,44	0,471	0,426	0,447	0,388	0,452	0,412	0,199	0,296	0,406	0,046	0,378	0,374	0,134	0,13	0,486	0,518	1,000	-	-	-	-	-	-
A22	0,473	0,285	0,354	0,448	0,481	0,448	0,447	0,399	0,457	0,409	0,176	0,289	0,389	0,053	0,381	0,372	0,142	0,15	0,505	0,523	0,838	1,000	-	-	-	-	-
A23	0,609	0,447	0,489	0,541	0,635	0,583	0,597	0,575	0,624	0,539	0,253	0,447	0,511	0,202	0,542	0,494	0,315	0,323	0,562	0,497	0,577	0,65	1,000	-	-	-	-
A24	0,517	0,346	0,389	0,509	0,52	0,489	0,479	0,458	0,464	0,412	0,133	0,259	0,406	0,064	0,359	0,376	0,155	0,159	0,633	0,589	0,651	0,684	0,624	1,000	-	-	-
A25	0,479	0,327	0,357	0,472	0,478	0,431	0,455	0,406	0,43	0,423	0,186	0,267	0,405	0,054	0,358	0,367	0,188	0,154	0,578	0,558	0,691	0,703	0,574	0,762	1,000	-	-
A26	0,417	0,18	0,259	0,407	0,413	0,39	0,397	0,343	0,367	0,376	0,134	0,232	0,355	-0,029	0,325	0,338	0,064	0,058	0,424	0,396	0,579	0,565	0,44	0,597	0,625	1,000	-
A27	0,534	0,294	0,399	0,518	0,53	0,501	0,5	0,448	0,491	0,432	0,169	0,296	0,417	0,023	0,371	0,385	0,114	0,116	0,543	0,529	0,678	0,68	0,548	0,748	0,742	0,73	1,000

A1-A9: Matematik Öğrenmeyi Sevme;
A10-A18: Matematikte Kendine Güven;
A19-A27: Matematiğe Değer Verme

EK-Ç: Araştırma Etik Komisyon İzin Muafiyeti Formu

	Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması/Araştırma Etik Komisyon İzin Muafiyeti Formu	F46
Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Ana Bilim Dalı Başkanlığına		15 / 01 / 2024
Tez/Araştırma Başlığı	TIMSS 2019 BAŞARI TESTİ VE DUYUŞSAL ÖZELLİK MODELİNİN ÖLÇME DEĞİŞMEZLİĞİNİN İNCELENMESİ	
Yukarıda başlığı/konusu verilen tez/araştırma çalışmam,		
1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır. 2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmaması gerektirmektedir. 3. Beden bütünlüğüne veya ruh sağlığına müdahale içermemektedir. 4. Anket, ölçek (test), mülakat, odak grup çalışması, gözlem, deney, görüşme gibi teknikler kullanılarak katılımcılardan veri toplamasını gerektiren nitel ya da nicel yaklaşımlarla yürütülen araştırmalar niteliğinde değildir. 5. Diğer kişi ve kurumlardan temin edilen veri kullanılması (kütap, belge vs.) gerektirmektedir. Ancak bu kullanım, diğer kişi ve kurumların izin verdiği ölçüde Kişisel Bilgilerin Korunması Kanununa riayet edilerek gerçekleştirilecektir.		
Çalışmada kullanılacak veriler: <input checked="" type="checkbox"/> Kamusal erişime açık (buraya yazınız): https://timss2019.org/international-database/ <input type="checkbox"/> Özel izin ve onaya tabii (buraya yazınız): <input type="checkbox"/> Üretilmiş veri (buraya yazınız): <input type="checkbox"/> Diğer (buraya yazınız):		
Yükseköğretim Kurumları Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre çalışmamın yürütülebilmesi için herhangi bir Etik Kurulundan/Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.		
Gereğini saygılarımla arz ederim.		
Ferda KOÇAK <small>(Araştırma Akademi Uzmanı)</small>		
Araştırmacı Bilgileri		
Adı Soyadı	Ferda KOÇAK	
Öğrenci İse No	N21223424	
Ana Bilim Dalı	Eğitim Bilimleri Ana Bilim Dalı	
Programı	Eğitimde Ölçme ve Değerlendirme	
Çalışma Türü	<input checked="" type="checkbox"/> Tez <input type="checkbox"/> Tezden Üretilen Yayın <input type="checkbox"/> Araştırma Makalesi	
Statüsü	<input checked="" type="checkbox"/> Yüksek Lisans <input type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr. <input type="checkbox"/> Diğer	
Danışman Görüşü ve Onayı*		
Öğrencimizin hazırladığı tez çalışması etik kurulların görüşleri doğrultusunda uygun görülmüştür.		
Prof. Dr. Nuri DOĞAN <small>(Doçent)</small> <small>(Danışman Üretim, Akademi Uzmanı)</small>		
*Tez ve tezden üretilen yayın ve araştırma makalelerinde geçerli		
Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Beytepe Yerleşkesi, 06000, Çankaya / ANKARA Telefon: 0312 297 85 72 Beşçeşme: 0312 297 85 66 e-Adı: etk@hacettepe.edu.tr e-Posta: etk@hacettepe.edu.tr		

EK-D: Etik Beyanı

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- * tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- * görsel, işitsel ve yazılı bütün bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- * başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- * atıfta bulunduğum eserlerin bütününe kaynak olarak gösterdiğimi,
- * kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- * bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

...../...../.....

Ferda KOÇAK

EK-E: Yüksek Lisans Tez Çalışması Orijinallik Raporu

HACETTEPE ÜNİVERSİTESİ
Eğitim Bilimleri Enstitüsü
Eğitim Bilimleri Ana Bilim Dalı Başkanlığına,

...../...../.....

Tez Başlığı : TIMSS 2019 MATEMATİK BAŞARI TESTİ VE MATEMATİĞE İLİŞKİN DUYUŞSAL ÖZELLİKLERİN ÖLÇME DEĞİŞMEZLİĞİNİN İNCELENMESİ

Yukarıda başlığı verilen tez çalışmamın tamamı (kapak sayfası, özetler, ana bölümler, kaynakça) aşağıdaki filtreler kullanılarak **Turnitin** adlı intihal programı aracılığı ile kontrol edilmiştir. Kontrol sonucunda aşağıdaki veriler elde edilmiştir:

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
25/03/2024	84	133,786	13/02/2024	%12	2330683091

Uygulanan filtreler:

1. Kaynaklar hariç
2. Alıntılar dâhil
3. 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esaslarını inceledim ve çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan eder, gereğini saygılarımla arz ederim.

Ad Soyadı: Ferda KOÇAK

Öğrenci No.: N21223424

Ana Bilim Dalı: Eğitim Bilimleri Ana Bilim Dalı

İmza

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü: Y.Lisans Doktora Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Prof. Dr. Nuri DOĞAN

EK-F: Thesis/Dissertation Originality Report

HACETTEPE UNIVERSITY
Graduate School of Educational Sciences
To The Department of

Thesis Title: INVESTIGATION OF MEASUREMENT INVARIANCE OF TIMSS 2019 MATHEMATICS ACHIEVEMENT TEST AND MATHEMATICAL ATTITUDINAL CHARACTERISTICS

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
25/03/2024	84	133,786	13/02/2024	%12	2330683091

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Name Lastname: **Ferda KOÇAK**

Student No.: N21223424

Department: Eğitim Bilimleri Ana Bilim Dalı

Program: Eğitimde Ölçme ve Değerlendirme

Status: Masters Ph.D. Integrated Ph.D.

Signature

ADVISOR APPROVAL

APPROVED

Prof. Dr. Nuri DOĞAN

EK-G: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren ... ay ertelenmiştir. ⁽²⁾
- Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

..... / /

(imza)

Ferda KOÇAK

"Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlerle ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir
*Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

