

**ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME PROGRAMI**

**PISA 2015 ÖĞRENCİ TUTUM ANKETLERİNİN DEĞİŞEN
MADDE FONKSİYONU VE ÖLÇME DEĞİŞMEZLİĞİNİN
İNCELENMESİ**

YÜKSEK LİSANS TEZİ

FURKAN TİRYAKİ

**ANKARA
HAZİRAN, 2019**



**ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME PROGRAMI**

**PISA 2015 ÖĞRENCİ TUTUM ANKETLERİNİN DEĞİŞEN
MADDE FONKSİYONU VE ÖLÇME DEĞİŞMEZLİĞİNİN
İNCELENMESİ**

YÜKSEK LİSANS TEZİ

FURKAN TİRYAKİ

DANIŞMAN: DOÇ. DR. CELAL DEHA DOĞAN

**ANKARA
HAZİRAN, 2019**

Eğitim Bilimleri Enstitüsü Müdürlüğü'ne

Furkan TİRYAKİ adlı öğrencinin hazırladığı “PISA 2015 Öğrenci Tutum Anketlerinin Değişen Madde Fonksiyonu ve Ölçme Değişmezliğinin İncelenmesi” başlıklı bu çalışma Eğitim Bilimleri Anabilim Dalı / Eğitimde Ölçme ve Değerlendirme Programı’nda jüri üyelerince oy birliği / oy çokluğu ile Yüksek Lisans Tezi olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan: Doç. Dr. Burcu ATAR



Üye: Doç. Dr. Hamide Deniz GÜLLEROĞLU



Üye: Doç. Dr. Celal Deha DOĞAN (Danışman)



ONAY

Bu tez Ankara Üniversitesi Lisansüstü Eğitim - Öğretim ve Sınav Yönetmeliği'nin ilgili maddeleri uyarınca yukarıdaki jüri üyeleri tarafından 19.06.2019 tarihinde, Enstitü Yönetim Kurulunca/...../20..... tarihinde kabul edilmiştir.

Prof. Dr. Yasemin KEPENEKÇİ

Enstitü Müdürü

ETİK İLKELERE UYGUNLUK BİLDİRİMİ

Tez içindeki bütün bilgileri akademik yazım kurallarına uygun biçimde raporlaştırdığımı ve bunları etik ilkelere (atıfta bulunulan tüm yapıtlara kaynaklarda yer verilmesi, tezde kullanılan bilgi ve belgelere resmi yollarla ulaşılması ve bunların aslı bozulmadan kullanılması vb.) uygun olarak elde ettiğimi ve sunduğumu bildiririm.



Furkan TİRYAKİ

ÖZET

PISA 2015 ÖĞRENCİ TUTUM ANKETLERİNİN DEĞİŞEN MADDE FONKSİYONU VE ÖLÇME DEĞİŞMEZLİĞİNİN İNCELENMESİ

TİRYAKİ, Furkan

Yüksek Lisans Tezi, Ölçme ve Değerlendirme Anabilim Dalı

Tez Danışmanı: Doç. Dr. Celal Deha DOĞAN

Haziran, 2019, xi + 64 Sayfa

Bu çalışmanın amacı OECD tarafından her üç yılda bir gerçekleştirilen Uluslararası Öğrenci Değerlendirme Programı (PISA) 2015 uygulamasının fen bilimlerine yönelik öğrenci tutum anketlerinin 37 madde içeren 6 alt testi için çoklu grup doğrulayıcı faktör analizi yöntemiyle ölçme değişmezliğinin, madde tepki kuramı olabilirlik oranı yöntemiyle ise değişen madde fonksiyonu açısından incelenmesidir. Araştırmanın örneklemini Türkiye ve ABD’de PISA 2015 uygulamasına katılan toplam 11.607 öğrenci oluşturmaktadır.

Ölçme değişmezliği analizinin sonucunda ST094 (Fen öğrenmekten zevk alma) ve ST113 (Fen öğreniminde araçsal güdülenme) alt testleri bütün değişmezlik aşamalarını karşılarken, ST095 (Geniş fen konularına ilgi), ST129 (Fene yönelik öz yeterlik), ST131 (Fen bilimlerine yönelik epistemik inançlar) ve ST146 (Bilim aktiviteleri) alt testlerinde yalnızca metrik değişmezlik boyutları sağlanmıştır.

Değişen madde fonksiyonu analizi sonucunda cinsiyet açısından Türkiye örnekleminde 30 maddede, ABD örnekleminde 20 maddede ihmal edilemeyecek boyutta değişen madde fonksiyonu gözlenmiştir. ABD ve Türkiye arasında kültürel açıdan yapılan analizde 34 maddede ihmal edilemeyecek boyutta değişen madde fonksiyonu gözlenmiştir.

Ölçme değişmezliği ve değişen madde fonksiyonu analizlerinin çıktıları incelendiğinde elde edilen sonuçlar birbiriyle tutarlı bulunmuştur.

Anahtar Kelimeler: PISA, değişen madde fonksiyonu, ölçme değişmezliği, madde tepki kuramı olabilirlik oran testi, çoklu grup doğrulayıcı faktör analizi

ABSTRACT

ANALYSIS ON DIFFERENTIAL ITEM FUNCTIONING AND MEASUREMENT INVARIANCE OF PISA 2015 STUDENT ATTITUDE SURVEYS

TİRYAKİ, Furkan

Master's Thesis, Department of Measurement and Evaluation

Supervisor: Assoc. Prof. Dr. Celal Deha DOĞAN

June, 2019, xi + 64 Pages

The aim of this study is to examine the Programme for International Student Assessment's (PISA) measurement invariance through the multi-group confirmatory factor analysis method and the differential item functioning through the item response theory likelihood-ratio test method. A total of 6 sub-tests consisting of 37 items of the student attitude surveys towards science in the 2015 implementation of PISA, which is carried out every three years by the OECD, have been examined.

The sample of the study consists of 11,607 students from Turkey and the U.S.A. who participated in the PISA 2015.

As a result of the measurement invariance analysis, the sub-tests of ST094 (enjoyment of science) and ST113 (instrumental motivation to learn science) met all criteria of the measurement invariance steps, while the sub-tests of ST095 (interest in science), ST129 (science self-efficacy), ST131 (epistemic beliefs about science) and ST146 (science activities) only ensured the metric invariance level.

As a result of differential item functioning analysis in terms of gender, a significant amount of items indicating DIF were observed in 30 items in the Turkey sample, and 20 items in the U.S.A. sample. A significant amount of items indicating DIF was observed in 34 items in the analysis made between the U.S.A. and Turkey in terms of culture.

When the outcome of the measurement invariance and differential item functioning analyzes were examined, the results were found to be consistent with each other.

Keywords: PISA, differential item functioning, measurement invariance, item response theory likelihood-ratio test, multi-group confirmatory factor analysis

ÖNSÖZ

Tez çalışmam ve ders dönemimde desteğini üstümden esirgemeyen sabırla ve özveriyle benim yanımda duran, değerli tez danışmanım Doç. Dr. Celal Deha DOĞAN'a,

Süreç boyunca tez çalışmamın şekillenmesinde desteklerini sunan değerli hocalarım Fatih İMROL ve Ömer KAMIŞ'a,

Yüksek lisans süresince kendilerinden ders alma fırsatı bulduğum, akademik katkılarının yanında hayat dersleri de edindiğim Prof. Dr. Nizamettin KOÇ, Dr. Öğr. Üyesi Ömer KUTLU, Prof. Dr. Nükhet ÇIKRIKÇI DEMİRTAŞLI ve Prof. Dr. Ömay ÇOKLUK BÖKEOĞLU'na,

Hayatım boyunca desteklerini benden esirgemeyen, bu günlere gelmemde sonsuz katkıları bulunan Babam Hamza TİRYAKİ'ye, Annem Hülya TİRYAKİ'ye, Kardeşim Ayşe İrem TİRYAKİ'ye,

Ve her zaman kalbimde olan değerli eşim Kubra TİRYAKİ'ye sonsuz teşekkürlerimi sunuyorum.

Sizinle hayat yolunda çizgilerimizin kesişmesi benim için onur verici olmuştur.

İÇİNDEKİLER

	Sayfa
ETİK İLKELERE UYGUNLUK BİLDİRİMİ.....	iii
ÖZET	iv
ABSTRACT	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
TABLolar DİZİNİ.....	ix
ŞEKİLLER DİZİNİ	x
KISALTMALAR/SİMGELER.....	xi
BÖLÜM 1	1
GİRİŞ.....	1
1.1. Problem Durumu.....	1
1.2. Amaç	4
1.3. Önem.....	5
1.4. Sayıtlılar.....	6
1.5. Sınırlılıklar	6
BÖLÜM 2	7
KAVRAMSAL ÇERÇEVE.....	7
2.1. Ölçme Değişmezliği.....	7
2.2. Çoklu Grup Doğrulayıcı Faktör Analizi	7
2.3. Yapısal Değişmezlik	8
2.4. Metrik Değişmezlik.....	9
2.5. Skalar Değişmezlik	9
2.6. Katı Değişmezlik	10
2.7. Uyum İyiliği İndeksleri.....	10
2.8. Madde Yanlılığı ve Değişen Madde Fonksiyonu	12
2.9. Madde Tepki Kuramı Olabilirlik Oranı (MTK-OO).....	13
2.10. İlgili Araştırmalar.....	14
2.11. Ölçme Değişmezliği Çalışmaları	14
2.12. Değişen Madde Fonksiyonu Çalışmaları	16

BÖLÜM 3	19
YÖNTEM	19
3.1. Araştırma Modeli	19
3.2. Evren-Örneklem.....	19
3.3. Veriler ve Toplanması.....	20
3.4. Verilerin Analizi	20
3.5. Ölçme Değişmezliği Varsayımlarının İncelenmesi	21
BÖLÜM 4.....	26
BULGULAR VE YORUMLAR	26
4.1. Birinci Alt Probleme Yönelik Bulgu ve Yorumlar	26
4.2. İkinci Alt Probleme Yönelik Bulgu ve Yorumlar.....	39
4.3. Problem 2a'ya Yönelik Bulgu ve Yorumlar	39
4.4. Problem 2b'ye Yönelik Bulgu ve Yorumlar.....	41
4.5. Problem 2c'ye Yönelik Bulgu ve Yorumlar	42
BÖLÜM 5	44
SONUÇLAR VE ÖNERİLER.....	44
5.1. Sonuçlar	44
5.1.1. Değişen Madde Fonksiyonuna Yönelik Sonuçlar.....	44
5.1.2. Ölçme Değişmezliğine Yönelik Sonuçlar.....	45
5.2. Öneriler	45
KAYNAKLAR	47
EKLER	51
Ek 1. PISA 2015 fen bilimlerine yönelik öğrenci tutum anketlerinin Türkçe Formları	52
Ek 2. PISA 2015 fen bilimlerine yönelik öğrenci tutum anketlerinin İngilizce Formları.....	55
Ek 3. Doğrusallık varsayımı için Saçılım Diyagramı Testleri	58
Ek 4. MTK-OO Örnek Çıktısı.....	61
Ek 5. Etik Kurul Kararı	62
BENZERLİK BİLDİRİMİ	63
ÖZGEÇMİŞ	64

TABLOLAR DİZİNİ

Tablo		Sayfa
Tablo 1	Ülkeler ve Cinsiyetler Açısından Örneklem.....	20
Tablo 2	Kayıp Verilerinin Düzenlenmesinden Sonra Veri Seti.....	22
Tablo 3	Kültüre Göre ABD ve Türkiye Örneklemeleri Çoklu Normallik Testi Çarpıklık Sonuçları.....	22
Tablo 4	Kültüre Göre ABD ve Türkiye Örneklemeleri Çoklu Normallik Testi Basıklık Sonuçları.....	22
Tablo 5	Kültüre Göre Türkiye ve ABD örnekleme Basıklık ve Çarpıklık Değerleri	23
Tablo 6	Kültüre Göre Türkiye ve ABD örnekleme Çoklu Bağlantılılık ve Eş Varyanslılık İstatistikleri (Tolerans Değerleri ve Varyans Artış Faktör Değerleri).....	24
Tablo 7	St094 Alt Testi Uyum İndeksleri.....	26
Tablo 8	St094 Ükelere Göre Model Uyum İndeksleri.....	27
Tablo 9	St094 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları.....	28
Tablo 10	St095 Alt Testi Uyum İndeksleri.....	29
Tablo 11	St095 Ükelere Göre Model Uyum İndeksleri.....	30
Tablo 12	St095 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları.....	30
Tablo 13	St113 Alt Testi Uyum İndeksleri.....	31
Tablo 14	St113 Ükelere Göre Model Uyum İndeksleri.....	32
Tablo 15	St113 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları.....	32
Tablo 16	St129 Alt Testi Uyum İndeksleri.....	33
Tablo 17	St129 Ükelere Göre Model Uyum İndeksleri.....	34
Tablo 18	St129 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları.....	34
Tablo 19	St131 Alt Testi Uyum İndeksleri.....	35
Tablo 20	St131 Ükelere Göre Model Uyum İndeksleri.....	36
Tablo 21	St131 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları.....	36
Tablo 22	St146 Alt Testi Uyum İndeksleri.....	37
Tablo 23	St146 Ükelere Göre Model Uyum İndeksleri.....	38
Tablo 24	St146 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları.....	38
Tablo 25	Cinsiyete göre Türkiye örnekleme MTK-OO analiz sonuçları	40
Tablo 26	Cinsiyete göre ABD örnekleme MTK-OO analiz sonuçları	41
Tablo 27	Kültüre Göre Türkiye ve ABD örnekleme MTK-OO analiz sonuçları.....	42

ŞEKİLLER DİZİNİ

Şekil	Sayfa
Şekil 1. ST094 Türkiye ve ABD örnekleme DFA Standartlaştırılmış Parametre Değerleri.....	27
Şekil 2. ST095 Türkiye ve ABD Örnekleme DFA Standartlaştırılmış Parametre Değerleri.....	29
Şekil 3. St113 Türkiye ve ABD örnekleme DFA Standartlaştırılmış Parametre Değerleri.....	31
Şekil 4. St129 Türkiye ve ABD örnekleme DFA Standartlaştırılmış Parametre Değerleri.....	33
Şekil 5. St131 Türkiye ve ABD Örnekleme DFA Standartlaştırılmış Parametre Değerleri.....	35
Şekil 6. ST146 Türkiye ve ABD örnekleme DFA Standartlaştırılmış Parametre Değerleri.....	37

KISALTMALAR/SİMGELER

ÇGDFA	: Çoklu Grup Doğrulayıcı Faktör Analizi
DMF	: Değişen Madde Fonksiyonu
GADM	: Genelleştirilmiş Aşamalı Doğrusal Modelleme
KTK	: Klasik Test Kuramı
MTK	: Madde Tepki Kuramı
MTK-OO	: Madde Tepki Kuramı Olabilirlik Oranı
OECD	: İktisadi İşbirliği ve Gelişme Teşkilatı
PIRLS	: Uluslararası Okuma Becerileri Projesi
PISA	: Uluslararası Öğrenci Değerlendirme Programı
TIMSS	: Uluslararası Matematik ve Fen Eğilimleri Araştırması

BÖLÜM 1

GİRİŞ

Bu bölümde problem durumu, araştırmanın amacı, problem cümlesi, sayıtlar ve sınırlılıklar yer almaktadır.

1.1. Problem Durumu

Ulusal politikalar açısından bakıldığında ülkeler, çağa uyum sağlamış eğitimli bireyler yetiştirme amacı güder. Eğitimli bireyler uluslararası toplumların her açıdan daha ileri gitmesini sağlamaktadır. Bu amaca ulaşmanın yolu ise sağlam ve etkili bir eğitim politikası uygulamaktır. Ülkeler kritik ve önemli kararlar verme aşamasında çeşitli ölçütlere ihtiyaç duyarlar. Eğitim politikalarının ve hedeflerinin belirlenmesi aşamasında ise bu kıstas genellikle uluslararası kıyaslamalar ve tanılayıcı kriterlerin belirlenmesinden geçer. Bu kıyaslamalar ve kriterlerin sağlanmasında en etkili araçlardan biri de uluslararası ölçeklerde uygulanan geniş katımlı sınavlardır. Geniş katımlı sınavlar ülkelerin eğitim sistemlerini ilerletmek ve iyileştirmekte kullanabilecekleri önemli veriler sunmaktadır.

Günümüz dünyasında ülkeler eğitime her geçen gün daha fazla önem vermektedirler. PISA, TIMSS ve PIRLS gibi geniş ölçekli sınavlardan elde edilen çıktılar, ülkelerin eğitim programlarını diğer ülkelere göre ne düzeyde olduklarını kıyaslaması, eğitim politikalarında yaptıkları veya yapacakları değişim, gelişim ve düzeltmelerin doğruluğu ve uygunluğu hakkında uluslara önemli bilgiler vermektedir (Berberoğlu ve Kalender, 2005).

Türkiye’de çeşitli amaçlarla kullanılmak üzere birçok geniş katımlı sınav uygulanmaktadır. Bu sınav uygulayıcılarının başında gelen MEB ülke çapındaki sınavlara ek olarak uluslararası toplumlar üzerinde uygulanan sınavlara da Türk öğrencilerin katılımını sağlamaktadır. Ülkemizin de katıldığı uluslararası sınavlardan TIMSS, Hollanda merkezli Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu tarafından ilk olarak 1995’te uygulanmıştır. Fen ve matematik başarısını ölçmek için 4. ve 8. sınıflar üzerinde uygulanıp her 4 senede bir tekrarlanmaktadır. PIRLS ise yine aynı kurum tarafından ilki

2001 yılında yapılan ve 5 yılda bir tekrarlanan 4-5 ve 6. Sınıf öğrencileri için okuduğunu anlama becerilerinin ölçülmesi amacıyla uygulanan bir sınavdır (MEB, 2016).

Türkiye'nin de katılımcısı olduğu bu sınavlardan biri de PISA'dır. PISA sınavı İktisadi İşbirliği ve Gelişme Teşkilatı tarafından finanse edilip uygulanmakta ve 2000 yılından itibaren her üç yılda bir tekrarlanmaktadır. PISA sınavının amacı 15 yaş grubu öğrencilerinin modern topluma uyumlu ve başarılı bir şekilde katılmaları için gereken temel bilgi ve becerilere ne kadar sahip olduklarının incelenmesi ve ölçülmesidir (MEB, 2016).

PISA sınavını diğer geniş katımlı uluslararası sınavlardan ayıran iki önemli farklılık vardır. Birinci farklılık içerik açısından ele alınmıştır. Diğer çalışmalarda genellikle öğretim programları ve müfredat ele alınırken PISA okuryazarlık adını verdiği yapıyı ölçmeyi amaçlar (MEB, 2013). Okuryazarlık kavramı ise, öğrencilerin eğitim ve günlük hayatın içerisinde karşılarına çıkabilecek problemleri, sorunları, güçlükleri çözme mekanizmaları olarak öğrendikleri becerileri kullanma, analiz ve çıkarımlar yapma, iletişim araçlarını etkili kullanma becerileri olarak belirlenmiştir (MEB, 2016). İkinci farklılık ise toplanan veri açısından ele alınmıştır. Diğer sınavlardan farklı olarak okulun yapısal işleyişi, okul iklimi, eğitimsel eşitlik kavramı, eğitim fırsatları, ülke olanakları, demografik bilgiler ve farklılıklar açısından eğitim programları ve politikaları belirlemede kullanılabilecek bilgiler üretmesi açısından veriler toplamaktır (MEB, 2013).

PISA sınavının odaklandığı temel alanları fen, matematik ve okuma becerileri oluşturmakta ve bu beceriler okuryazarlık kavramı üzerinden tanımlanmaktadır. PISA uygulamalarının her döngüsünde incelenen temel becerilerden birisi odak noktası olarak ele alınmaktadır. PISA 2015 döngüsünün odak noktası fen okuryazarlığı olarak belirlenmiştir. (MEB, 2016).

Fen okuryazarlığı PISA açısından ele alındığında “bilim ile ilgili konulara, bilimin fikirler ve buluşlarına yakınlık gösterip değer veren ve onları günlük yaşamına uygulayan vatandaş” olarak tanımlanmaktadır. Bilimsel olarak okuryazar olan bir kişi bilim ve teknoloji konularında mantıklı bir söylem geliştirmeye yatkın biridir. Kısaca fen okuryazarlığı ya da bilimsel okuryazarlık, kişinin olayları bilimsel olarak açıklama, bilimle ilgili durumları değerlendirme ve tasarlama, verileri ve kanıtları bilimsel olarak yorumlama için gerekli yeterliliğe sahip olma durumudur (OECD, 2016).

PISA'ya katılan ülkeler elde edilen verileri eğitim politikalarını geliştirmek, kendi ülkelerindeki öğrencilerin bilgi ve beceri düzeylerini başka ülkeler ile karşılaştırmak, eğitim sistemlerinin içinde bulunduğu durumu tespit etmek ve nihai olarak eğitim

çıktılarının kalitesini arttırmak amacı ile kullanılmaktadırlar. PISA Türkiye'nin de aralarında bulunduğu en geniş katılımlı eğitim uygulamasıdır. Türkiye'nin 2003'ten beri katıldığı PISA uygulamasına 2000 yılında 43, 2003'te 41, 2006'da 57, 2009'da 75, 2012'de 65 ve 2015'te 73 ülke katılmıştır. Bu katılımın sonraki PISA döngülerinde daha da artacağı öngörülmektedir (OECD, 2016).

PISA sınavı salt biçimde ülkeler arası karşılaştırmaların yanında ülke içinde de tanılayıcı veriler sunarak karşılaştırmalara da olanak sağlar. Bu karşılaştırmaların alt boyutları arasında okul türleri, cinsiyet ve istatistiki bölge birimleri vardır. Bu açıdan ulusal farklılıkların ortaya konulması, her bir alt boyutta yer alan bireylerin eşit eğitim imkânlarına ulaşmasının öncülüdür. Bu bilginin ışığında ülkeler PISA verileri ile mikro müdahaleler yapabilmekte ve alt boyutlarda aksayan durumları tespit edip sorunu giderebilmektedirler.

PISA uygulaması, temel becerileri ölçmenin yanında anketler aracılığıyla öğrencilerin öz yeterlilikleri, motivasyonları ve öğrenme sürecindeki psikolojik durumları, kariyer planları, okul ve aile ortamları ile ilgili veriler de toplar (MEB, 2016). İlgili anketler, eğitim politikalarının belirlenmesi ve ortaya çıkan sonuçların daha iyi okunması açısından yarar sağlar.

Çıktıları açısından kritik öneme sahip sınavların geçerliği ve güvenilirliğinin incelenmesi, sınav sonuçlarının etki edeceği kararların isabetliliği açısından önem arz etmektedir. Duruma bu açıdan bakıldığında PISA sınavı, hem ülke çapında bilinirliğinin yüksek olması hem de ülkemizin OECD üyesi olması sebebiyle bu araştırmanın konusu olmuştur. Tüm bu hususlar dikkate alınarak PISA 2015 uygulaması fen bilimlerine yönelik öğrenci tutum anketlerinin değişen madde fonksiyonu ve ölçme değişmezliği açısından incelenmesinin sınavın geçerlik ve güvenilirliği açısından önemli bulunmuştur.

Ulusal ve uluslararası çapta yapılan araştırmalar ve alan yazın incelendiğinde PISA sınavının öğrenci anketleri için yapılan psikometrik çalışmaların sayısının görece yetersiz olduğu ve bu konuda bir eksiklik olduğu tespit edilmiştir. PISA, öğrencilerin akademik başarılarını maksimum performans testleriyle ölçerken bir taraftan da akademik başarılarına ilişkin duyuşsal özelliklerin ölçümlerini de öğrenci anketleri yoluyla yapmaktadır. Duyuşsal özelliklerin öğrenci başarısı üzerinde önemli bir etkisinin olduğu bilinmektedir.

PISA açısından bakıldığında değişen madde fonksiyonu analizinin alan yazında çoğunlukla iki kategorili maddeler ve başarı testleri üzerinde yoğunlaştığı, bununla birlikte çok kategorili maddeler için değişen madde fonksiyonunun incelenmesi

alışmalarının lkemizde sayıca ok az olduėu ve olanların da analiz yntemlerinin genellikle KTK zerinden olduėu grlmř ve bu durum gz nne alınarak MTK zerinden yapılacak analizlerin alana katkı saėlayacaėı ngrlmřtr.(Atalay, 2006; Ayan, 2011; Demir, 2013; Akın, 2015; Gk, Kabasakal ve Kelecioėlu, 2014; Yıldırım, 2006).

PISA sınavının amalarından biri de lkelerin kendi eėitim durumlarını farklı lkelerle karřılařtırma imkânı sunmasıdır. Bu karřılařtırmanın saėlıklı olması aısından birok dile evrilen ve farklı kltrlerde uygulanan sınavın lme deėiřmezliėi tartıřmaya aık bir konudur. PISA sınavı iin lme deėiřmezliėi analizi alışmaları incelendiėinde ise yine bu alışmaların sayısının az olduėu, olanların genellikle anketlerin veya sınavın tmne deėil sadece 1 alt boyutunu ele aldıėı ve PISA 2015 uygulaması ėrenci anketleri iin lme deėiřmezliėi analizleri zelinde alıřma yapılmadıėı grlmřtr (Asil ve Gelbal, 2012; Glleroėlu, 2016; Kıbrıřlıoėlu, 2015; Yandı, Kse ve Uysal, 2017; Yandı, Kse, Uysal ve Oėul, 2017).

1.2. Ama

Bu alıřmanın amacı 2015 yılı PISA testi ėrenci anketlerindeki maddelerin Amerika Birleřik Devletleri ve Trkiye rneklemeleri baėlamında cinsiyete ve kltre dayalı deėiřen madde fonksiyonu ierip iermediėinin belirlenmesi ve bahsi geen iki lke rneklemeleri iin kltre dayalı lme deėiřmezliėinin saėlanıp saėlanmadıėının belirlenmesidir.

Bu genel ama doėrultusunda ařaėıdaki sorulara yanıt aranmıřtır.

- 1) PISA 2015 uygulamasında fen bilimlerine ynelik ėrenci tutum anketlerinde yer alan maddeler Amerika Birleřik Devletleri ve Trkiye kltrnde lme deėiřmezliėi gstermekte midir?
 - a) Yapısal deėiřmezlik boyutunda lme deėiřmezliėi gstermekte midir?
 - b) Metrik deėiřmezlik boyutunda lme deėiřmezliėi gstermekte midir?
 - c) Skalar deėiřmezlik boyutunda lme deėiřmezliėi gstermekte midir?
 - d) Katı deėiřmezlik boyutunda lme deėiřmezliėi gstermekte midir?
- 2) PISA 2015 uygulamasında kullanılan ėrenci anketlerinde yer alan maddeler cinsiyete ve farklı kltre dayalı deėiřen madde fonksiyonu gstermekte midir?

- a) PISA 2015 Türkiye uygulamasındaki öğrenci anketlerinde yer alan maddeler cinsiyete dayalı olarak MTK-OO yöntemine göre değişen madde fonksiyonu göstermekte midir?
- b) PISA 2015 Amerika Birleşik Devletleri uygulamasındaki öğrenci anketlerinde yer alan maddeler cinsiyete dayalı olarak MTK-OO yöntemine göre değişen madde fonksiyonu içermekte midir?
- c) PISA 2015 uygulamasında yer alan öğrenci anketlerindeki maddeler Türkiye ve Amerika Birleşik Devletleri arasında kültüre dayalı olarak MTK-OO yöntemine göre değişen madde fonksiyonu içermekte midir?

1.3. Önem

PISA, son yıllarda modern dünyada eğitim politikalarının belirlenmesi ve diğer ülkelerle kıyaslanması açısından kullanılan ana akım standartlardan biri olmuştur. Aynı anda 70'ten fazla ülkeye ve 500.000'den fazla öğrenciye uygulanan sınavın bu kadar ciddiyetle ele alınması ve sonuçlarının hayati değerler taşıması bu sınavın geçerlik ve güvenilirlik açısından incelenmesini zorunlu kılmaktadır. Zira geçerli ve güvenilir olmayan bir sınavın uygulanması ve sonuçlarının değerlendirilmesi zaman kaybından başka bir şey değildir. Tek bir dilde yazılıp daha sonra diğer dillere çevrilmesi ve kültürel açıdan uyarlanması PISA'yı değişen madde fonksiyonu ve ölçme değişmezliği konularında kırılgan hale getirmiştir.

Alan yazın incelendiğinde PISA'yı psikometrik açıdan ele alan çalışmaların az olduğu, mevcut çalışmaların da genellikle başarı testleri ve iki kategorili değişkenler üzerinden yapıldığı görülmektedir. Bu durum alan açısından bir eksiklik teşkil etmektedir. Tutum anketlerinin PISA sonuçlarının okunmasındaki katkısı yadsınamaz boyuttadır. PISA 2015 uygulaması öğrenci anketleri üzerinden henüz değişen madde fonksiyonu ve ölçme değişmezliği çalışmalarının yapılmadığı gözlenmiştir. Bu açıdan bakıldığında yapılacak bu çalışmanın hem değişen madde fonksiyonu hem de ölçme değişmezliği analizini içermesi açısından bütüncül bir yaklaşım sergilemesi öngörülmüş ve alandaki bu boşluğu kapatması hedeflenmiştir. Çalışma ülkeleri karşılaştırmanın yanında cinsiyetler arasındaki değişen madde fonksiyonunu da incelemiştir.

Tüm bu durumlar bir araya gelerek yapılacak çalışmanın alanda yapılan diğer çalışmalar da göz önüne alınarak PISA sınavlarının sonuçlarının kullanılma ve

yorumlanma sürecine destek olması açısından önemli olduđu ve birçok yönden alana katkı sağlayacağı düşünülmektedir.

1.4. Sayıtlar

PISA 2015 uygulamasına katılan öğrencilerin, öğrenci anketi tutum maddelerini yanıtlarken gerçek tutumlarını yansıttıkları varsayılmıştır.

1.5. Sınırlılıklar

Bu çalışmada değışen madde fonksiyonu incelemesi iki kültür ve tek bir yöntemle, ölçme değışmezliğı incelemesi ise iki kültür ile sınırlıdır.

BÖLÜM 2

KAVRAMSAL ÇERÇEVE

Bu bölümde ölçme değişmezliği ve değişen madde fonksiyonu üzerine alan yazından kavramsal bilgiler verilmiş, analiz yöntemleri incelenmiş ve alanda çalışmanın konusuyla ilişkilendirilebilecek diğer çalışmalardan bahsedilmiş ve bu çalışmalar özetlenmiştir.

2.1. Ölçme Değişmezliği

Ölçme değişmezliğinin test edilmesi son yıllarda sosyal bilimlerde ve testler açısından önemli bir analiz haline gelmiştir. Kısaca tanımlamak gerekirse ölçme değişmezliği birbirinden farklı gruplarda uygulanan ortak bir testin maddelerinin alt gruplarda aynı anlama gelmesidir (Cheung ve Rensvold, 2002). Maddelerin farklı gruplarda aynı anlama gelmesi, testin ilgili grupların kıyaslanmasında kullanımının sorun yaratmayacağı anlamına gelmektedir. Ölçme değişmezliğinin sağlanamaması gruplar arasındaki puan farklılıklarının kaynağının (akademik performans, psikometrik farklılıklar vb.) bilinemeyeceği anlamına gelir ve dolayısı ile testten elde edilen puanların yorumlanması anlamsızlaşır.

Ölçme değişmezliği alt gruplar açısından incelenemediği gibi testin faktör yapısının zamana göre değişip değişmemesi veya testin uygulanma metodu açısından da (kağıt-kalem testi, internet üzerinden çevrimiçi bilgisayar testleri, vb) incelenebilir (Kline, 2015).

Cheung ve Rensvold'a (2002) göre ölçme değişmezliği çalışmalarının konularını genellikle kültürler, diller ve cinsiyetler arası çalışmalar, uyarlama ve çeviri testler oluşturur. Bu açıdan bakılarak PISA 2015'te kullanılan öğrencilere yönelik tutum maddelerinin hem Türkiye açısından çeviri bir test olması hem de kültürlerarası bir çalışma olması açısından ölçme değişmezliğinin doğrulayıcı faktör analizi yöntemiyle test edilmesi uygun görünmektedir.

2.2. Çoklu Grup Doğrulayıcı Faktör Analizi

Yapısal eşitlik modeli analizlerinden biri olan çoklu grup doğrulayıcı faktör analizinde amaç, basitçe, aynı anda birden çok gruba doğrulayıcı faktör analizi

uygulamaktır (Tabachnick ve Fidell, 2014). Elde edilen parametrelerin gruplar arasındaki yakınsaklığı analizin inceleme konusudur. Bu parametrelerin belirli sınırlandırmalar içinde yakınsaklık sağlaması ilgili testin alt grupları arasındaki ölçme değişmezliği için aşama aşama kanıtlar sağlar.

ÇGDFA analizi ölçme değişmezliğini tespit etmekte esnek bir yaklaşım sunar. Belli parametrelerin sınırlandırılması, görece daha geniş bir modelden başlar ve daha sınırlı bir modele doğru incelenir. Bu aşamaların her biri kendisinden sonraki aşamalar için ön koşul olarak kabul edilir. Eğer bir aşama veri uyumu sağlanıyorsa, bir üstündeki daha sınırlı bir aşama tespit edilir. Bu uyum katı değişmezliğe kadar incelenir. Katı değişmezlikten önceki herhangi bir aşama uyum indekslerine göre uygunluk göstermiyorsa analiz orada sonlandırılır ve testin yalnızca o aşamaya kadar ölçme değişmezliği gösterdiği sonucuna ulaşılır (Kline, 2015; Önen, 2009).

Meredith (1993) ÇGDFA ile ölçme değişmezliğini yapısal, metrik, skalar ve katı değişmezlik olmak üzere 4 boyutta test etmeyi önermiştir. Araştırmanın üzerinde duracağı aşamalar bunlar olmuştur ve bir sonraki kısımda her bir alt boyut ayrıntılı olarak incelenmiştir.

2.3. Yapısal Değişmezlik

Yapısal değişmezlik, ölçme değişmezliği hipotezlerinin test edilmesi için uygulanan ve analiz parametrelerini en az sınırlandıran ilk aşamadır. Alt gruplardaki faktör sayısının aynı olması ve faktöre ait göstergelerin benzer olması beklenmektedir. Bunun dışında kalan diğer indeksler (Regresyon sabitleri, hata varyansları ve faktör yükleri) serbesttir. Gruplardan elde edilen verinin yapısal değişmezlik aşamasının uyum indekslerine uymaması o testin ölçme değişmezliğini sağlamadığı ve ilgili testin alt gruplar açısından farklı şekillerde anlamlandırıldığı anlamına gelir (Kline, 2015).

Model oluşturulma aşamasında ihtiyaç duyulması halinde her bir alt grup için farklı modifikasyonlara gidilebilir. Burada dikkat edilmesi gereken nokta, yapılan işlem ve modifikasyonların artık temel modelin kesinleşmiş bir parçası olması gerekliliğidir ve ölçme değişmezliğinin diğer aşamalarında da değiştirilmeden aynen uygulanması gerekmektedir (Wang ve Wang, 2012). Yapısal değişmezlik aşamasının hipotezlerinin kabul edilmesi ardından metrik değişmezlik aşamasına geçilir.

2.4. Metrik Değişmezlik

Metrik değişmezlik testinin yapılabilmesi için öncelik o testin alt gruplarda yapısal değişmezliğinin sağlanmasıdır (Kline, 2015). Wang ve Wang (2012) metrik değişmezliği ya da başka bir ismiyle zayıf değişmezliği, kısaca faktör yükü parametrelerinin testin uygulandığı alt gruplarda değişmez olup olmadığının ölçülmesi olarak tanımlamaktadır. Bu durumdan da anlaşılmaktadır ki faktör yüklerinin değişmez olması, farklı grupların teste ait maddeleri benzer şekilde yorumladığı anlamına gelmektedir.

Gregorich (2006) metrik değişmezliğin sağlanamaması durumunda iki farklı çıkarım yapılabileceğini belirtir. İlki, ortak faktörlerden bir ya da birden fazla faktörün gruplar arasında aynı anlama gelmediğidir. Yapılabilecek diğer yorum ise ilgili yapıya ait maddelerin alt gruplardan bir ya da birden fazlası için yanlılık içerebileceği yönündedir. Metrik değişmezlik aşamasının hipotezlerinin kabul edilmesinin ardından skalar değişmezlik aşamasına geçilir.

2.5. Skalar Değişmezlik

Skalar değişmezlik testinin ön koşulu ilgili testin metrik değişmezliğinin kabulüdür. Faktör yükü parametrelerinin yanı sıra bu aşamada regresyon sabitlerinin de alt gruplar açısından değişmez olup olmadığı test edilir (Wang ve Wang, 2012).

Skalar değişmezliğin reddedilmesi, testin alt gruplar için yönü belli olmayan şekilde avantaj veya dezavantaj sağlaması olarak yorumlanabilir. Skalar değişmezliğin sağlanması durumunda teste ait maddelerin alt gruplar için yanlılık göstermediği yorumu yapılabilir. Tersten okunduğunda bu bilgi, skalar değişmezliği sağlayan bir testin alt gruplar için yansız maddeler içermesi olarak yorumlanabilir. (Gregorich, 2006).

Kline (2015) skalar değişmezliği bir test üzerinden belli alt gruplara yönelik olarak çeşitli çıkarımları kesin olarak yapabilmek için testin sahip olması gereken en minimal boyut olarak görmektedir. Skalar değişmezliğin sağlanamadığı bir test alt gruplar açısından anlamlı bir şekilde karşılaştırılabilir değildir. Skalar değişmezlik aşamasının hipotezlerinin kabul edilmesinin ardından katı değişmezlik aşamasına geçilir.

2.6. Katı Değişmezlik

Katı değişmezlik, ölçme değişmezliğinin en yüksek seviyesidir. Katı değişmezliğin ön koşulu skalar değişmezliğinin sağlanmasıdır. Katı değişmezlikte diğer alt değişmezlik türlerinin getirdiği ön kabullerin üzerine ek olarak hata varyansları da sabitlenmektedir. Katı değişmezliğin kabulünün oldukça zor olduğu ve birçok araştırmada sağlanamadığı görülmüştür. Katı değişmezliğin kabulünde test alt gruplar bağlamında her türlü karşılaştırma için manidar ve tutarlı sonuçlar sunacaktır. Araştırmacıların bir kısmı ise katı değişmezliğin pratikte önemini ve gerekliliğini sorgulamış ve uygulanmasının kritik önem arz etmediği sonucuna varmışlardır (Bentler, 2006; Kline, 2015; Karaduman, 2017).

2.7. Uyum İyiliği İndeksleri

Araştırmada kullanılan veri uyum indeksleri ve kabul boyutları bu kısımda kısaca özetlenmiştir.

Yapısal eşitlik modelinde yıllar süren çalışmalar ve geliştirmeler sonucunda ortaya koyulan literatür oldukça geniştir. Bu durum genel olarak her çalışmanın birbiriyle bazı ortak yönleri olmasına rağmen uyum indeksi kabul boyutları açısından farklılığa sebep olmaktadır. Bu açıdan bakıldığında çalışmalarda rapor edilecek değerler ve bu değerlerin kesme noktaları konusunda ortak bir kanıdan söz edilememektedir. Her bir uyum indeksi boyutunun kendi açısından güçlü ve zayıf noktaları vardır. Bu durum göz önünde bulundurularak raporlanacak değerler bir takım ortak noktalar da göz önünde bulundurularak araştırmacının inisiyatifindedir (Bentler, 2006; Tabachnick ve Fidell, 2014; Kline, 2015; Hooper, Coughlan ve Mullen, 2008).

Kline (2015) ölçme değişmezliği incelemelerinde kullanılmak üzere 4 uyum indeksinin kullanımının makul olduğunu belirtmiştir. Bunlar ki-kare, RMSEA CFI ve SRMR değerleridir. Ancak ki-kare değerinin örneklem büyüklüğünden oldukça etkilenmesi dolayısıyla analizlerde ki-kare değerinin modeller arası karşılaştırma amacıyla kullanılmamasına karar verilmiştir. Bu değerlerin kesme noktaları ve kısaca anlamları aşağıda verilmiştir.

RMSEA (Yaklaşık hataların ortalama karekökü): Bu indeks verilerin modele karşı verdiği hata büyüklüğü olarak tanımlanabilir. Oldukça güçlü bir indekstir ve alan yazında uyum kötülüğü indeksi olarak da bilinir. Alt değeri 0 dan başlar ve üst değerinin bir sınırı

yoktur. Çeşitli araştırmacılar bu değer için bir kabul aralığı belirlerken farklı kesme noktaları kullanmışlardır (Kline, 2015). RMSEA için kabul edilen kesme noktaları son araştırmacılar açısından 15 yılda ciddi anlamda düşüş göstermiştir (Hooper, Coughlan ve Mullen, 2008). Maccallum, Browne ve Sugawara (1996) RMSEA için kesme noktasını .08 ile .10 arasında kabul edilebilir, 0.08'in altında ise iyi uyum gösteren durum olarak belirlemişlerdir. Browne ve Cudeck (1993) ise bu değerleri .05 ve altı ise iyi uyum, .10 ve üzeri ise ciddi bir problem olarak görmekle birlikte kesin bir kanıya varılamayacağını belirtmişlerdir. Steiger (2007) .07'nin altındaki değerlerin iyi uyum için uygun olduğunu belirtmişlerdir (akt. Hooper, Coughlan ve Mullen, 2008). Hu ve Bentler (1999) .06'nın altının iyi uyuma işaret ettiğini belirtmişlerdir. Bu çalışmada da .10'un altında görülen değerler veri uyumu için uygun kabul edilmiştir.

SRMR (Artık ortalamaların standardize edilmiş karekökü) örneklemin kovaryans matrisinin artıkları ile hipotezlenmiş kovaryans modeli arasındaki farkın karekökünün standardize edilmiş halidir. SRMR'ın Sıfıra yakın olması veri uyumunun iyiliğini arttırmakla beraber modeldeki parametre ve örneklem büyüklüğü arttıkça düşmesi beklenir. (Hooper, Coughlan ve Mullen, 2008). 0 ile 1 arasında değer alan bu uyum indeksi için genellikle kabul edilen kesme noktası .08'in altındaki değerlerdir (Hu ve Bentler, 1999). Bu çalışmada da .08'in altında görülen değerler SRMR açısından veri uyumu için uygun kabul edilmiştir.

CFI (Karşılaştırmalı uyum indeksi) mevcut model ile yokluk modeli arasında karşılaştırma yapar. 0 ile 1 arasında değer alır. Hu ve Bentler (1999) için iyi uyum göstergesi .95 ve üzeri olarak kabul edilmiştir. Bu çalışmada da CFI için .95 ve üzeri iyi uyum göstergesi olarak kabul edilmiştir.

TLI (Tucker-Lewis İndeksi) normlaştırılmamış bir indekstir ve ki-kare ve serbestlik derecesini de hesaba katar ve 0 ile 1 arasında değer alır. Hu ve Bentler (1999) .95 ve üzeri değerlerin iyi uyumu işaret ettiğini belirtmişlerdir. Çalışmada TLI açısından kesme noktası .95 ve üzeri olarak kabul edilmiştir.

Modeller arasındaki ölçme değişmezliği aşamalarının kabulünde kesme noktaları olarak alan yazında genel olarak kabul gören Cheung ve Rensvold'un (2002) önerdiği üzere modeller arasındaki CFI farkının değişimi (ΔCFI) gözlenmiştir. Bu değişimin -.01 ile +.01 aralığında olması halinde modelin o aşama için ölçme değişmezliği gösterdiği şeklinde yorumlanmıştır.

2.8. Madde Yanlılığı ve Değişen Madde Fonksiyonu

Günümüzde geniş katılımlı sınavların önemi günden güne artmaktadır ve bu durum ilgili testler açısından geçerlik sorunlarını gündeme getirmektedir. Testlerin geçerliğini tehdit eden bir durum olarak madde yanlılığı bu sorunlar açısından öne çıkmaktadır. Madde yanlılığı, ölçme değişmezliğinde olduğu gibi test açısından bir geçerlik sorunudur. Kısaca tanımlamak gerekirse bir testi alan alt gruplardan birinin diğer gruptaki katılımcılara göre belli bir maddeyi sistematik olarak doğru yanıtlama ihtimalinin düşmesi durumudur (Zumbo, 1999).

Madde yanlılığının tarihi oldukça eskiye dayanmaktadır. 1910 yılında Alfred Binet'nin gerçekleştirdiği bir test esnasında ilgili teste tabii çocukların zihinsel kapasiteleriyle ilgili olamayacak bir takım cevaplar verdikleri görülmüştür. Araştırmasını derinleştiren Binet bu farkın kabaca toplum içinde farklı sosyoekonomik özelliklere ait ailelerden gelen çocuklar arasında ortaya çıktığını fark etmiştir. Binet yanlılık gösterdiğini düşündüğü maddeleri testten çıkarmıştır (akt. Camilli ve Shepard, 1994).

Değişen Madde Fonksiyonu ise farklı alt gruplardan gelen aynı yetenek düzeyine ait katılımcıların belli bir maddeye verdikleri yanıtların farklılaşmasıdır. Ancak bu farklılaşma iki şekilde ortaya çıkabilir. Birinci farklılık madde etkisi olarak adlandırılan belli bir yeteneğe ait gerçek farkları gösteren farklılıktır. İkinci farklılık ise maddenin adil olmayan şekilde bir alt gruba avantaj sağlaması durumudur. Bu açıdan bakıldığında DMF gösteren maddelerin tamamının yanlılık gösterdiği gibi bir çıkarım yapmak yanlıştır. (Clauser ve Mazor,1998; Zumbo, 1999).

Alan yazın incelendiğinde 2 farklı DMF tipi görülür. Birinci tip tek biçimli olmayan DMF (non-uniform DIF) olarak adlandırılır. Bu tipte maddenin belirli bir düzeye kadar belli bir gruba farklı, o düzeyden sonra başka bir gruba farklı davranması söz konusudur. Tek biçimli olan DMF'de (uniform DIF) ise bir maddenin alt gruplarda bütün yetenek gruplarına farklı davranması durumudur (Karami, 2012).

DMF analizlerinde çok kategorili maddeler için KTK'ya dayanan Ordinal Lojistik Regresyon, POLY-SIBTEST ve MTK'ya dayanan BILOG-MG ve MTK-OO gibi farklı yöntemler kullanılmaktadır. Bu çalışmada da madde tepki kuramının klasik test kuramına göre avantajları göz önünde bulundurularak analizlerin Madde Tepki Kuramı Olabilirlik Oranı üzerinden yapılmasına karar verilmiştir.

2.9. Madde Tepki Kuramı Olabilirlik Oranı (MTK-OO)

MTK-OO yönteminin uygulanmasında öncelikle referans ve odak grupları belirlenir ve bu gruplardan elde edilen madde parametreleri karşılaştırılır. MTK-OO testi için yokluk hipotezi “belirlenen gruplardan yola çıkarak hesaplanan madde parametreleri arasında anlamlı farklılık yoktur.” şeklindedir. Bu amaç doğrultusunda sınırlandırılmış model ve genişletilmiş model adı altında iki model oluşturulur. Sınırlandırılmış modelde odak ve referans gruplarının madde parametrelerinin eşit olduğu varsayılır. Genişletilmiş modelde ise i. maddesine ait parametrelerin farklılaştığı, diğer madde parametrelerinin ise eşit olduğu varsayımından yola çıkılır. Her bir alt test için bir tane sınırlandırılmış model ve madde sayısı kadar genişletilmiş model kurgulanır ve olabilirlik fonksiyonu elde edilir. Sınırlandırılmış ve genişletilmiş modellerden elde edilen olabilirlik fonksiyonunun logaritmalarının farkları alınır ve G^2 değerlerine ulaşılır (Thissen, 2001).

$$G^2 = -2LL_C - (-2LL_A)$$

Elde edilen G^2 değerinin 3,84 üzerinde olması ilgili maddenin değişen madde fonksiyonu gösterdiği şeklinde yorumlanır. (Thissen, 2001). Greer (2004) ise bu kesme noktalarını şu şekilde sınıflandırmıştır (akt. Akın, 2015):

$3,84 < G^2 < 9,4$ İhmal edilebilir düzeyde değişen madde fonksiyonu

$9,4 \leq G^2 < 41,9$ Orta düzeyde değişen madde fonksiyonu

$G^2 \geq 41,9$ Yüksek düzeyde değişen madde fonksiyonu

Zieky (1993) yaptığı çalışmada A seviyesinde DMF gösteren maddelerin yanlılık açısından yorumlanmamasını, B ve C seviyesindeki DMF gösteren maddelerin ise yanlılık açısından incelenmesini önermiştir (akt. Gierl, 1999).

Analizin uygulanmasında çok kategorili maddeleri destekleyen Aşamalı Tepki Modeli (Graded Response Model) kullanılmıştır. Alfa parametrelerindeki farklılaşma tek biçimli olmayan DMF’yi işaret ederken, beta parametrelerindeki farklılaşma tek biçimli DMF’yi işaret etmektedir. DMF’nin avantaj sağladığı grubu tespit etmek için b_1 , b_2 ve b_3 değerlerindeki farklılaşma incelenir. Bu değerlerdeki yük iki grup için karşılaştırılır ve düşük değer gösteren gruptaki katılımcıların, aynı yetenek düzeyini gösteren diğer gruptaki katılımcılara göre ilgili maddeyi işaretleme eğilimlerinin daha yüksek olduğu sonucuna ulaşılır (Baranik, Meade, Lakey, Lance, Hu, Hua, ve Michalos, 2008).

2.10. İlgili Araştırmalar

Bu bölümde literatür incelenmiş, alanda yapılan ölçme değişmezliği ve değişen madde fonksiyonu çalışmaları özetlenmiştir.

2.11. Ölçme Değişmezliği Çalışmaları

Uyar (2011) PISA 2009 uygulamasında öğrenmeyi öğrenme stratejileri anketinde bulunan maddelerden oluşturulan modelin ölçme değişmezliğini Türkiye örneklemini için cinsiyet, okul türü ve istatistiki bölgeler üzerinden çoklu grup doğrulayıcı faktör analizi yöntemi ile yapmıştır. Sonuç olarak cinsiyet ve okul türü açısından yalnızca yapısal ve metrik değişmezlik aşamaları sağlanmış. İstatistiki bölge birimleri açısından ise yapısal metrik skalar ve katı değişmezlik aşamalarının hepsi sağlanmıştır.

Hansson ve Gustafsson (2013) TIMSS 2003 uygulamasının 8. Sınıflar için olan matematik testinin İsveç örneklemini üzerinde sosyo-ekonomik gruplar açısından karşılaştırma yapma amacını gütmektedir. Alt grupları İsveç'te doğmayan göçmen ailelerden gelen öğrenciler, İsveç'te doğan göçmen ailelerden gelen öğrenciler ve İsveçli öğrenciler oluşturmaktadır. Sonuç olarak çalışmada sadece yapısal boyutta ölçme değişmezliğinin sağlandığı görülmüş, ayrıca üst boyuttaki değişmezlik basamaklarının sağlanamamasının olası sebepleri ve eğitim açısından politik ve bilimsel çözümleri tartışılmıştır.

İmrol (2017) PISA 2012 uygulamasının Türkiye örnekleminde matematiğe yönelik uygulanan tutum anketlerinden oluşturduğu 2 farklı modelin (motivasyon ve öz-inanç) ölçme değişmezliğini çoklu grup doğrulayıcı faktör analizi ile sosyo-ekonomik düzey, okul türü ve istatistiki bölgeler açısından incelemiştir. Araştırmanın sonucunda motivasyon yapısı için bütün alt grup ve düzeyler için yapısal, metrik, skalar ve katı değişmezlik aşamalarının tümünü karşılamış, öz-inanç yapısı için ise okul türü için yapısal, sosyo-ekonomik düzeye göre metrik boyutlarına kadar ölçme değişmezliği sağlarken istatistiki bölge birimlerinde ÇGDFA aşamalarının tüm boyutlarını karşılamıştır.

Karaduman (2017) Sınav Stresi Ölçeğinin Türk kültürüne uyarlama çalışmasını yaptıktan sonra aynı ölçeğin ölçme değişmezliğini Türkiye örneklem için çoklu grup doğrulayıcı faktör analizi ile incelemiştir. Ölçeğin inceleme yapılan tüm alt gruplarda

(cinsiyet, sınıf düzeyi, okul türü) ölçme değişmezliğinin tüm aşamaları katı değişmezliğe kadar başarıyla sağlanmıştır.

Uzun ve Öğretmen (2010) 1999 yılı TIMSS-R uygulaması Türkiye örnekleminde fen başarısına etki eden faktörlerin incelenmesi için uygulanan anketin cinsiyetler arası ölçme değişmezliğini ÇGDFA yöntemiyle analiz etmişlerdir. Sonuç olarak tutum alt boyutu için skalar değişmezliğe kadar, sınıf için öğrenci etkinlikleri, önem ve öz yeterlik alt boyutları için ise metrik değişmezliğe kadar ölçme değişmezliği sağlanmıştır.

Karakoç Alatl ı ve Çokluk Bökeoğlu (2018) PISA 2012 uygulamasında fen okuryazarlığı, matematik okuryazarlığı ve okuma becerileri testlerinin ÇGDFA ile ölçme değişmezliğini test etmiş ve bunun yanı sıra GADM, SIBTEST ve MTK-OO yöntemleriyle DMF analizlerini gerçekleştirmişlerdir. Sonuç olarak fen okuryazarlığı testi yalnızca yapısal değişmezlik göstermekte ve %34 oranında DMF’li madde içermektedir. Matematik okuryazarlığı testi ise yalnızca yapısal değişmezlik göstermekte ve %35 oranında DMF’li madde içermektedir. Son olarak okuma becerileri testlerinin analizi yapılmış yalnızca yapısal değişmezlik aşamasın sağlandığı ve %22 oranında DMF’li madde içerdiği görülmüştür. Olası yanlışlık kaynakları ise çeviri hataları, belli bir grubun ilgili dile aşinalığı ve anlamsal farklılıklar olarak belirlenmiştir.

Luo (2010) Rosenberg Benlik Saygısı Ölçeğinin İngiliz ve Çin Halk Cumhuriyeti kökenli öğrenciler arasında ÇGDFA yöntemiyle ölçme değişmezliğini incelemiş ve ölçeğin yalnızca yapısal ve metrik değişmezliği sağladığı, skalar ve buna bağlı olarak katı değişmezlik aşamasında ölçme değişmezliği sağlanamamıştır ve bundan sonraki kültürler arası çalışmalarda kullanım açısından dikkatli olunması önerilmiştir.

Gülleroğlu (2016) PISA 2012 matematiğe yönelik duyuşsal özelliklerin ölçme değişmezliğini Türkiye örneklemini açısından cinsiyet değişkenine göre ÇGDFA yöntemiyle incelemiştir. Sonuç olarak matematik öz yeterliği boyutu hiç bir aşamada ölçme değişmezliği sağlamazken, matematik benlik algısı boyutu yapısal, matematik kaygısı boyutu metrik ve matematiğe yönelik ilgi boyutu skalar aşamada ölçme değişmezliği sağlamıştır.

Pauwels (2018) Achievement Motive Ölçeğinin ekranda tek madde ile ekranda birden çok madde gösterilen alt boyutlarında ölçme değişmezliğini ÇGDFA yöntemiyle ve 3’lü bir sınıflandırmayla incelemiş, alt boyutların hepsinin skalar düzeyde ölçme değişmezliği gösterdiği sonucuna varmıştır.

Karakoç Alatl ı, Ayan, Polat Demir ve Uzun (2016) tarafından TIMSS 2011 uygulaması 4. sınıf matematik testinin Türkiye, Amerika Birleşik Devletleri, Japonya ve

İngiltere örneklemi için ölçme değişmezliği incelemesi yapılmış ve ölçme değişmezliğinin bütün alt gruplar için yalnızca yapısal boyutta karşılanabildiği görülmüştür. Bunun üzerine ölçme değişmezliğini bozan alt grubu tespit için alt grupların 2'li ve 3'lü incelemesi yapılmış ancak metrik değişmezlik bu durumda da sağlanamamıştır. Sonuç olarak TIMSS 2011 için ilgili ülkeler arasında sağlıklı bir karşılaştırma yapılamayacağı kanısına varılmıştır.

2.12. Değişen Madde Fonksiyonu Çalışmaları

Ayan (2011) PISA 2009 fen okuryazarlığı testinin 2. Kitapçığı için Mantel Haenszel ve Lojistik Regresyon yöntemleriyle Türkiye örnekleminde cinsiyet üzerinden DMF belirleme çalışması yapmıştır. Sonuç olarak 4 maddenin DMF içerdiği bulgusuna ulaşılmış ve bu maddelerin 3 tanesi kadınlar lehine çalışırken 1 madde ise erkekler lehine çalışmıştır. MH ve LR yöntemlerinin genel olarak uyumlu sonuç gösterdiği sonucuna ulaşılmıştır.

Langer ve arkadaşları (2008) Sağlıkla İlgili Hayat Kalitesi Testi'nin geliştirilmesi bağlamında DMF içeren maddeleri MTK-OO yöntemiyle tespit etmek ve genel olarak sağlık bilimleri için kullanılan test ve envanterlerin geçerliğini yükseltme amacı ile çalışma yürütmüşlerdir. Çalışmanın bir diğer amacı ise MTK-OO testinin gücünü ve önemini vurgulamaktır. Sağlık sorunu yaşayan ve sağlıklı olan iki ayrı alt gruptaki toplam 5429 çocuk üzerinde gerçekleştirilen çalışma sonucu testin genelinde neredeyse bütün maddelerde DMF bulunmuş ve olası yanlılık sebepleri incelenmiştir.

Gierl, Khaliq ve Boughton (1999) Alberta Education Mathematics and Science başarı testinin 1994, 1995, 1996 ve 1997 yıllarında 3, 6 ve 9. sınıflar için rastgele seçilen 6000 kız ve 6000 erkek toplam 12000 öğrenciden oluşan örneklem için MH, LR ve SIBTEST üzerinden cinsiyetler arasında madde yanlılığı incelemesi yapmışlardır. Sonuç olarak, üç yöntem de DMF açısından uyum gösterirken MH yöntemi daha tutucu bir yöntem olarak gözlenmiştir. DMF'li madde oranı %8.4 ile %17.3 arasında bulunmuştur.

Lyons-Thomas, Sandilands ve Ercikan (2014) PISA 2009 uygulamasının matematik okuryazarlığı alt testi için 35 madde üzerinden Şangay, Kanada, Türkiye ve Finlandiya örneklemelerinden elde edilen 21670 kişilik 15 yaş grubu öğrencilerini, Linn-Harnisch, SIBTEST ve MH yöntemleriyle cinsiyetler açısından değişen madde fonksiyonu bağlamında incelemişlerdir. Sonuç olarak çoktan seçmeli maddeler erkek öğrenciler lehine

çalışırken yazılı cevap beklenen maddeler kadın öğrenciler lehine çalışmıştır. Cinsiyetler arasındaki DMF içeren madde oranı en çok Finlandiya örnekleminde elde edilmiştir.

Atalay (2010) PISA 2006 uygulaması fen bilimlerine yönelik öğrenci tutum anketlerinin maddelerinde Türkiye ve ABD örneklemi için DMF gösteren maddelerin POLY-SIBTEST ve OLR yöntemleriyle tespiti üzerine bir çalışma yürütmüştür. Her iki yöntemle yapılan analizlerde 30 maddede DMF görülmüştür. Analiz yöntemleri birbiriyle tutarlı sonuç vermekle birlikte DMF'nin düzeyleri üzerinden bir takım farklılaşmalar bulunmuştur.

Woods (2008) normalliğin sağlanamadığı veri setlerinin MTK-OO üzerindeki etkisi üzerine çalışmıştır. Sonuç olarak MTK-OO yöntemi normalliğin sağlandığı veriler için iyi performans gösterirken bir ya da birden fazla grubun çarpıklığının yüksek olduğu durumunda tutarsız sonuçlar vermiştir.

Sheppard, Han, Colarelli, Dai ve King (2006) 13 alt ölçek ve 206 maddeden oluşan Hogan Kişilik Envanterinin madde yanlılığını cinsiyet ve 2 ırksal grup (siyah ve beyaz ırk) açısından MH yöntemiyle test etmişlerdir. Maddelerin %38.4'ü cinsiyet açısından, %37.3'ü ise ırksal gruplar açısından DMF göstermiştir. Araştırmacılar bu sonuçlar üzerinden kişilik envanterlerinin yorumlanma ve oluşturulma süreçlerinde DMF açısından pratik ve teori üzerinde çeşitli açıklamalar ve önerilerde bulunmuşlardır.

Baranik ve arkadaşları (2008) Rosenberg Self-Esteem ölçeğini 8 ülke örneklemini üzerinden değişen madde fonksiyonu için MTK-OO yöntemiyle analiz etmişler ve neredeyse bütün maddelerde belli gruplar için DMF'ye rastlamışlardır. Ortaya çıkan sonuçlar kültürler arası uygulamalar, ölçekler, çok kategorili maddeler ve madde tepki kuramı açısından tartışılmıştır.

Demir (2013) PISA 2009 matematik okuryazarlığı testinin cinsiyet ve kültür üzerinden DMF'li madde gösterip göstermediğini araştırmıştır. Analizde kullanılan yöntemler MH, SIBTEST ve LR'dir. Araştırmanın örneklemini Almanya, Türkiye, ABD ve Finlandiya'dan 5. kitapçığa cevap veren öğrenciler oluşturmaktadır. Sonuç olarak matematik okuryazarlığı testinin %60 oranında DMF'li madde gösterdiği, yöntemlerin ortalama olarak birbiriyle tutarlı sonuçlar verdiği, cinsiyete yönelik DMF gösteren madde sayısında yıllara göre bir azalma olduğu belirlenmiştir.

Alandaki araştırmalar incelenmiş ve sonuç olarak görülmüştür ki, ölçme değişmezliği ve değişen madde fonksiyonu genellikle geniş katılımlı ve bireysel ölçekler üzerinden incelenmiş, cinsiyet, kültür, dil, sosyo-ekonomik düzey, okul türü gibi yapılar üzerinden analizler yapılmıştır. Analizler birçok farklı kuram ve yöntem üzerinden

gerçekleşmiş ve genellikle birbiriyle örtüşen sonuçlar elde edilmiştir. Çalışmaların sonuçları incelendiğinde; dil ve kültürün diğer yapılardan daha çok DMF gösterdiği ve ölçme değişmezliği açısından katı değişmezlik aşamasını sağlayamadığı görülmüştür. Tüm bu bilgilerin ışığında farklı dil ve kültürler üzerinden yapılan çalışmaların önemi ortaya çıkmıştır.

DMF ve ölçme değişmezliği çalışmaları PISA, TIMSS ve PIRLS gibi geniş ölçekli çalışmalar açısından ekstra önem taşımaktadır. Geniş katılımlı sınavların ana amaçlarından birisi ülke ve alt grupları birbiriyle kıyaslamaktır. Yanlılık içeren ve ölçme değişmezliği sağlanmamış testlerin alt gruplar açısından karşılaştırılması anlamsız olacaktır ve ilgili testin geçerliğini önemli ölçüde ve olumsuz yönde etkileyecektir.



BÖLÜM 3

YÖNTEM

Bu bölümde araştırmanın türü, çalışma grubu, verilerin elde edilmesi, verilerin analizi ve analiz için kullanılan bilgisayar programı hakkında bilgi verilmiştir.

3.1. Araştırma Modeli

Çalışma, 2015 yılı PISA Testi öğrenci anketlerinin kültüre ve cinsiyete göre değişen madde fonksiyonu içerip içermediğinin ve iki kültür arasında ölçme değişmezliğinin sağlanıp sağlanmadığının incelenmesine yöneliktir. Çalışma, bu yapısı itibariyle betimsel bir çalışmadır.

3.2. Evren-Örneklem

2000 yılından beri yapılmakta olan ve 2015 yılında 6.'sı yapılan PISA 2015 uygulamasına aralarında Türkiye'nin de bulunduğu 35 OECD ülkesi ile birlikte toplam 72 ülkede 15 yaşındaki 29 milyon öğrenciyi temsil etmek üzere 540.000'e yakın öğrencinin katılımı sağlanmıştır (MEB, 2016).

Çalışmanın Türkiye örnekleminde evreni oluşturan 1.324.089 öğrenci adına İstatistiki Bölge Birimleri Sınıflamasına göre belirlenen ve tabakalı seçkisiz örnekleme yoluyla seçilen 61 il ve 187 okuldan sağlanan 5895 öğrencinin katılımı sağlanmıştır (MEB, 2016).

Çalışmanın Amerika Birleşik Devletleri ayağında ise ABD evrenini temsilen 177 okuldan 5712 öğrencinin katılımı sağlanmıştır (OECD, 2016).

Araştırmaya konu olan örneklemelerin Türkiye ve ABD olarak seçilmesinin sebebi anketlerin kültürel farklılıklara ve çeviriye bağlı olumsuzluklara açık olmasıdır ve bu durumunda değişen madde fonksiyonunun ve ölçme değişmezliğinin olası nedenleri arasında yer almasıdır. Analizde kullanılan verilerin ülkeler ve cinsiyetler açısından frekans ve yüzde istatistikleri Tablo 1'de verilmiştir.

Tablo 1

Ülkeler ve Cinsiyetler Açısından Örneklem

Ülkeler	Erkek		Kadın		Toplam	
	f	%	f	%	f	%
Türkiye	2957	50.2	2938	49.8	5895	50.8
ABD	2858	50	2854	50	5712	49.2

Tablo 1 incelendiğinde hem ülke örneklemlerinin hem de cinsiyet değişkenlerinin katılımcılar açısından eşit dağıldığı görülmüştür.

3.3. Veriler ve Toplanması

Araştırmada, Ekonomik Kalkınma ve İşbirliği Örgütü'nün 2015'te uyguladığı PISA sınavının öğrenci anketi verileri kullanılmıştır. Veriler OECD'nin PISA için hazırladığı resmi internet sitesinden elde edilmiştir (<http://www.oecd.org/pisa/>).

3.4. Verilerin Analizi

PISA 2015 uygulamasında fen okuryazarlığına yönelik duyuşsal performansı ölçmeyi amaçlayan 4'lü likert şeklinde oluşturulmuş 37 maddeden oluşan 6 alt boyut ele alınmıştır. Bu alt boyutlardan ST094 kodlu alt boyut “fen öğrenmekten zevk alma” (enjoyment of science), ST095 kodlu alt boyut “geniş fen konularına ilgi” (interest in science), ST113 kodlu alt boyut “fen öğreniminde araçsal güdülenme” (instrumental motivation to learn science), ST129 kodlu alt boyut “fene yönelik öz yeterlik” (science self-efficacy), ST131 kodlu alt boyut “fen bilimlerine yönelik epistemik inançlar” (epistemic beliefs about science) ve ST146 kodlu alt boyut “bilim aktiviteleri” (science activities) olarak adlandırılmıştır (OECD, 2016). Analizler için tersten kodlanan maddelerin verileri düzeltilmiştir. ST095 testindeki son seçenek olan “bunun ne olduğunu bilmiyorum” seçeneği OECD'nin PISA 2015 Results (Volume I) kitapçığında önerildiği üzere kayıp veri olarak kodlanmıştır.

Verilerin çözümlenmesi ve analizinde betimsel istatistikler ve varsayım incelemeleri için SPSS 23.0 yazılımı ve değişen madde fonksiyonu analizinde MTK-OO yöntemi için ise IRTLRDIF2 yazılımı kullanılmıştır.

Kültürler arası ölçme değişmezliğinin ve varsayımların incelenmesinde ÇGDFA yöntemi kullanılmış ve analizler Windows işletim sistemi için Mplus 7 programıyla yapılmıştır.

ÇGDFA yöntemi yapısal eşitlik modellemesi olarak ölçme değişmezliği tespitinde kullanılmaktadır. Ölçme değişmezliğinde belirli bir ölçme aracına ait iki farklı gruptan gelen verinin farklı kriterler açısından birbirine benzerliği beklenmektedir (Tabachnick ve Fidell, 2014). Belirli bir kültür ve dilde geliştirilen bir testin faktör yapısının, testin uygulandığı başka bir kültürde denk veya benzer olması beklenir. ÇGDFA sırasıyla dört alt yeterlilikte test edilir. Bunlar sırasıyla, yapısal, metrik, skalar ve katı değişmezliktir. (Meredith, 1993). Bu çalışmada sırasıyla bahsi geçen dört hipotez test edilmiştir.

Yapısal değişmezlik aşamasında regresyon sabitleri, hata varyansları ve faktör yükleri serbest tahminlenir. Metrik değişmezlik aşamasında faktör yükleri sabitlenirken regresyon sabitleri ve hata varyansları serbest tahminlenir. Skalar değişmezlik aşamasında yalnızca hata varyansları serbest tahminlenir. Katı değişmezlik aşamasında ise faktör yükleri ve Regresyon sabitlerinin yanı sıra hata varyansları da sabitlenmektedir.

3.5. Ölçme Değişmezliği Varsayımlarının İncelenmesi

Yapısal eşitlik modeli uygulamalarında analizlere başlamadan önce test edilmesi gereken bir takım varsayımlar vardır. Bu varsayımlar sırasıyla kayıp değerlerin incelenmesi, uç değerlerin şncelenmesi, normallik, doğrusallık, eş varyanslılık ve çoklu bağlantılılıktır (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Bu bölümde ilgili varsayımlar sırasıyla incelenmiştir.

Analizlere başlamadan önce verilerin analize uygun hale getirilmesi gerekmektedir. Bu açıdan ilk olarak her bir alt test için hiçbir maddeye cevap vermeyen (unit nonresponse) birimler veri setinden atılmıştır. Veriler daha sonra Little's MCAR yöntemiyle incelenmiş ve kayıp verilerin St113 alt testi haricinde “tamamıyla rastgele kayıp” (missing completely at random) olmadığı görülmüş ancak veri seti üzerinde yapılan inceleme sonucunda kayıp verilerin örüntüsünün tamamıyla rastgele kayıp olduğuna karar verilmiştir. Veri kaybını azaltmak amacıyla Expectation-Maximization algoritması kullanılarak kayıp verilere atama yapılmıştır.

Tekli uç değer incelemesinde verilerin her bir alt test için z puanlarına bakılmıştır. 6 alt test içinde 3.29 sınır değerini geçen bir veriye rastlanmamıştır.

Bu aşamadan sonra verilere çoklu uç değer incelemesi için Mahalanobis uzaklığı analizi uygulanmıştır. Analiz sonucunda $p < .001$ için St146 testinden 705, St131 testinden 447, St129 testinden 268, St113 testinden 300, St095 testinden 126, St094 testinden 427 veri uç değer olarak belirlenmiş ve veri setinden çıkarılmıştır. Bu işlemin sonucunda Türkiye ve ABD arasındaki 6 alt testin veri setindeki değişim Tablo 2’de verilmiştir.

Tablo 2

Kayıp Verilerin Düzenlenmesinden Sonra Veri Seti

Alt Test	Tüm Veri Sayısı	Hiçbir Maddeye Cevap Vermeyen Birimler Atıldıktan Sonra Kalan Veri Sayısı	Madde Bazında Kayıp Veri Oranları	Mahalanobis Uzaklığı Analizinden Sonra Kalan Veri Sayısı
St146	11607	11045	%0.4 - %1.8	10340
St131	11607	11066	%0.4 - %1.5	10619
St129	11607	11031	%0.6 - %1.6	10763
St113	11607	11070	%0.3 - %1.1	10770
St095	11607	10947	%2.5 - %4.8	10821
St094	11607	11177	%0.3 - %1.0	10750

Çok değişkenli normallik Mplus paket programı üzerinden Mardia’nın (1970) çok değişkenli çarpıklık ve basıklık analizine göre yapılmış ve sonuçlar Tablo 3 ve 4’te verilmiştir. Veriler her bir alt test için ayrı ayrı incelenmiş ve $p < .05$ için verilerin çok değişkenli normallik varsayımı sağlanamamış ve reddedilmiştir.

Tablo 3

Kültüre Göre ABD ve Türkiye Örneklemeleri Çoklu Normallik Testi Çarpıklık Sonuçları

Alt Test	Örneklem Değeri	Ortalama	sd	p
St146	10.127	0.096	0.010	0.0000
St131	2.642	0.032	0.006	0.0000
St129	3.093	0.067	0.009	0.0000
St113	2.663	0.011	0.004	0.0000
St095	1.140	0.019	0.005	0.0000
St094	2.506	0.019	0.005	0.0000

Tablo 4

Kültüre Göre ABD ve Türkiye Örneklemeleri Çoklu Normallik Testi Basıklık Sonuçları

Alt Test	Örneklem Değeri	Ortalama	sd	p
St146	155.006	99.009	0.270	0.0000
St131	71.366	48.000	0.191	0.0000
St129	106.256	80.006	0.242	0.0000
St113	36.842	23.998	0.126	0.0000
St095	38.885	35.002	0.170	0.0000
St094	51.226	34.999	0.163	0.0000

Doğrusallık varsayımı için saçılım diyagramları EK 3'te verilmiştir. Sonuçlar incelendiğinde doğrusallığın karşılandığı sonucuna ulaşılmıştır.

Tek değişkenli normallik incelemesi çarpıklık ve basıklık üzerinden yapılmış ve sonuçlar tablo 5'te verilmiştir.

Tablo 5

Kültüre Göre Türkiye ve ABD Örnekleme Basıklık ve Çarpıklık Değerleri

Alt Boyutlar	Maddeler	Çarpıklık	Basıklık
Fen Öğrenmekten Zevk Alma (st094)	Q01	-0.474	-0.396
	Q02	-0.254	-0.584
	Q03	-0.386	-0.437
	Q04	-0.589	-0.129
	Q05	-0.541	-0.283
Geniş Fen Konularına İlgi (st095)	Q01	0.107	-1.010
	Q02	-0.022	-0.948
	Q03	-0.111	-0.957
	Q04	-0.355	-0.941
	Q05	-0.522	-0.638
Fen Öğreniminde Araçsal Güdülenme (st113)	Q01	-0.697	0.291
	Q02	-0.552	-0.040
	Q03	-0.542	-0.149
	Q04	-0.474	-0.358
	Q01	-0.672	0.128
Fene Yönelik Öz Yeterlik (st129)	Q02	-0.661	-0.006
	Q03	-0.455	-0.448
	Q04	-0.410	-0.415
	Q05	-0.605	-0.153
	Q06	-0.460	-0.376
Fen Bilimlerine Yönelik Epistemik İnançlar (st131)	Q07	-0.373	-0.628
	Q08	-0.342	-0.721
	Q01	-0.861	0.620
	Q02	-0.794	0.773
	Q03	-0.857	0.704
Bilim Aktiviteleri (st146)	Q04	-0.888	0.588
	Q05	-0.692	0.522
	Q06	-0.690	0.436
	Q01	0.721	-0.057
	Q02	0.974	0.264
	Q03	0.674	-0.224
	Q04	0.836	0.027
	Q05	1.373	1.059
	Q06	1.074	0.366
	Q07	1.115	0.472
	Q08	1.041	0.330
	Q09	1.007	0.172

Tablo 5 incelendiğinde maddeler için çarpıklık ve basıklık değerleri st146 alt boyutunda ± 1 e oldukça yakın olduğu, diğer alt boyutlarda ise ± 1 değerinin içinde olduğu için tek değişkenli normallik varsayımını sağladığı çıkarımı yapılır. Çoklu bağlantılılık ve eş varyanslılık incelemesinin sonuçları Tablo 6'da verilmiştir.

Tablo 6

Kültüre Göre Türkiye ve ABD Örnekleme Çoklu Bağlantılılık ve Eş Varyanslılık İstatistikleri (Tolerans Değerleri ve Varyans Artış Faktör Değerleri)

Alt Boyutlar	Maddeler	Tolerans	VIF	Durbin-Watson
Fen Öğrenmekten Zevk Alma (st094)	Q01	0.251	3.980	0.103
	Q02	0.275	3.637	
	Q03	0.191	5.235	
	Q04	0.165	6.052	
	Q05	0.175	5.705	
Geniş Fen Konularına İlgi (st095)	Q01	0.590	1.696	0.529
	Q02	0.436	2.295	
	Q03	0.395	2.532	
	Q04	0.610	1.638	
	Q05	0.592	1.691	
Fen Öğreniminde Araçsal Güdülenme (st113)	Q01	0.308	3.244	0.016
	Q02	0.232	4.305	
	Q03	0.246	4.065	
	Q04	0.326	3.066	
	Q01	0.559	1.789	
Fene Yönelik Öz Yeterlik (st129)	Q02	0.509	1.965	1.056
	Q03	0.485	2.060	
	Q04	0.455	2.199	
	Q05	0.469	2.134	
	Q06	0.468	2.136	
	Q07	0.499	2.005	
	Q08	0.512	1.954	
	Q01	0.375	2.670	
Fen Bilimlerine Yönelik Epistemik İnançlar (st131)	Q02	0.294	3.401	0.118
	Q03	0.268	3.730	
	Q04	0.271	3.688	
	Q05	0.316	3.167	
	Q06	0.327	3.057	
	Q01	0.438	2.282	
Bilim Aktiviteleri (st146)	Q02	0.276	3.623	0.117
	Q03	0.315	3.177	
	Q04	0.277	3.608	
	Q05	0.330	3.031	
	Q06	0.199	5.019	
	Q07	0.192	5.214	
	Q08	0.260	3.845	
	Q09	0.315	3.171	

Eş varyanslılık incelemesi için Tabachnick ve Fidell (2014) Durbin-Watson katsayısının incelenmesini ve 0-4 arasındaki değerlerin varsayımı sağlayacağını belirtmişlerdir. Tablo 6 incelendiğinde her bir alt boyut için bu değerler kestirilmiş ve varsayımın sağlandığı görülmüştür.

Çoklu bağlantılılık varsayımının incelenmesinde ise her bir alt boyut için tolerans ve varyans şişkinlik faktörü incelenmiştir. Tablo 6 incelendiğinde her bir alt testteki maddeler için tolerans değerleri 0.01'den yüksek, VIF değerleri 10'dan küçük olduğu görülmüş, veri setlerinin çoklu bağlantılılık göstermediği ve ilgili varsayımın karşılandığı görülmüştür (Tabachnick ve Fidell, 2014).

ÇGDFA analizi için gerekli varsayımların çok deęişkenli normallik harici karşılandığı belirtilebilir. Çok deęişkenli normalliğin sağlanamaması durumu göz önüne alınarak Satorra-Bentler düzeltmeli en çok olabilirlik teknięi kullanılarak analizler gerçekleştirilmiştir.



BÖLÜM 4

BULGULAR VE YORUMLAR

Bu kısımda ilgili araştırma sorularına yönelik analizler yapılmış ve bulgular incelenmiştir. PISA 2015 uygulaması öğrenci tutum anketlerinde yer alan ve 6 alt testte bulunan 37 madde ÇGDFA ile ölçme değişmezliği ve MTK-OO ile değişen madde fonksiyonu açısından incelenmiştir.

4.1. Birinci Alt Probleme Yönelik Bulgu ve Yorumlar

Problem 1: PISA 2015 uygulamasında fen bilimlerine yönelik tutum ölçmekte kullanılan öğrenci anketlerinde yer alan maddeler Amerika Birleşik Devletleri ve Türkiye kültüründe ölçme değişmezliği göstermekte midir?

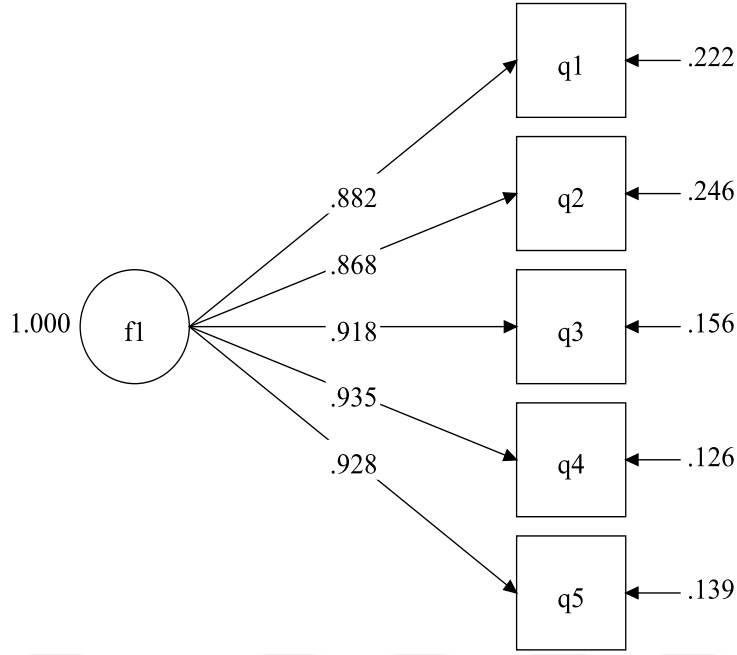
Bu bölümde PISA 2015 uygulaması fen bilimlerine yönelik öğrenci tutum anketlerine ait toplamda 37 maddeden oluşan 6 alt test, ÇGDFA yöntemine göre ölçme değişmezliği açısından ABD ve Türkiye örneklemeleri üzerinden incelenmiştir.

Tablo 7

St094 Alt Testi Uyum İndeksleri

Alt Test	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
St094	640,923	5	.993	.985	.010	.082(.075-.090)

Tablo 7 incelendiğinde uyum indekslerine ilişkin elde edilen sonuçlar CFI=.993, TLI=.985, SRMR=.010, RMSEA=.082 olarak bulunmuş ve modelle iyi uyum sağladığı görülmüştür (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu durum incelendikten sonra şekil 1’de alt testin doğrulayıcı faktör analizi için standartlaştırılmış parametre değerleri incelenmiştir.



Şekil 1. ST094 Türkiye ve ABD örnekleme DFA Standartlaştırılmış Parametre Değerleri

Şekil 1’de görüldüğü üzere ST094 (fen öğrenmekten zevk alma) alt testine ait maddelerin tek faktör üzerinde toplandığı, faktör yüklerinin .935 ve .882 arasında değiştiği, hata varyanslarının .246’yı geçmediği görülmüştür. Bu durum uyum indekslerinden elde edilen bilgilerle bir araya getirilerek model veri uyumunun iyi olduğu anlamına gelmektedir. Ölçme değişmezliği aşamalarına geçmeden önce alt testin her bir örnekleme gösterdiği uyum Tablo 8’de incelenmiştir.

Tablo 8

St094 Ülkelere Göre Model Uyum İndeksleri

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
Türkiye	259.450	5	.990	.980	.011	.097(.087-.108)
ABD	167.897	5	.993	.986	.009	.078(.068-.088)

Tablo 8 incelendiğinde alt testin örneklemini oluşturan her iki ülkenin de CFI değerleri .990 ile .993 arasında, TLI değerleri .980 ile .986 arasında, SRMR değerleri .009 ile .011 arasında, RMSEA değerleri ise .078 ile .097 arasında bulunmuştur. Bu bilgiler Türkiye ve ABD örneklemelerinin model veri uyumunun iyi olduğu anlamına gelmektedir (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu kısımdan sonra modelin ABD ve Türkiye örneklemelerinde ölçme değişmezliği aşamaları sırasıyla test edilmiş ve sonuçlar Tablo 9’da verilmiştir.

Tablo 9

St094 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA (%90 C.I)	Δ CFI
Yapısal Değişmezlik	755.602	10	.991	.983	.010	.089(.082-.096)	-
Metrik Değişmezlik	771.253	14	.989	.985	.033	.066(.063-.069)	.002
Skalar Değişmezlik	1322.316	19	.980	.979	.039	.098(.093-.103)	.009
Katı Değişmezlik	1322.316	19	.980	.979	.039	.098(.093-.103)	.000

Tablo 9 incelendiğinde ilk aşama olan yapısal değişmezlik aşamasının (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10) sağlandığı görülmüştür. Bu aşamada yapılabilecek yorum, St094 alt testi açısından ABD ve Türkiye örneklemelerinin benzer faktör yapılarına ve değerlerine sahip olduğu yönündedir. Yapısal değişmezlik aşamasının sağlanmasının ardından metrik değişmezlik aşamasına geçilir.

St094 alt testi için metrik değişmezliğin CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10 aralıklarının içinde olması ve Δ CFI<.01 sınırını aşmaması sebebiyle sağlandığı sonucuna ulaşılmıştır. Bu durumda yapılacak çıkarım faktör yükü parametrelerinin ABD ve Türkiye örneklemeleri açısından değişmez olduğu ve iki ülke örneklemelerinin teste ait maddeleri benzer şekilde yorumladığı şeklindedir. Metrik değişmezliğin sağlanmasının ardından bir sonraki aşama olan skalar değişmezliğin test edilmesi aşamasına geçilir.

Skalar değişmezlik aşamasının sağlanıp sağlanmadığının tespiti için Tablo 9 incelenmiş ve ilgili değerler kesme noktalarının üst sınırlarına yakın olsa da uygun sınırlar içerisinde bulunmuştur (CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve Δ CFI<.01). Skalar değişmezliğin sağlanmasıyla metrik aşamada test edilen faktör parametrelerinin benzerliğinin yanında regresyon sabitleri de örnekleme oluşturan ABD ve Türkiye açısından değişmez bulunmuştur. Skalar değişmezliğin sağlanmasının ardından bir sonraki ve son aşama olan katı değişmezlik test edilmiştir.

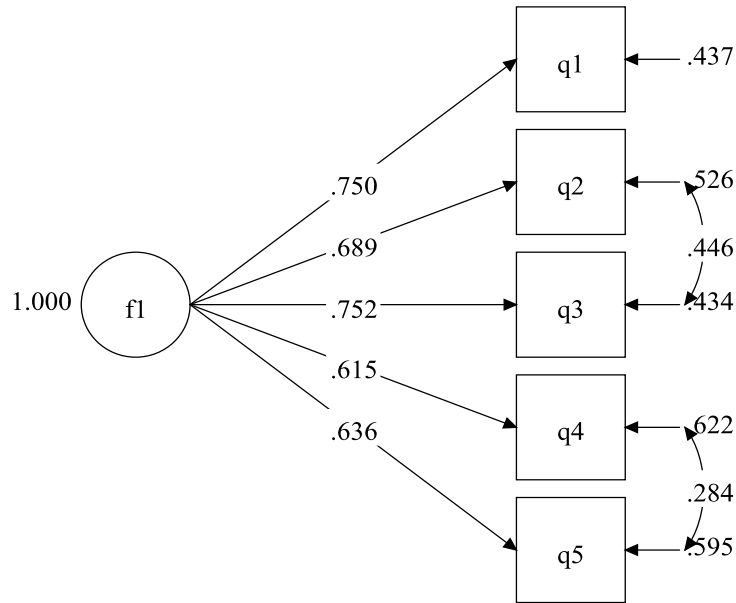
Tablo 9 da verilen parametreler incelenmiş ve tüm parametrelerin belirlenen değer aralıklarında olduğu görülmüştür (CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve Δ CFI<.01). Sonuç olarak St094 testinin diğer bütün ön kabullerle birlikte hata varyansları açısından da değişmez olup katı değişmezlik aşamasını sağladığı sonucuna ulaşılmıştır. Özetle St094 alt testi Türkiye ve ABD örneklemeleri açısından karşılaştırılabilir ve bu karşılaştırmalardan yapılacak çıkarımlar manidardır.

Tablo 10

St095 Alt Testi Uyum İndeksleri

Alt Test	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I.)
St095	148,769	3	.994	.981	.011	.060(.051-.069)

St095 testi örneklemindeki analiz sonucu incelenmiş (CFI=.921, TLI=.842, SRMR=.057, RMSEA=.173) ve Q3 ve Q2 maddeleri için hata varyanslarında uygulanacak bir kovaryans modifikasyonunun verinin uyum istatistiklerini önemli bir biçimde etkilediği görülmüş ve χ^2 değerinde 1435,507'lik önemli bir düşüş sağlamıştır. Ancak yapılan bu işlemin alt testin uyum istatistiklerini uygun aralığa getirmediği görülmüş ve Q5 ile Q4 maddelerinin hata varyansları için bir kovaryans modifikasyonu daha uygulanmıştır. Bu durumun sonucunda χ^2 değerinde 509,781'lik bir düşüş sağlanmıştır. Bu işlemlerden sonra testin uyum istatistikleri Tablo 10 üzerinden incelenmiş ve uygun bulunmuştur (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Ardından şekil 2'de alt testin doğrulayıcı faktör analizi için standartlaştırılmış parametre değerleri incelenmiştir.



Şekil 2. ST095 Türkiye ve ABD Örnekleme DFA Standartlaştırılmış Parametre Değerleri

Şekil 2'de görüldüğü üzere ST095 (geniş fen konularına ilgi) alt testine ait maddelerin tek faktör üzerinde toplandığı, faktör yüklerinin .752 ve .615 arasında değiştiği, hata varyanslarının .622'yı geçmediği görülmüştür. Bu durum uyum indekslerinden elde edilen bilgilerle bir araya getirilerek model ile veri uyumunun iyi

olduğu anlamına gelmektedir. Ölçme değişmezliği aşamalarına geçmeden önce alt testin her bir örnekleme gösterdiği uyum Tablo 11’de incelenmiştir.

Tablo 11

St095 Ülkelere Göre Model Uyum İndeksleri

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
Türkiye	65,481	3	.996	.987	.010	.052(.040-.066)
ABD	91,838	3	.992	.973	.012	.069(.056-.082)

Tablo 11 incelendiğinde alt testin örneklemini oluşturan her iki ülkenin de CFI değerleri .996 ile .992 arasında, TLI değerleri .987 ile .973 arasında, SRMR değerleri .010 ile .012 arasında, RMSEA değerleri ise .052 ile .069 arasında bulunmuştur. Bu bilgiler Türkiye ve ABD örneklemelerinin model veri uyumunun iyi olduğu anlamına gelmektedir (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu kısımdan sonra modelin ABD ve Türkiye örneklemelerinde ölçme değişmezliği aşamaları sırasıyla test edilmiş ve sonuçlar Tablo 12’de verilmiştir.

Tablo 12

St095 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)	ΔCFI
Yapısal Değişmezlik	157,313	6	.994	.981	.011	.060(.051-.070)	-
Metrik Değişmezlik	212,976	10	.991	.983	.024	.057(.050-.064)	.003
Skalar Değişmezlik	622,865	15	.973	.964	.056	.083(.077-.089)	.018

Tablo 12 incelendiğinde St095 alt testinin yapısal değişmezlik aşamasının (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10) sağlandığı görülmüştür. ABD ve Türkiye örneklemelerinin St095 alt testi açısından benzer faktör yapılarına ve değerlerine sahip olduğu görülmektedir. Yapısal değişmezliğin sağlanmasının ardından metrik değişmezlik aşamasına geçilmiştir.

St095 alt testi için metrik değişmezliğin CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve ΔCFI <.01 değerlerini geçmemesi sebebiyle sağlandığı sonucuna ulaşılmıştır. Faktör yükü parametrelerinin ABD ve Türkiye örneklemeleri açısından değişmez olduğu anlaşılmıştır. Metrik değişmezliğin sağlanmasının ardından bir sonraki aşama olan skalar değişmezlik aşamasına geçilmiştir.

Skalar değişmezliğin sağlanıp sağlanmadığını incelemek için Tablo 12 incelenmiş ve elde edilen değerlerin kabul sınırlarının dışında olduğu görülmüştür (ΔCFI <.01). Bu

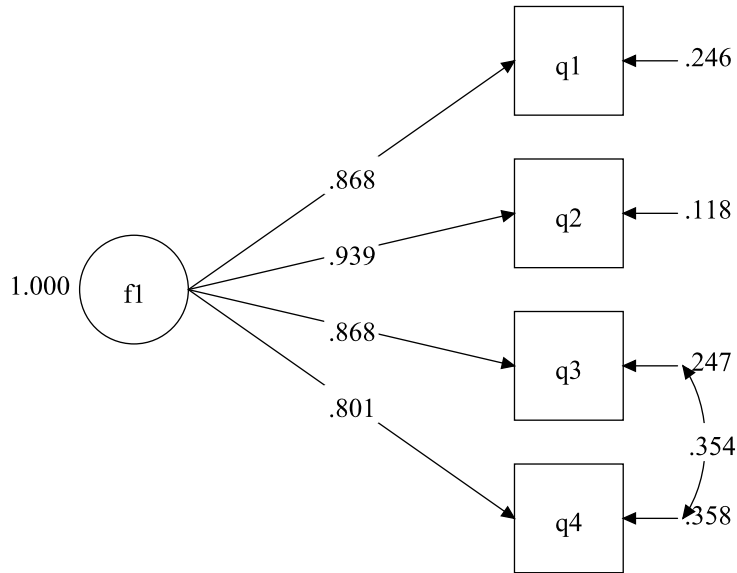
durumda regresyon sabitleri Türkiye ve ABD örneklemleri arasında değişmez değildir. Yapılabilecek bir başka yorum ise St095 testinin maddelerinin madde yanlılığı gösterebileceğidir. Katı değişmezliğin test edilebilmesi için skalar değişmezliğin sağlanabilmesi ön koşulu bulunmaktadır. Skalar değişmezliğin sağlanamamasının ardından analizler St095 alt testi açısından durdurulmuştur.

Tablo 13

St113 Alt Testi Uyum İndeksleri

Alt Test	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
St113	7.250	1	1.000	.999	.001	.016(.001-.035)

St113 testi model uyum indeksleri (CFI=.986, TLI=.958, SRMR=.018, RMSEA=.132) incelenmiş ve Q4 ve Q3 maddeleri arasında hata varyansları için uygulanacak bir kovaryans modifikasyonunun verinin χ^2 değerinde 807,265'lik manidar bir düşüşe sebep olacağı görülmüş ve bu durumda modifikasyona gidilmiştir. Bu işlemlerden sonra testin uyum istatistikleri Tablo 13 üzerinden incelenmiş ve uygun bulunmuştur (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Ardından şekil 3'te alt testin doğrulayıcı faktör analizi için standartlaştırılmış parametre değerleri incelenmiştir.



Şekil 3. St113 Türkiye ve ABD örneklemleri DFA Standartlaştırılmış Parametre Değerleri

Şekil 3'te görüldüğü üzere ST113 (fen öğreniminde araçsal güdülenme) alt testine ait maddelerin tek faktör üzerinde toplandığı, faktör yüklerinin .939 ile .801 arasında değiştiği, hata varyanslarının .358'i geçmediği görülmüştür. Bu durum uyum indekslerinden elde edilen bilgilerle bir araya getirilerek model veri uyumunun iyi olduğu anlamına gelmektedir. Ölçme değişmezliği aşamalarına geçmeden önce alt testin her bir örnekleme gösterdiği uyum Tablo 14'te incelenmiştir.

Tablo 14

St113 Ülkelere Göre Model Uyum İndeksleri

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
Türkiye	7.390	1	1.000	.999	.002	.022(.000-.048)
ABD	1.254	1	1.000	1.000	.001	.000(.000-.034)

Tablo 14 incelenmiş ve St113 testinin örneklemini oluşturan her iki ülkenin de CFI değerleri 1.000, TLI değerleri .999 ve 1.000, SRMR değerleri .002 ile .001 arasında, RMSEA değerleri ise .022 ile .000 olarak bulunmuştur. Bu bilgiler Türkiye ve ABD örneklemelerinin model veri uyumunun iyi olduğu anlamına gelmektedir (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu kısımdan sonra modelin ABD ve Türkiye örneklemelerinde ölçme değişmezliği aşamaları sırasıyla test edilmiş ve sonuçlar Tablo 15'te verilmiştir.

Tablo 15

St113 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)	ΔCFI
Yapısal Değişmezlik	8.644	2	1.000	.999	.002	.016(.000-.035)	-
Metrik Değişmezlik	36.495	5	.999	.998	.013	.030(.020-.041)	.001
Skalar Değişmezlik	124.263	9	.996	.995	.021	.045(.038-.053)	.003
Katı Değişmezlik	124,263	9	.996	.995	.021	.045(.038-.053)	.000

Tablo 15 incelendiğinde yapısal değişmezlik aşaması için elde edilen parametre değerlerinin uygun kesme noktalarının aralığında bulunduğu (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10) görülmüştür. İlgili alt gruplar açısından faktör yapılarının ve örüntülerinin benzer olduğu görülmüş ve St113 testinin ABD ve Türkiye örneklemelerindeki yapısal değişmezliği aşaması kabul edilmiştir.

Metrik değişmezlik aşamasında ise yapılan analizin incelenmesi sonucu parametre değerlerinin uygun aralıktan sapmadığı görülmüştür (CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve ΔCFI <.01). Sonuç olarak St113 bağlamında Türkiye ve ABD

örneklemelerinin faktör yükleri değişmezdir. Metrik değişmezliğin sağlanmasının ardından skalar değişmezlik aşaması analiz edilmiştir.

Skalar değişmezliğin incelenmesi için analizden elde edilen parametreler kesme noktalarıyla karşılaştırılmış (CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve Δ CFI<.01) ve St113 testinin skalar değişmezliği kabul sınırları içerisinde yer aldığı için doğrulanmıştır. St113 testinin regresyon sabitleri örneklemin alt grupları açısından değişmezdir. Skalar değişmezliğin sağlanmasının ardından katı değişmezlik aşamasının analizine geçilmiştir.

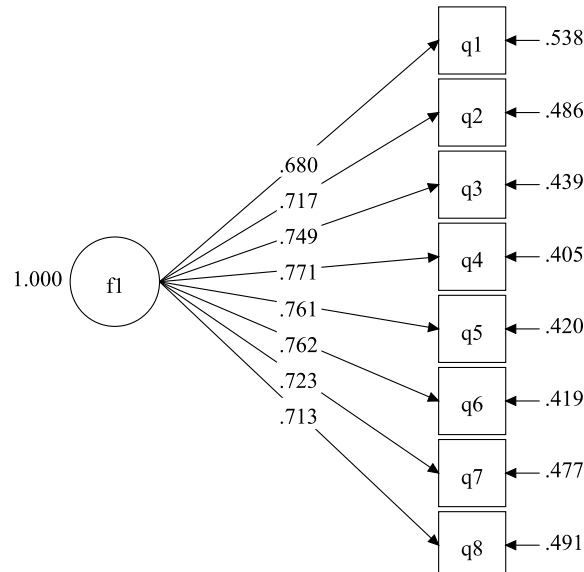
St113 alt testinin katı değişmezlik açısından analizi sonucunda ortaya çıkan parametreler uyum kabul değerlerinin içerisinde yer almaktadır (CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve Δ CFI<.01). Bu durum sonucunda St113 alt testinin ABD ve Türkiye alt grupları için tüm boyutlarda yapılacak karşılaştırmaların sonuçları manidar olacaktır.

Tablo 16

St129 Alt Testi Uyum İndeksleri

Alt Test	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
St129	1691.172	20	.966	.953	.029	.072(.068-.075)

Tablo 16 incelendiğinde uyum indekslerine ilişkin elde edilen sonuçlar CFI=.966, TLI=.953, SRMR=.029, RMSEA=.072 olarak bulunmuş ve modelle iyi uyum sağladığı görülmüştür (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu durum incelendikten sonra şekil 4'te alt testin doğrulayıcı faktör analizi için standartlaştırılmış parametre değerleri incelenmiştir.



Şekil 4. St129 Türkiye ve ABD örneklemleri DFA Standartlaştırılmış Parametre Değerleri

Şekil 4 incelenmiş ve St129 (fene yönelik öz-yeterlik) alt testine ait maddelerin tek faktör üzerinde toplandığı, faktör yüklerinin .680 ile .771 arasında değiştiği, hata varyanslarının .538'i geçmediği görülmüştür. Bu durum uyum indekslerinden elde edilen bilgilerle bir araya getirilerek model veri uyumunun iyi olduğu anlamına gelmektedir. Ölçme değişmezliği aşamalarına geçmeden önce alt testin her bir örnekleme gösterdiği uyum Tablo 17'de incelenmiştir.

Tablo 17

St129 Ülkelere Göre Model Uyum İndeksleri

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
Türkiye	809.384	19	.970	.955	.028	.070(.065-.075)
ABD	687.012	20	.973	.963	.025	.066(.061-.071)

St129 testi Türkiye örneklemindeki analiz sonucu incelenmiş (CFI=.953, TLI=.934, SRMR=.036, RMSEA=.085) ve Q2 ile Q1 maddeleri için hata varyanslarında tanımlanacak bir kovaryans modifikasyonunun verinin uyum istatistiklerini önemli bir biçimde etkilediği görülmüş ve χ^2 değerinde 443,897'lik bir düşüş sağlamıştır. Daha sonrasında St129 alt testinin genelinde ve alt gruplarında uyum indeksleri incelendikten sonra uygun aralıkta bulunmuş (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10) ve iki kültür arasındaki ölçme değişmezliği aşamaları sırasıyla test edilmiştir.

Tablo 18

St129 ABD ve Türkiye Örneklemlerinde Ölçme Değişmezliği Aşamaları

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)	ΔCFI
Yapısal Değişmezlik	1496.408	39	.971	.959	.027	.068(.064-.072)	-
Metrik Değişmezlik	1560.106	46	.968	.961	.033	.066(.063-.069)	.003
Skalar Değişmezlik	2643.4399	54	.943	.941	.047	.081(.078-.085)	.025

Yapısal değişmezlik aşaması St129 alt testi için Tablo 18'de incelenmiş ve ilgili değişmezlik aşamasının gerektirdiği kesme noktalarının içerisinde yer aldığı görülmüştür (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu durumda St129 alt testinin yapısal değişmezliğinin karşılandığı ve ABD ile Türkiye örnekleminin benzer faktör örüntüleri gösterdiği sonucuna varılmıştır. Yapısal değişmezliğin sağlanmasının ardından metrik değişmezlik aşamasının analizine geçilmiştir.

Tablo 18 incelenmiş ve St129 alt testinin parametreleri metrik değişmezliğin sağlanıp sağlanamaması açısından test edilmiştir. İlgili parametrelerden elde edilen sonuçlarının metrik değişmezliğin kabul sınırları içerisinde yer aldığı görülmüştür

(CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Sonuç olarak St129 alt testi Türkiye ve ABD örneklemeleri açısından faktör yükleri boyutuyla değişmez bulunmuştur.

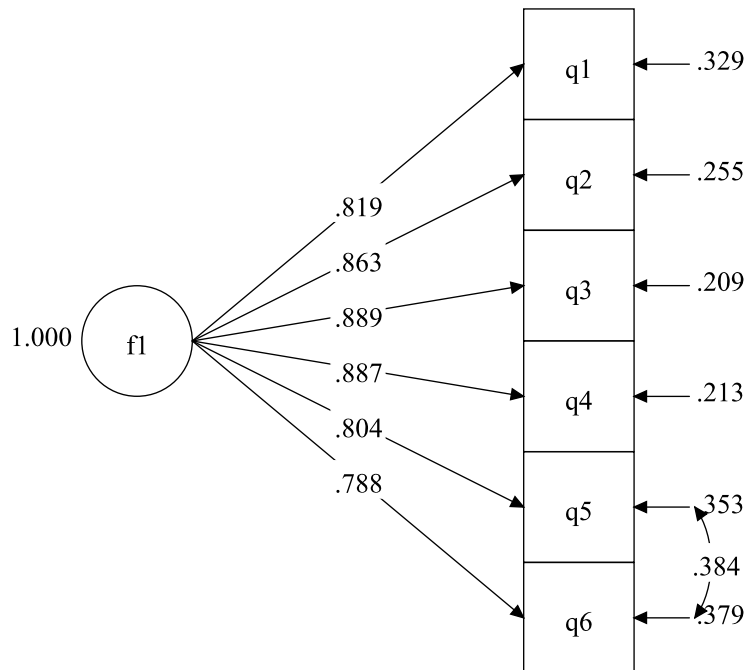
St129 alt testi skalar değişmezlik aşaması için test edilmiş ve elde edilen parametreler, kabul aralıklarının dışında bulunmuştur (CFI>.95, TLI>.95, SRMR<.08 RMSEA<.10 ve $\Delta CFI < .01$). Bu durumda yapılabilecek yorum ABD ve Türkiye örnekleminin regresyon sabitleri açısından değişmezliğinin reddedilmesi yönündedir. Skalar değişmezlik sağlanmadığı için katı değişmezlik aşamasının analizi durdurulmuştur.

Tablo 19

St131 Alt Testi Uyum İndeksleri

Alt Test	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
St131	726,727	8	.989	.979	.017	.072(.067-.078)

St131 testi örneklemindeki analiz sonucu incelenmiş (CFI=.968, TLI=.947, SRMR=.026, RMSEA=.114) ve Q6 ile Q5 maddeleri için hata varyanslarında uygulanacak bir kovaryans modifikasyonunun verinin uyum istatistiklerini önemli ölçüde etkilediği görülmüş ve χ^2 değerinde 1319,815'lik önemli bir düşüş sağlamıştır. Bu işlemin sonucunda testin uyum istatistikleri Tablo 19'da verildiği şekliyle uygun bulunmuştur (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu işlemten sonra şekil 5'te St131 testin doğrulayıcı faktör analizi için standartlaştırılmış parametre değerleri incelenmiştir.



Şekil 5. St131 Türkiye ve ABD Örneklemini DFA Standartlaştırılmış Parametre Değerleri

Şekil 5 incelenmiş ve St131 (fen bilimlerine yönelik epistemik inançlar) alt testine ait maddelerin tek faktör üzerinde toplandığı, faktör yüklerinin .788 ile .889 arasında değiştiği, hata varyanslarının .384'ü geçmediği görülmüştür. Veriden elde edilen parametrelerle birlikte incelendiğinde verinin analize uygun olduğu görülmüştür. Ölçme değişmezliği aşamalarına geçmeden önce alt testin her bir örnekleme gösterdiği uyum Tablo 20'de incelenmiştir.

Tablo 20

St131 Ülkelere Göre Model Uyum İndeksleri

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
Türkiye	342,243	8	.990	.981	.016	.072(.064-.080)
ABD	517,338	8	.982	.966	.021	.084(.076-.092)

Tablo 20'de verilen değerler CFI için .990 ve .982, TLI için .981 ve .966, SRMR için .016 ve .021, RMSEA için .072 ve .084 olarak bulunmuştur. Bu durumda verinin uyum indeksleri iyi uyum için belirlenen aralıktadır ($CFI > .95$, $TLI > .95$, $SRMR < .08$ ve $RMSEA < .10$). Bu işlemin ardından St131 testinin ABD ve Türkiye verileri arasındaki ölçme değişmezliği aşamaları test edilmiştir ve Tablo 21'de ilgili sonuçlar verilmiştir.

Tablo 21

St131 ABD ve Türkiye Örneklemelerinde Ölçme Değişmezliği Aşamaları

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)	ΔCFI
Yapısal Değişmezlik	859,581	16	.986	.974	.019	.079(.073-.084)	-
Metrik Değişmezlik	1007,855	21	.982	.974	.036	.079(.074-.084)	.004
Skalar Değişmezlik	1644,257	27	.968	.964	.098	.092(.088-.097)	.014

Tablo 21 incelenmiş ve yapısal değişmezlik için elde edilen parametrelerin veri ile iyi uyum gösterdiği görülmüştür ($CFI > .95$, $TLI > .95$, $SRMR < .08$ ve $RMSEA < .10$). Bu açıdan bakılarak ABD ve Türkiye örneklemelerinin faktör örüntülerinin benzer olduğu belirtilebilir. Dolayısıyla St131 testinin yapısal değişmezliği ABD ve Türkiye örneklemeleri açısından sağlanmıştır. Bu işlemin akabinde testin metrik değişmezlik aşamasının analizine geçilmiştir.

Metrik değişmezlik aşaması kabul noktaları açısından incelenmiş ve testin yapısal değişmezliği kabul edilmiştir ($CFI > .95$, $TLI > .95$, $SRMR < .08$, $RMSEA < .10$ ve $\Delta CFI < .01$). Sonuç olarak St131 testi metrik değişmezliği sağlayarak ABD ve Türkiye örneklemeleri açısından faktör yükleri boyutuyla değişmez olarak tespit edilmiştir. Metrik değişmezliğin sağlanmasının ardından bir sonraki aşama olan skalar değişmezlik test edilmiştir.

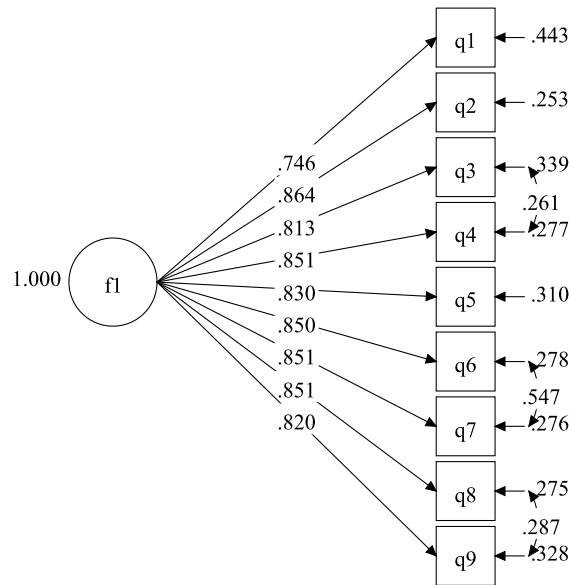
Tablo 21 incelendiğinde St131 testinin ABD ve Türkiye örneklemeleri açısından skalar değişmezliğinin sağlanamadığı görülmüştür (CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve Δ CFI<.01). Sonuç olarak skalar değişmezliğin sağlanamamasıyla iki alt grubun regresyon sabitlerinin değişmezliği reddedilmiştir. Bu nedenle bir sonraki değişmezlik boyutunun (katı değişmezlik) analizleri skalar değişmezlik sağlanmadığı için gerçekleştirilmemiştir.

Tablo 22

St146 Alt Testi Uyum İndeksleri

Alt Test	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
St146	3234,395	25	.967	.952	.029	.088(.085-.091)

St146 testi örneklemindeki analiz sonucu incelenmiş (CFI=.929, TLI=.906, SRMR=.036, RMSEA=.124) ve Q7 ile Q6 maddeleri için hata varyanslarında uygulanacak bir kovaryans modifikasyonu verinin uyum istatistiklerini önemli bir biçimde etkilediği görülmüş ve χ^2 değerinde 3028,028’lik önemli bir düşüş sağlamıştır. Ancak yapılan bu işlemin alt testin uyum istatistiklerini uygun aralığa getirmediği görülmüş ve Q9 ile Q8 maddelerinin hata varyansları için bir kovaryans modifikasyonu daha uygulanmıştır ve bu durumun sonucunda χ^2 değerinde 783,297’lik bir düşüş sağlanmıştır. Bu işlemlerin sonucunda Tablo 22’de görülen değer ortaya çıkmış ve verinin modelle uyumunun iyi olduğuna karar verilmiştir (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu durum incelendikten sonra şekil 6’da St146 testinin doğrulayıcı faktör analizi için standartlaştırılmış parametre değerleri incelenmiştir.



Şekil 6. ST146 Türkiye ve ABD örneklemleri DFA Standartlaştırılmış Parametre Değerleri

Şekil 6 incelenmiş ve St146 (bilim aktiviteleri) alt testine ait maddelerin tek faktör üzerinde toplandığı, faktör yüklerinin .746 ile .864 arasında değiştiği, hata varyanslarının .547'yi geçmediği görülmüştür. Veriden elde edilen parametreler incelendiğinde verinin analize uygun olduğu görülmüştür. Ölçme değişmezliği aşamalarına geçmeden önce alt testin her bir örnekleme gösterdiği uyum Tablo 23'te incelenmiştir.

Tablo 23

St146 Ülkelere Göre Model Uyum İndeksleri

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)
Türkiye	1717.166	24	.967	.950	.029	.093(.088-.097)
ABD	1233.825	25	.969	.956	.030	.077(.072-.081)

St146 testi Türkiye örneklemindeki analiz sonucu incelenmiş (CFI=.958, TLI=.939, SRMR=.033, RMSEA=.103) ve Q2 ile Q1 maddeleri için hata varyanslarında uygulanacak bir kovaryans modifikasyonu verinin uyum istatistiklerini önemli bir biçimde etkilediği görülmüş ve χ^2 değerinde 473,324'lük bir düşüş sağlanmıştır. Daha sonra St146 alt testinin genelinde ve alt gruplarında uyum indeksleri Tablo 23'te verilen şekliyle incelendikten sonra uygun bulunmuş ve iki kültür arasındaki ölçme değişmezliği aşamaları sırasıyla test edilmiştir (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). İlgili sonuçlar Tablo 24'te verilmiştir.

Tablo 24

St146 ABD ve Türkiye Örneklemlerinde Ölçme Değişmezliği Aşamaları

Ülke	χ^2 (SB Düzeltmeli)	sd	CFI	TLI	SRMR	RMSEA(%90 C.I)	ΔCFI
Yapısal Değişmezlik	2950,974	49	.968	.952	.030	.085(.082-.088)	-
Metrik Değişmezlik	3273,599	57	.962	.952	.048	.085(.082-.088)	.006
Skalar Değişmezlik	4909,235	66	.940	.935	.142	.100(.097-.102)	.022

Tablo 24 incelenmiş ve yapısal değişmezlik için elde edilen parametrelerin veri ile iyi uyum gösterdiği görülmüştür (CFI>.95, TLI>.95, SRMR<.08 ve RMSEA<.10). Bu açıdan bakılarak ABD ve Türkiye örneklemlerinin faktör örüntülerinin benzer olduğu belirtilebilir. Dolayısıyla ile St146 testinin yapısal değişmezliği ABD ve Türkiye örneklemleri açısından sağlanmıştır. Bu işlemin akabinde testin metrik değişmezlik aşamasının analizine geçilmiştir.

Metrik değişmezlik aşaması kabul noktaları açısından incelenmiş ve elde edilen parametreler uygun bulunarak testin metrik değişmezliği kabul edilmiştir (CFI>.95, TLI>.95, SRMR<.08, RMSEA<.10 ve ΔCFI <.01). Sonuç olarak St146 testi ve Türkiye

örneklemeleri açısından faktör yükleri boyutuyla değişmez olarak tespit edilmiştir. Metrik değişmezliğin sağlanmasının ardından bir sonraki aşama olan skalar değişmezlik aşaması test edilmiştir.

Tablo 21 incelendiğinde St146 testinin ABD ve Türkiye örneklemeleri açısından skalar değişmezliğinin sağlanamadığı görülmüştür ($CFI > .95$, $TLI > .95$, $SRMR < .08$, $RMSEA < .10$ ve $\Delta CFI < .01$). Sonuç olarak skalar değişmezliğin sağlanamamasıyla iki alt grubun regresyon sabitlerinin değişmezliği reddedilmiştir. Bu nedenle bir sonraki değişmezlik boyutunun (katı değişmezlik) analizleri skalar değişmezlik sağlanamadığı için gerçekleştirilmemiştir.

4.2. İkinci Alt Probleme Yönelik Bulgu ve Yorumlar

Bu bölümde PISA 2015 uygulaması fen bilimlerine yönelik öğrenci tutum anketlerine ait 6 alt testte bulunan 37 madde MTK-OO yöntemine göre DMF açısından cinsiyet ve kültüre göre incelenmiştir. Maddelerin çok kategorili yapısı gereği aşamalı tepki modeli (graded response model) kullanılmıştır. ST094 cinsiyete göre Türkiye örneklemine DMF analizi çıktısı örnek teşkil etmesi açısından EK 4’te verilmiştir.

MTK-OO yönteminde karşılanması gereken varsayımlar yerel bağımsızlık ve tek boyutluluktur. Tek boyutluluk analizi DFA ile incelenmiş ve her bir alt testin tek boyutlu olduğu bulunmuştur. Yerel bağımsızlık varsayımı ise maddelerin yanıtlanma olasılıklarının birbirinden bağımsız olması durumudur. Bu durum tek boyutluluk analizi ile de alakalıdır. Maddelerin sıralaması için de bu inceleme yapılmış ve birbirinin yanıtlanmasını etkileyen madde bulunmamıştır. Tüm bu durumlar bir araya gelerek alt testlerin MTK-OO varsayımlarını karşıladığı sonucuna ulaşılmıştır.

4.3. Problem 2a’ya Yönelik Bulgu ve Yorumlar

Alt Problem 2a: PISA 2015 Türkiye uygulamasındaki öğrenci anketlerinde yer alan maddeler cinsiyete dayalı olarak MTK-OO yöntemine göre değişen madde fonksiyonu göstermekte midir?

PISA 2015 uygulaması fen bilimlerine yönelik öğrenci tutum anketi verileri Türkiye örneğinde MTK-OO yöntemine göre cinsiyete dayalı DMF içeren madde analizi sonuçları Tablo 25’te verilmiştir.

Tablo 25

Cinsiyete Göre Türkiye Örnekleme MTK-OO Analiz Sonuçları

Alt Boyutlar	Maddeler	G ²	DMF	Düzye	DMF Türü	Avantajlı Grup
Fen Öğrenmekten Zevk Alma (st094)	Q01	34.5	Var	B	TB	Erkekler
	Q02	2.5	Yok	-		
	Q03	29.8	Var	B	TB	Kadınlar
	Q04	25.3	Var	B	TBO	Erkekler
	Q05	12.5	Var	B	TBO	Kadınlar
Geniş Fen Konularına İlgi (st095)	Q01	17.9	Var	B	TB	Erkekler
	Q02	112.3	Var	C	TB	Kadınlar
	Q03	13.1	Var	B	TB	Kadınlar
	Q04	11.5	Var	B	TB	Kadınlar
	Q05	161.2	Var	C	TB	Erkekler
Fen Öğreniminde Araçsal Güdülenme (st113)	Q01	28.9	Var	B	TB	Kadınlar
	Q02	11.0	Var	B	TB	Erkekler
	Q03	6.6	Var	A	TBO	Kadınlar
	Q04	10.7	Var	B	TB	Kadınlar
Fene Yönelik Öz Yeterlik (st129)	Q01	14.6	Var	B	TBO	Erkekler
	Q02	33.5	Var	B	TBO	Erkekler
	Q03	33.9	Var	B	TBO	Erkekler
	Q04	11.6	Var	B	TBO	Erkekler
	Q05	20.9	Var	B	TB	Erkekler
	Q06	9.6	Var	B	TB	Kadınlar
	Q07	83.3	Var	C	TB	Kadınlar
	Q08	12.9	Var	B	TB	Erkekler
Fen Bilimlerine Yönelik Epistemik İnançlar (st131)	Q01	10.8	Var	B	TB	Erkekler
	Q02	14.8	Var	B	TBO	Kadınlar
	Q03	19.2	Var	B	TB	Erkekler
	Q04	14.4	Var	B	TB	Erkekler
	Q05	15.0	Var	B	TB	Kadınlar
	Q06	16.9	Var	B	TB	Kadınlar
Bilim Aktiviteleri (st146)	Q01	8.8	Var	A	TB	Kadınlar
	Q02	40.5	Var	B	TB	Erkekler
	Q03	22.5	Var	B	TBO	Erkekler
	Q04	32.6	Var	B	TB	Erkekler
	Q05	8.9	Var	A	TB	Kadınlar
	Q06	6.5	Var	A	TB	Kadınlar
	Q07	14.7	Var	B	TB	Kadınlar
	Q08	7.2	Var	A	TB	Kadınlar
	Q09	5.9	Var	A	TB	Kadınlar

*Referans Grup: Kadınlar *Odak Grup: Erkekler

Tablo 25 incelendiğinde maddelerin G² değerlerine göre çıkarımlar yapılmış ve 6 maddede ihmal edilebilir düzeyde (A), 27 maddede orta düzeyde (B), 3 maddede ise yüksek düzeyde (C) DMF içeren madde bulunmuştur. İhmal edilebilir düzeyde olan maddeler çıkarıldığında 30 maddede DMF bulunmuş ve DMF'li maddelerin testteki oranı %81.08 olarak belirlenmiştir. DMF belirlenen maddelerin 22'si tek biçimli 8'i ise tek biçimli olmayan şekilde DMF göstermiştir ve bu maddelerin 17'si erkekler 13'ü kadınlar lehine avantaj sağlamaktadır.

4.4. Problem 2b'ye Yönelik Bulgu ve Yorumlar

Alt Problem 2b: PISA 2015 Amerika Birleşik Devletleri Uygulamasındaki öğrenci anketlerinde yer alan maddeler cinsiyete dayalı olarak MTK-OO yöntemine göre değişen madde fonksiyonu içermekte midir?

PISA 2015 uygulaması fen bilimlerine yönelik öğrenci tutum anketi verileri ABD örnekleminde MTK-OO yöntemine göre cinsiyete dayalı DMF içeren madde analizi sonuçları Tablo 26'da verilmiştir.

Tablo 26

Cinsiyete göre ABD Örneklemini MTK-OO Analiz Sonuçları

Alt Boyutlar	Maddeler	G ²	DMF	Düzye	DMF Türü	Avantajlı Grup
Fen Öğrenmekten Zevk Alma (st094)	Q01	10.5	Var	B	TB	Erkekler
	Q02	18.5	Var	B	TB	Kadınlar
	Q03	4.9	Var	A	TB	Erkekler
	Q04	14.4	Var	B	TBO	Kadınlar
	Q05	5.3	Var	A	TB	Erkekler
Geniş Fen Konularına İlgi (st095)	Q01	31.9	Var	B	TB	Erkekler
	Q02	128.2	Var	C	TB	Kadınlar
	Q03	7.8	Var	A	TB	Kadınlar
	Q04	6.0	Var	A	TB	Kadınlar
	Q05	330.8	Var	C	TB	Erkekler
Fen Öğreniminde Araçsal Güdülenme (st113)	Q01	4.3	Var	A	TB	Erkekler
	Q02	0.3	Yok	-		
	Q03	8.5	Var	A	TB	Kadınlar
	Q04	3.5	Yok	-		
	Q01	18.4	Var	B	TB	Erkekler
Fene Yönelik Öz Yeterlik (st129)	Q02	32.5	Var	B	TBO	Erkekler
	Q03	79.5	Var	C	TBO	Erkekler
	Q04	14.4	Var	B	TB	Kadınlar
	Q05	30.0	Var	B	TB	Erkekler
	Q06	40.5	Var	B	TB	Erkekler
	Q07	60.0	Var	C	TB	Kadınlar
	Q08	60.3	Var	C	TB	Kadınlar
	Q01	14.0	Var	B	TB	Erkekler
Fen Bilimlerine Yönelik Epistemik İnançlar (st131)	Q02	2.0	Yok	-		
	Q03	1.8	Yok	-		
	Q04	17.4	Var	B	TB	Erkekler
	Q05	8.2	Var	A	TB	Kadınlar
	Q06	2.0	Yok	-		
	Q01	51.1	Var	C	TB	Kadınlar
Bilim Aktiviteleri (st146)	Q02	11.4	Var	B	TB	Erkekler
	Q03	6.3	Var	A	TB	Erkekler
	Q04	4.6	Var	A	TB	Erkekler
	Q05	5.8	Var	A	TB	Kadınlar
	Q06	9.2	Var	A	TB	Kadınlar
	Q07	15.1	Var	B	TB	Kadınlar
	Q08	41.1	Var	B	TBO	Erkekler
	Q09	3.7	Yok	-		

*Referans Grup: Kadınlar *Odak Grup: Erkekler

Tablo 26 incelendiğinde maddelerin G^2 değerlerine göre çıkarımlar yapılmış ve 11 maddede ihmal edilebilir düzeyde (A), 14 maddede orta düzeyde (B), 6 maddede ise yüksek düzeyde (C) DMF içeren madde bulunmuştur. İhmal edilebilir düzeyde olan maddeler çıkarıldığında 20 maddede DMF bulunmuş ve DMF'li maddelerin testteki oranı %54.05 olarak belirlenmiştir. DMF belirlenen maddelerin 16'sı tek biçimli 4'ü ise tek biçimli olmayan şekilde DMF göstermiştir ve bu maddelerin 12'si erkekler 8'i kadınlar lehine çalışır şekildedir.

4.5. Problem 2c'ye Yönelik Bulgu ve Yorumlar

Alt Problem 2c: PISA 2015 uygulamasında yer alan öğrenci anketlerindeki maddeler Türkiye ve Amerika Birleşik Devletleri arasında kültüre dayalı olarak MTK-OO yöntemine göre değişen madde fonksiyonu içermekte midir?

PISA 2015 uygulaması fen bilimlerine yönelik öğrenci tutum anketi verileri ABD ve Türkiye örnekleminde MTK-OO yöntemine göre kültüre dayalı DMF içeren madde analizi sonuçları Tablo 27'de verilmiştir.

Tablo 27

Kültüre Göre Türkiye ve ABD Örneklemini MTK-OO Analiz Sonuçları

Alt Boyutlar	Maddeler	G^2	DMF	Düzye	DMF Türü	Avantajlı Grup
Fen Öğrenmekten Zevk Alma (st094)	Q01	292.3	Var	C	TBO	Türkiye
	Q02	385.1	Var	C	TB	ABD
	Q03	129.7	Var	C	TBO	Türkiye
	Q04	56.4	Var	C	TBO	Türkiye
	Q05	75.2	Var	C	TB	ABD
Geniş Fen Konularına İlgi (st095)	Q01	5.2	Var	A	TBO	ABD
	Q02	174.3	Var	C	TB	ABD
	Q03	72.4	Var	C	TB	ABD
	Q04	168.9	Var	C	TB	Türkiye
	Q05	256.6	Var	C	TBO	Türkiye
Fen Öğreniminde Araçsal Güdülenme (st113)	Q01	44.7	Var	C	TBO	Türkiye
	Q02	154.7	Var	C	TBO	ABD
	Q03	135.5	Var	C	TBO	Türkiye
	Q04	144.5	Var	C	TBO	Türkiye
	Q01	114.0	Var	C	TBO	Türkiye
Fene Yönelik Öz Yeterlik (st129)	Q02	164.9	Var	C	TB	Türkiye
	Q03	63.0	Var	C	TB	Türkiye
	Q04	321.1	Var	C	TB	ABD
	Q05	320.2	Var	C	TB	Türkiye
	Q06	54.3	Var	C	TB	Türkiye
	Q07	148.7	Var	C	TBO	Türkiye
	Q08	631.9	Var	C	TB	ABD

(Devam ediyor)

Tablo 27 (Devamı)

Kültüre Göre Türkiye ve ABD Örnekleme MTK-OO Analiz Sonuçları

Alt Boyutlar	Maddeler	G ²	DMF	Düzye	DMF Türü	Avantajlı Grup
Fen Bilimlerine Yönelik Epistemik İnançlar (st131)	Q01	411.9	Var	C	TB	Türkiye
	Q02	391.9	Var	C	TBO	Türkiye
	Q03	52.6	Var	C	TBO	ABD
	Q04	186.4	Var	C	TBO	ABD
	Q05	45.7	Var	C	TB	ABD
	Q06	163.7	Var	C	TBO	ABD
Bilim Aktiviteleri (st146)	Q01	164.7	Var	C	TBO	Türkiye
	Q02	147.3	Var	C	TB	ABD
	Q03	54.8	Var	C	TB	Türkiye
	Q04	52.7	Var	C	TB	Türkiye
	Q05	582.1	Var	C	TB	ABD
	Q06	21.2	Var	B	TB	ABD
	Q07	37.9	Var	B	TBO	Türkiye
	Q08	56.1	Var	C	TBO	Türkiye
	Q09	299.7	Var	C	TBO	Türkiye

*Referans Grup: Türkiye *Odak Grup: ABD

Tablo 27 incelendiğinde maddelerin G² değerlerine göre çıkarımlar yapılmış ve 1 maddede ihmal edilebilir düzeyde (A), 2 maddede orta düzeyde (B), 34 maddede ise yüksek düzeyde (C) DMF içeren madde bulunmuştur. İhmal edilebilir düzeyde olan maddeler çıkarıldığında 36 maddede DMF bulunmuş ve DMF'li maddelerin testteki oranı %97.29 olarak belirlenmiştir. DMF belirlenen maddelerin 18'i tek biçimli 18'i ise tek biçimli olmayan şekilde DMF göstermiştir ve bu maddelerin 22'si Türkiyeli öğrencilerin, 14'ü ABD'li öğrencilerin lehine çalışır şekildedir.

BÖLÜM 5

SONUÇLAR VE ÖNERİLER

Bu bölümde araştırmadan edinilen sonuç ve önerilere yer verilmiştir. Sonuçlar ölçme değişmezliği ve değişen madde fonksiyonu açısından iki başlık altında incelenmiştir. Elde edilen sonuçlardan yola çıkarak gelecekte yapılması planlanan uygulamalara ve araştırmalara yönelik önerilere yer verilmiştir.

5.1. Sonuçlar

Bu bölümde değişen madde fonksiyonu ve ölçme değişmezliğine ilişkin sonuçlar iki başlık altında irdelenmiştir.

5.1.1. Değişen Madde Fonksiyonuna Yönelik Sonuçlar

2015 PISA uygulaması fen bilimlerine yönelik öğrenci tutum anketleri 6 testten oluşmuş ve değişen madde fonksiyonu yönünden Türk ve ABD örneklemi için kültür açısından, Türk ve ABD örnekleme için cinsiyet açısından incelenmiştir. Sonuç olarak cinsiyet açısından Türkiye örnekleme için 37 maddenin 30'unda (%81.08), ABD örnekleme için 37 maddenin 20'sinde (%54.05) ihmal edilemeyecek düzeyde DMF görülmüştür. Bu açıdan bakıldığında Türkiye örnekleminde DMF içeren madde sayısının ABD örnekleminde DMF içeren madde sayısına göre manidar şekilde yüksek olduğu görülmüştür. Le'ye göre (2009) uluslararası geniş katılımlı sınavlar açısından cinsiyet değişkeninde DMF gösteren maddelerin olabileceğini belirtmiştir. Demir (2013) ise PISA 2009 üzerinden yaptığı çalışmada bu çalışmayla benzer sonuçlara ulaşmış ve hem kültür hem de cinsiyet açısından oldukça yüksek oranda DMF'li madde tespit etmiştir.

ABD ve Türkiye örneklemi arasında kültürel açıdan yapılan analizlerde 37 maddenin 36'sında (%97.29) ihmal edilemeyecek düzeyde DMF içeren madde görülmesinin yanı sıra bu maddelerin de 34'ünün yüksek düzeyde DMF içerdiği görülmüştür. Elde edilen bu sonuç dilsel ve kültürel farklılıklar arttıkça maddelerin DMF

ve olası yanlılık gösterme ihtimalinin arttığı yönünde yorumlanabilir. Bu durum benzer örneklemeler üzerinden yapılan başka bir uygulama olan Atalay'ın 2010'da yaptığı çalışmanın sonuçları ile tutarlılık göstermiştir.

5.1.2. Ölçme Değişmezliğine Yönelik Sonuçlar

PISA 2015 uygulamasında yer alan fen bilimlerine yönelik öğrenci tutum anketlerinin ölçme değişmezliği Türkiye ve ABD örneklemeleri açısından 6 alt testte incelenmiştir. ST094 (Fen öğrenmekten zevk alma) ve ST113 (Fen öğreniminde araşsal güdülenme) alt testleri bütün değişmezlik aşamalarını karşılarken, ST095 (Geniş fen konularına ilgi), ST129 (Fene yönelik öz yeterlik), ST131 (Fen bilimlerine yönelik epistemik inançlar) ve ST146 (Bilim aktiviteleri) alt testlerinde yalnızca metrik değişmezlik boyutları sağlanmıştır. Sonuç olarak ST094 ve ST113 testleri kültürler açısından benzerdir ve elde edilen puanların farklılığı alt gruplardan kaynaklanmaktadır. Ancak ST095, ST129, ST131 ve ST146 testleri regresyon sabitleri açısından kültürler arasında değişmez değildir.

Metrik değişmezliğin sağlandığı ama skalar değişmezliğin sağlanmadığı durumlarda Gregorich (2006) o testin yönü belli olmayan şekilde alt gruplardan birine avantaj sağladığı yorumunu yapmıştır. Araştırmada 4 alt test skalar değişmezlik boyutunu sağlamamıştır. Araştırma için yapılan DMF analizlerinden elde edilen çıktılar testlerin kültürel açıdan oldukça yüksek oranda DMF gösteren madde içerdiğini göstermiştir. Bu iki bilgi birleştiğinde ortaya tutarlı bir sonuç çıkarmaktadır. Bu araştırmaya benzer olarak Karakoç Alatl ve Çokluk Bökeoğlu (2018) yaptıkları araştırmada PISA 2012 uygulamasını DMF ve ölçme değişmezliği açısından incelemişler ve sonuç olarak alt testlerde yüksek oranda DMF bulmuşlardır. İlgili çalışmada ÇGDFA ile ölçme değişmezliği analizi de yapılmış ve sonuçlar testlerin ölçme değişmezliğinin tüm aşamaları sağlayamadığı yönünde çıkmıştır. Olası yanlılık kaynakları ise çeviri hataları, belli bir grubun ilgili dile aşinalığı ve anlamsal farklılıklar olarak belirlenmiştir.

5.2. Öneriler

1. Çalışmada PISA 2015 fen bilimlerine yönelik öğrenci tutum anketlerinin Türkiye ve ABD kültürleri bağlamında ölçme değişmezliğini tam olarak sağlayamadığı ve birçok maddesinin DMF gösterdiği için PISA 2015

verilerine dayanarak iki kültürün tutum anket sonuçlarının karşılaştırılması hatalı sonuçlar verebilir. İleriki yıllarda yapılacak PISA uygulamalarına dayalı karşılaştırmalar yapılmadan önce ölçme değişmezliği ve DMF çalışmalarının yapılması önerilir.

2. PISA 2015 uygulaması verilerine göre kültüre dayalı olarak birçok alt testte ölçme değişmezliği sağlanamadığı ve DMF gösteren maddelerin çokluğu dikkate alındığında ileriki uygulamalarda çeviri aşamasında kültürel ve dilsel farklılıklara daha çok dikkat edilmesi önerilir.
3. Bu çalışmada ölçme değişmezliği çalışması yalnızca Türkiye ve ABD örneklemi üzerinden, DMF çalışması ise Türkiye ve ABD arasında kültürel ve iki kültür için cinsiyet değişkenine göre yapılmıştır. Geniş katılımlı sınavların ölçme değişmezliği ve DMF çalışmalarının aynı kültür farklı dil, aynı dil aynı kültür gibi çeşitli varyasyonlar açısından incelenmesi alana yarar sağlayacaktır.
4. Araştırmada DMF analizi yalnızca MTK-OO yöntemi üzerinden çalışılmıştır. Ordinal Lojistik Regresyon, POLY SIBTEST, genelleştirilmiş Mantel-Haenszel, BILOG-MG gibi birden çok yöntemle benzer çalışmalar yapılabilir.
5. Araştırmada alt testlerin belli aşamalarda skalar değişmezliği sağlamadığı görülmüştür. Gelecekteki çalışmalarda alt testler kısmi değişmezlik açısından test edilebilir.
6. Çalışmada ölçme değişmezliği ÇGDFA yöntemiyle incelenmiştir. Benzer çalışmalar farklı yöntemler kullanılarak gerçekleştirilebilir.
7. Araştırmada yalnızca kültür ve cinsiyete dayalı alt gruplar karşılaştırılmıştır. Gelecek çalışmalarda sosyo-ekonomik düzey, okul türü ve iktisadi bölge birimleri gibi farklı değişkenler dikkate alınabilir.
8. Çalışmada DMF bulunan maddelerin yanlışlık gösterip göstermediği, gösteriyor ise yanlışlığın olası nedenleri ve yanlışlığın giderilmesine yönelik öneriler alan uzmanlarının görüşlerine dayalı olarak belirlenebilir.

KAYNAKLAR

- Akın, Ç. A. (2015). *Değişen madde fonksiyonu belirlemede MTK-olabilirlik oranı, ordinal lojistik regresyon ve poly-sibtest yöntemlerinin karşılaştırılması*. E-Uluslararası Eğitim Araştırmalar Dergisi, 6(1), 1-16.
- Asil, M. ve Gelbal, S. (2012) *PISA öğrenci anketinin kültürler arası eşdeğerliği*. Eğitim ve Bilim, 37(166), 236-249.
- Atalay, K. (2010). *PISA 2006 Öğrenci Anketinde Yer Alan Tutum Maddelerinin Değişen Madde Fonksiyonu Açısından İncelenmesi*, Yüksek Lisans Tezi, Ankara.
- Ayan, C. (2011). *PISA 2009 Fen Okuryazarlığı Alt Testinin Değişen Madde Fonksiyonu Açısından İncelenmesi*, Yüksek Lisans Tezi, Ankara.
- Baranik, L. E., Meade, A. W., Lakey, C. E., Lance, C. E., Hu, C., Hua, W. and Michalos, A. (2008). Examining the Differential Item Functioning of the Rosenberg Self-Esteem Scale Across Eight Countries. *Journal of Applied Social Psychology*, 38 (7), 1867-1904.
- Bentler, P.M. (2006) EQS 6.1: Structural Equations Program Manual, Multivariate Software, Inc., Encino, CA.
- Berberoğlu, G. ve Kalender, İ. (2005). *Öğrenci başarısının yıllara, okul türlerine, bölgelere göre incelenmesi: ÖSS ve PISA analizi*. Eğitim Bilimleri ve Uygulama, 4(7), 21-35.
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). New York, NY: The Guilford Press.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equations models* (pp. 136–162). Newbury Park, CA: Sage.
- Camilli, G. ve Shepard, L. A. (1994). *Methods for Identifying Biased Test Items*. Hollywood: Sage Publications.
- Cheung, G. W., and Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255.
- Clauser, E. B., & Mazor, M. K. (1998). *Using statistical procedures to identify differentially functioning test items*. Educational Measurement: Issues and Practice. 17, 31-44.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal Bilimler İçin Çok Değişkenli İstatistik (SPSS ve LISREL Uygulamaları)*. Ankara: Pegem Akademi.

- Demir, S. (2013). *Pisa 2009 Matematik Okuryazarlığı Alt Testinde Bulunan Maddelerinin Mantel-Haenszel, Sibtest ve Lojistik Regresyon Yöntemleri İle Değişen Madde Fonksiyonunun İncelenmesi*, Yüksek Lisans Tezi, Bolu.
- Evran, D. (2019). Assessing measurement invariance: multiple group confirmatory factor analysis for differential item functioning detection in polytomous measures of Turkish and American students, *Harran Maarif Dergisi*, 4 (1), 1-20.
- Gierl, M., Khaliq, S. N., Boughton, K. (1999). *Gender Differential Item Functioning in Mathematics and Science: Prevalence and Policy Implications*, Paper presented at the Canadian Society for the Study of Education.
- Gök, B., Kabasakal, K. A. ve Kelecioğlu, H. (2014). *PISA 2009 öğrenci anketi tutum maddelerinin kültüre göre değişen madde fonksiyonu açısından incelenmesi*, *Eğitimde Ölçme ve Değerlendirme Dergisi* 5(1), 72-87.
- 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, 78-94.
- Gülleroğlu, H. D. (2016). *PISA 2012 Matematik Uygulamasına Katılan Türk Öğrencilerin Duyuşsal Özelliklerinin Cinsiyete Göre Ölçme Değişmezliğinin İncelenmesi*, *GEFAD / GUJGEF* 37(1): 151-175 (2017).
- Hansson, Å. & Gustafsson, J. (2013). Measurement Invariance of Socioeconomic Status Across Migrational Background, *Scandinavian Journal of Educational Research*, 57:2, 148-166.
- Holland P. W. & Wainer H. (1993). *Differential Item Functioning*, United States of American Lawrence Erlbaum Associates, Inc.
- Hooper, D., Coughlan, J., Mullen, M. (2008). *Structural Equation Modelling: Guidelines for Determining Model Fit*. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- İmrol, F. (2017). *PISA 2012 Türkiye Örnekleminde Matematiğe Yönelik Motivasyon ve Öz-İnanç Yapılarının Ölçme Değişmezliğinin İncelenmesi*. Yüksek Lisans Tezi, Ankara.
- Karaduman, B. (2017). *Sınav Stresi Ölçeğinin Uyarlanması ve Ölçme Değişmezliğinin İncelenmesi*. Yüksek Lisans Tezi, Bolu.
- Karakoc Alatl, B., Ayan, C., Polat Demir, B., & Uzun, G. (2016). Examination of the TIMSS 2011 Fourth Grade Mathematics Test in terms of cross-cultural measurement invariance. *Eurasian Journal of Educational Research*, 66, 389-406.

- Karakoç Alatl, B., Çokluk Bökeoğlu, Ö. (2018). Uluslararası Öğrenci Değerlendirme Programı (PISA-2012) Okuryazarlık Testlerinin Ölçme Değişmezliğinin İncelenmesi. *İlköğretim Online*, 17(2), 1096-1115.
- Karami, H. (2012). *An Introduction to Differential Item Functioning*. The International Journal of Educational and Psychological Assessment September 2012, Vol. 11(2)
- Kıbrıslıoğlu, N. (2015). *PISA 2012 matematik öğrenme modelinin kültürlere ve cinsiyete göre ölçme değişmezliğinin incelenmesi: Türkiye-Çin(Şangay)-Endonezya örneği*. Yüksek Lisans Tezi, Ankara.
- Kline, R. B. (2015). *Principles and practices of structural equation modelling* (4th ed.). Guilford.
- Langer, M. M., Hill, C. D., Thissen, D., Burwinkle, T. M., Varni, J. W. and DeWalt, D. A. (2008). Item response theory detected differential item functioning between healthy and ill children in quality-of-life measures, *Journal of Clinical Epidemiology* 61, 268-276.
- Le, L.T. (2009). Investigation gender differential item functioning across countries and test languages for PISA science items. *International Journal of Testing*, 9(2), 122- 133.
- Luo, C. (2010) *Measurement invariance of Rosenberg Self-Esteem Scale between British and Chinese college students*. Yüksek Lisans Tezi, Edinburgh.
- Lyons-Thomas, J., Sandilands, D. and Ercikan, K. (2014). Gender Differential Item Functioning in Mathematics in Four International Jurisdictions, *Eğitim ve Bilim*, 39 (172), 20-32.
- Maccallum, R. C., Browne, M. W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling of fit involving a particular measure of model. *Psychological Methods*, 13(2), 130–149.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.
- MEB, (2013). *PISA 2012 ulusal ön raporu*. MEB, Ankara.
- MEB (2016). *TIMSS 2015 ulusal matematik ve fen bilimleri ön raporu 4. ve 8. sınıflar*, MEB, Ankara
- MEB (2016). *Uluslararası öğrenci değerlendirme programı pisa 2015 ulusal raporu*. MEB, Ankara.
- Meredith, W. (1993). *Measurement Invariance, Factor Analysis and Factorial Invariance*. *Psychometrika*, 58, 525-543.
- OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264266490-en>

- Önen, E. (2009). Ölçme Değişmezliğinin Yapısal Eşitlik Modellemesi Teknikleri ile İncelenmesi. Doktora Tezi, Ankara Üniversitesi, Ankara.
- Pauwels, J. (2018) *Presentation Mode of Personality Questionnaire Items Online and Measurement Invariance*, Yüksek Lisans Tezi, Ghent.
- Sheppard, R., Han, K., Colarelli, S. M., Dai, G. and King, D. W. (2006) Differential Item Functioning by Sex and Race in the Hogan Personality Inventory, *Assessment*, 13 (4), 442-453.
- Tabachnick, B. & Fidell, L. S. (2014). *Using Multivariate Statistics* (6. Baskı). Boston: Allyn & Bacon.
- Thissen, D. (2001). *IRTLRDIF v.2.0b: Software for the computation of the statistics involved in item response theory likelihood-ratio tests for differential item functioning*. Retrieved November 06, 2018, from <http://www.unc.edu/~dthissen/dl.html>.
- Uyar, Ş. (2011). *PISA 2009 Türkiye Örnekleminde Öğrenme Stratejileri Modelinin Farklı Gruplarda Ölçme Değişmezliğinin İncelenmesi*. Yüksek Lisans Tezi, Ankara.
- Uzun, B., Öğretmen, T. (2010). Fen Başarısı İle İlgili Bazı Değişkenlerin TIMSS-R Türkiye Örnekleminde Cinsiyete Göre Ölçme Değişmezliğinin Değerlendirilmesi. *Eğitim ve Bilim*, 35(155), 26-35.
- Wang, J., and Wang, X. (2012). *Structural Equation Modeling: Applications Using Mplus*. John Wiley & Sons.
- Woods, C. M. (2008). Likelihood-Ratio DIF Testing: Effects of Nonnormality, *Applied Psychological Measurement*, 32 (7), 511-526.
- Yandı, A., Köse, İ. A. ve Uysal, Ö. (2017) Farklı yöntemlerle ölçme değişmezliğinin incelenmesi: PISA 2012 örneği. *Mersin Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 13(1), 243-253.
- Yandı, A., Köse, İ. A., Uysal, Ö. ve Oğul, G. V. (2017) PISA 2015 öğrenci anketinin (St094Q01na-St094Q05na) ölçme değişmezliğinin farklı yöntemlerle incelenmesi. *Küreselleşen dünyada eğitim* (1. baskı) içinde (s.333-344). Ankara: Pegem Akademi Yayıncılık.
- Yıldırım, H. H. (2006). *Uluslararası değerlendirme çalışmalarında kullanılan matematik sorularının madde yanlılığı yöntemleri kullanılarak incelenmesi*. Doktora Tezi, Ankara.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Directorate of Human Resources Research and Evaluation, Department of National Defense. Ottawa: Canada.



EKLER

Ek 1. PISA 2015 fen bilimlerine yönelik öğrenci tutum anketlerinin Türkçe Formları

St095: Geniş Fen Konularına ilgi

Aşağıda yer alan fen bilimleri ile ilgili konularla ne ölçüde ilgileniyorsunuz?					
(Lütfen her bir satırdan bir cevap seçiniz.)					
	<i>İlgilenmiyorum</i>	<i>Fazla ilgi duymuyorum</i>	<i>İlgileniyorum</i>	<i>Çok ilgileniyorum</i>	<i>Bunun ne olduğunu bilmiyorum</i>
Biyosfer (ör. ekosistem hizmetleri, sürdürülebilirlik)	ST095Q04NA01 ☐	ST095Q04NA02 ☐	ST095Q04NA03 ☐	ST095Q04NA04 ☐	ST095Q04NA05 ☐
Hareket ve kuvvetler (ör. hız, sürtünme, manyetik ve yer çekimi kuvvetleri)	ST095Q07NA01 ☐	ST095Q07NA02 ☐	ST095Q07NA03 ☐	ST095Q07NA04 ☐	ST095Q07NA05 ☐
Enerji ve dönüşümü (ör. enerjinin korunması, kimyasal reaksiyonlar)	ST095Q08NA01 ☐	ST095Q08NA02 ☐	ST095Q08NA03 ☐	ST095Q08NA04 ☐	ST095Q08NA05 ☐
Evren ve tarihçesi	ST095Q13NA01 ☐	ST095Q13NA02 ☐	ST095Q13NA03 ☐	ST095Q13NA04 ☐	ST095Q13NA05 ☐
Bilimin, hastalıkların önlememizde bize nasıl yardımcı olduğu	ST095Q15NA01 ☐	ST095Q15NA02 ☐	ST095Q15NA03 ☐	ST095Q15NA04 ☐	ST095Q15NA05 ☐

St131: Fen Bilimlerine Yönelik Epistemik İnançlar

Aşağıdaki ifadelerle ne derece katılıyorsunuz ya da katılmıyorsunuz?				
(Lütfen her bir satırdan bir cevap seçiniz.)				
	<i>Kesinlikle katılmıyorum</i>	<i>Katılmıyorum</i>	<i>Katılıyorum</i>	<i>Kesinlikle katılıyorum</i>
Bir şeyin doğru olup olmadığını bilmenin en iyi yolu deney yapmaktır.	ST131Q01NA01 ☐	ST131Q01NA02 ☐	ST131Q01NA03 ☐	ST131Q01NA04 ☐
Fen bilimlerindeki fikirler bazen değişmektedir.	ST131Q03NA01 ☐	ST131Q03NA02 ☐	ST131Q03NA03 ☐	ST131Q03NA04 ☐
İyi cevaplar çok farklı deneylerden elde edilen kanıtlara dayanmaktadır.	ST131Q04NA01 ☐	ST131Q04NA02 ☐	ST131Q04NA03 ☐	ST131Q04NA04 ☐
Bulgularınızdan emin olmak için birden fazla deney yapmak iyi bir yoldur.	ST131Q06NA01 ☐	ST131Q06NA02 ☐	ST131Q06NA03 ☐	ST131Q06NA04 ☐
Bazen fen bilimlerinde çalışan bilim insanları, bilimdeki doğrular hakkındaki düşüncelerini değiştirmektedirler.	ST131Q08NA01 ☐	ST131Q08NA02 ☐	ST131Q08NA03 ☐	ST131Q08NA04 ☐
Fen bilimleri kitaplarında yer alan fikirler bazen değişmektedir.	ST131Q11NA01 ☐	ST131Q11NA02 ☐	ST131Q11NA03 ☐	ST131Q11NA04 ☐

St094: Fen Öğrenmekten Zevk Alma

Aşağıda yer alan kendiniz ile ilgili ifadelere ne derece katılıyorsunuz ya da katılmıyorsunuz?				
(Lütfen her bir satırdan bir cevap seçiniz.)				
	Kesinlikle katılmıyorum	Katılmıyorum	Katılıyorum	Kesinlikle katılıyorum
Fen bilimleri ile ilgili konuları öğrenirken genellikle eğlenirim.	ST094Q01NA01 ☐	ST094Q01NA02 ☐	ST094Q01NA03 ☐	ST094Q01NA04 ☐
Fen bilimleri ile ilgili yazıları okumaktan zevk alırım.	ST094Q02NA01 ☐	ST094Q02NA02 ☐	ST094Q02NA03 ☐	ST094Q02NA04 ☐
Fen bilimleri ile ilgili konularda çalışmak beni mutlu ediyor.	ST094Q03NA01 ☐	ST094Q03NA02 ☐	ST094Q03NA03 ☐	ST094Q03NA04 ☐
Fen bilimleri hakkında yeni bilgiler edinmekten zevk alırım.	ST094Q04NA01 ☐	ST094Q04NA02 ☐	ST094Q04NA03 ☐	ST094Q04NA04 ☐
Fen bilimleri ile ilgili bir şeyler öğrenmeye ilgi duyarım.	ST094Q05NA01 ☐	ST094Q05NA02 ☐	ST094Q05NA03 ☐	ST094Q05NA04 ☐

St129: Fene Yönelik Öz Yeterlik

Aşağıdaki çalışmaları kendi başınıza yapmanızın ne kadar kolay olacağını düşünürsünüz?				
(Lütfen her bir satırdan bir cevap seçiniz.)				
	Bunu kolaylıkla yapabilirim	Bunu uğraşırsam yapabilirim	Bunu yapabilmek için çok çaba göstermem gerekir	Bunu yapamam
Bir sağlık sorununu ele alan gazete raporunun temelinde yatan, fen bilimleri ile ilgili soruyu tanıma	ST129Q01TA01 ☐	ST129Q01TA02 ☐	ST129Q01TA03 ☐	ST129Q01TA04 ☐
Bazı alanlardaki depremlerin diğer alanlardan niçin daha sık meydana geldiğini açıklama	ST129Q02TA01 ☐	ST129Q02TA02 ☐	ST129Q02TA03 ☐	ST129Q02TA04 ☐
Antibiyotiklerin hastalıkların tedavisindeki rolünü belirtme	ST129Q03TA01 ☐	ST129Q03TA02 ☐	ST129Q03TA03 ☐	ST129Q03TA04 ☐
Fen bilimleri açısından, atıkların uygun şekilde toplanarak işlenmesi ile ilgili sorunu belirleme	ST129Q04TA01 ☐	ST129Q04TA02 ☐	ST129Q04TA03 ☐	ST129Q04TA04 ☐
Çevredeki değişimlerin belli canlı türlerinin hayatta kalmasını nasıl etkileyebileceğini önceden tahmin etme	ST129Q05TA01 ☐	ST129Q05TA02 ☐	ST129Q05TA03 ☐	ST129Q05TA04 ☐
Gıda maddelerinin üzerine konan etiketlerde verilen bilimsel bilgileri yorumlama	ST129Q06TA01 ☐	ST129Q06TA02 ☐	ST129Q06TA03 ☐	ST129Q06TA04 ☐
Yeni kanıtların, Mars'ta hayat olup olmadığına ilişkin anlayışınızı nasıl değiştirebileceğini tartışma	ST129Q07TA01 ☐	ST129Q07TA02 ☐	ST129Q07TA03 ☐	ST129Q07TA04 ☐
Asit yağmurlarının nasıl oluştuğunu açıklayan iki görüşten hangisinin daha iyi olduğunu belirleme	ST129Q08TA01 ☐	ST129Q08TA02 ☐	ST129Q08TA03 ☐	ST129Q08TA04 ☐

St146: Bilim Aktiviteleri

Aşağıdakileri ne kadar sık yaparsınız?				
(Lütfen her bir satırdan bir cevap seçiniz.)				
	Çok sık	Düzenli aralıklarla	Bazen	Hiç ya da hemen hemen hiç
Fen bilimleri ile ilgili televizyon programlarını izleme	ST146Q01TA01 ☐	ST146Q01TA02 ☐	ST146Q01TA03 ☐	ST146Q01TA04 ☐
Fen bilimleri konuları ele alan kitapları ödünç alma ya da satın alma	ST146Q02TA01 ☐	ST146Q02TA02 ☐	ST146Q02TA03 ☐	ST146Q02TA04 ☐
Fen bilimleri konuları ile ilgili internet sitelerini ziyaret etme	ST146Q03TA01 ☐	ST146Q03TA02 ☐	ST146Q03TA03 ☐	ST146Q03TA04 ☐
Fen bilimleri ile ilgili dergileri ya da gazete makalelerini okuma	ST146Q04TA01 ☐	ST146Q04TA02 ☐	ST146Q04TA03 ☐	ST146Q04TA04 ☐
Bilim kulübüne katılma	ST146Q05TA01 ☐	ST146Q05TA02 ☐	ST146Q05TA03 ☐	ST146Q05TA04 ☐
Bilgisayar programlarında ve sanal laboratuvarlarda doğal olayları yapay biçimde canlandırma	ST146Q06NA01 ☐	ST146Q06NA02 ☐	ST146Q06NA03 ☐	ST146Q06NA04 ☐
Bilgisayar programlarında ve sanal laboratuvarlarda teknik süreçleri yapay biçimde yapma	ST146Q07NA01 ☐	ST146Q07NA02 ☐	ST146Q07NA03 ☐	ST146Q07NA04 ☐
Ekoloji kuruluşlarının internet sitelerini ziyaret etme	ST146Q08NA01 ☐	ST146Q08NA02 ☐	ST146Q08NA03 ☐	ST146Q08NA04 ☐
Bloglar ve mikrobloglama üzerinden bilim, çevre veya ekoloji kuruluşlarının haberlerini takip etme	ST146Q09NA01 ☐	ST146Q09NA02 ☐	ST146Q09NA03 ☐	ST146Q09NA04 ☐

St113: Fen Öğreniminde Araçsal Güdülenme

Aşağıdaki ifadelere ne ölçüde katılıyorsunuz?				
(Lütfen her bir satırdan bir cevap seçiniz.)				
	Kesinlikle katılıyorum	Katılıyorum	Katılmıyorum	Kesinlikle katılmıyorum
Daha sonra yapacağım çalışmalarda bana yardımcı olacağından, okuldaki fen dersi konularını (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrenmek için çaba göstermeye değer.	ST113Q01TA01 ☐	ST113Q01TA02 ☐	ST113Q01TA03 ☐	ST113Q01TA04 ☐
Sonradan yapmak isteyeceğim şeyler için ihtiyaç duyacağımdan, okuldaki fen dersi konularında (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrendiklerim benim için önemlidir.	ST113Q02TA01 ☐	ST113Q02TA02 ☐	ST113Q02TA03 ☐	ST113Q02TA04 ☐
Öğreneceklerim geleceğe yönelik mesleki beklentilerime yardımcı olacağından, okuldaki fen dersi konularını (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrenmek benim için önemlidir.	ST113Q03TA01 ☐	ST113Q03TA02 ☐	ST113Q03TA03 ☐	ST113Q03TA04 ☐
Fen dersi konularında (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrendiğim bir çok şey, iş bulmamda bana yardımcı olacak.	ST113Q04TA01 ☐	ST113Q04TA02 ☐	ST113Q04TA03 ☐	ST113Q04TA04 ☐

Ek 2. PISA 2015 fen bilimlerine yönelik öğrenci tutum anketlerinin İngilizce Formları

St095: Interest in Science

To what extent are you interested in the following science topics?					
(Please select one response in each row.)					
	Not interested	Hardly interested	Interested	Highly interested	I don't know what this is
Biosphere (e.g., ecosystem services, sustainability)	ST095Q04NA01 ☐	ST095Q04NA02 ☐	ST095Q04NA03 ☐	ST095Q04NA04 ☐	ST095Q04NA05 ☐
Motion and forces (e.g., velocity, friction, magnetic and gravitational forces)	ST095Q07NA01 ☐	ST095Q07NA02 ☐	ST095Q07NA03 ☐	ST095Q07NA04 ☐	ST095Q07NA05 ☐
Energy and its transformation (e.g., conservation, chemical reactions)	ST095Q08NA01 ☐	ST095Q08NA02 ☐	ST095Q08NA03 ☐	ST095Q08NA04 ☐	ST095Q08NA05 ☐
The universe and its history	ST095Q13NA01 ☐	ST095Q13NA02 ☐	ST095Q13NA03 ☐	ST095Q13NA04 ☐	ST095Q13NA05 ☐
How science can help us prevent disease	ST095Q15NA01 ☐	ST095Q15NA02 ☐	ST095Q15NA03 ☐	ST095Q15NA04 ☐	ST095Q15NA05 ☐

St131: Epistemic Beliefs About Science

How much do you disagree or agree with the statements below?				
(Please select one response in each row.)				
	Strongly disagree	Disagree	Agree	Strongly agree
A good way to know if something is true is to do an experiment.	ST131Q01NA01 ☐	ST131Q01NA02 ☐	ST131Q01NA03 ☐	ST131Q01NA04 ☐
Ideas in science sometimes change.	ST131Q03NA01 ☐	ST131Q03NA02 ☐	ST131Q03NA03 ☐	ST131Q03NA04 ☐
Good answers are based on evidence from many different experiments.	ST131Q04NA01 ☐	ST131Q04NA02 ☐	ST131Q04NA03 ☐	ST131Q04NA04 ☐
It is good to try experiments more than once to make sure of your findings.	ST131Q06NA01 ☐	ST131Q06NA02 ☐	ST131Q06NA03 ☐	ST131Q06NA04 ☐
Sometimes scientists change their minds about what is true in science.	ST131Q08NA01 ☐	ST131Q08NA02 ☐	ST131Q08NA03 ☐	ST131Q08NA04 ☐
The ideas in science books sometimes change.	ST131Q11NA01 ☐	ST131Q11NA02 ☐	ST131Q11NA03 ☐	ST131Q11NA04 ☐

St094: Enjoyment of Science

How much do you disagree or agree with the statements about yourself below?				
(Please select one response in each row.)				
	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
I generally have fun when I am learning science topics.	ST094Q01NA01 ☐	ST094Q01NA02 ☐	ST094Q01NA03 ☐	ST094Q01NA04 ☐
I like reading about science.	ST094Q02NA01 ☐	ST094Q02NA02 ☐	ST094Q02NA03 ☐	ST094Q02NA04 ☐
I am happy working on science topics.	ST094Q03NA01 ☐	ST094Q03NA02 ☐	ST094Q03NA03 ☐	ST094Q03NA04 ☐
I enjoy acquiring new knowledge in science.	ST094Q04NA01 ☐	ST094Q04NA02 ☐	ST094Q04NA03 ☐	ST094Q04NA04 ☐
I am interested in learning about science.	ST094Q05NA01 ☐	ST094Q05NA02 ☐	ST094Q05NA03 ☐	ST094Q05NA04 ☐

St129: Science self-efficacy

How easy do you think it would be for you to perform the following tasks on your own?				
(Please select one response in each row.)				
	<i>I could do this easily</i>	<i>I could do this with a bit of effort</i>	<i>I would struggle to do this on my own</i>	<i>I couldn't do this</i>
Recognize the science question that underlies a newspaper report on a health issue	ST129Q01TA01 ☐	ST129Q01TA02 ☐	ST129Q01TA03 ☐	ST129Q01TA04 ☐
Explain why earthquakes occur more frequently in some areas than in others	ST129Q02TA01 ☐	ST129Q02TA02 ☐	ST129Q02TA03 ☐	ST129Q02TA04 ☐
Describe the role of antibiotics in the treatment of disease	ST129Q03TA01 ☐	ST129Q03TA02 ☐	ST129Q03TA03 ☐	ST129Q03TA04 ☐
Identify the science question associated with the disposal of garbage	ST129Q04TA01 ☐	ST129Q04TA02 ☐	ST129Q04TA03 ☐	ST129Q04TA04 ☐
Predict how changes to an environment will affect the survival of certain species	ST129Q05TA01 ☐	ST129Q05TA02 ☐	ST129Q05TA03 ☐	ST129Q05TA04 ☐
Interpret the scientific information provided on the labels of food items	ST129Q06TA01 ☐	ST129Q06TA02 ☐	ST129Q06TA03 ☐	ST129Q06TA04 ☐
Discuss how new evidence can lead you to change your understanding about the possibility of life on Mars	ST129Q07TA01 ☐	ST129Q07TA02 ☐	ST129Q07TA03 ☐	ST129Q07TA04 ☐
Identify the better of two explanations for the formation of acid rain	ST129Q08TA01 ☐	ST129Q08TA02 ☐	ST129Q08TA03 ☐	ST129Q08TA04 ☐

St146: Science Activities

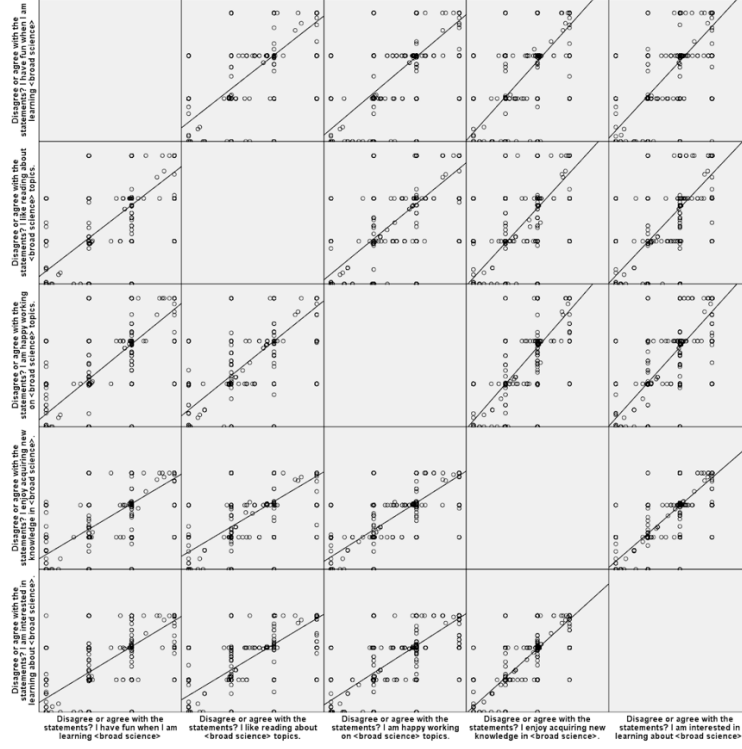
How often do you do these things?				
(Please select one response in each row.)				
	Very often	Regularly	Sometimes	Never or hardly ever
Watch TV programs about science	ST146Q01TA01 ☐	ST146Q01TA02 ☐	ST146Q01TA03 ☐	ST146Q01TA04 ☐
Borrow or buy books on science topics	ST146Q02TA01 ☐	ST146Q02TA02 ☐	ST146Q02TA03 ☐	ST146Q02TA04 ☐
Visit web sites about science topics	ST146Q03TA01 ☐	ST146Q03TA02 ☐	ST146Q03TA03 ☐	ST146Q03TA04 ☐
Read science magazines or science articles in newspapers	ST146Q04TA01 ☐	ST146Q04TA02 ☐	ST146Q04TA03 ☐	ST146Q04TA04 ☐
Attend a science club	ST146Q05TA01 ☐	ST146Q05TA02 ☐	ST146Q05TA03 ☐	ST146Q05TA04 ☐
Use computer programs or virtual lab to simulate natural phenomena	ST146Q06NA01 ☐	ST146Q06NA02 ☐	ST146Q06NA03 ☐	ST146Q06NA04 ☐
Use computer programs or virtual lab to simulate technical processes	ST146Q07NA01 ☐	ST146Q07NA02 ☐	ST146Q07NA03 ☐	ST146Q07NA04 ☐
Visit websites of environmental organizations	ST146Q08NA01 ☐	ST146Q08NA02 ☐	ST146Q08NA03 ☐	ST146Q08NA04 ☐
Follow news of science, environmental, or ecology organizations via blogs and microblogging	ST146Q09NA01 ☐	ST146Q09NA02 ☐	ST146Q09NA03 ☐	ST146Q09NA04 ☐

St113: Instrumental Motivation to Learn Science

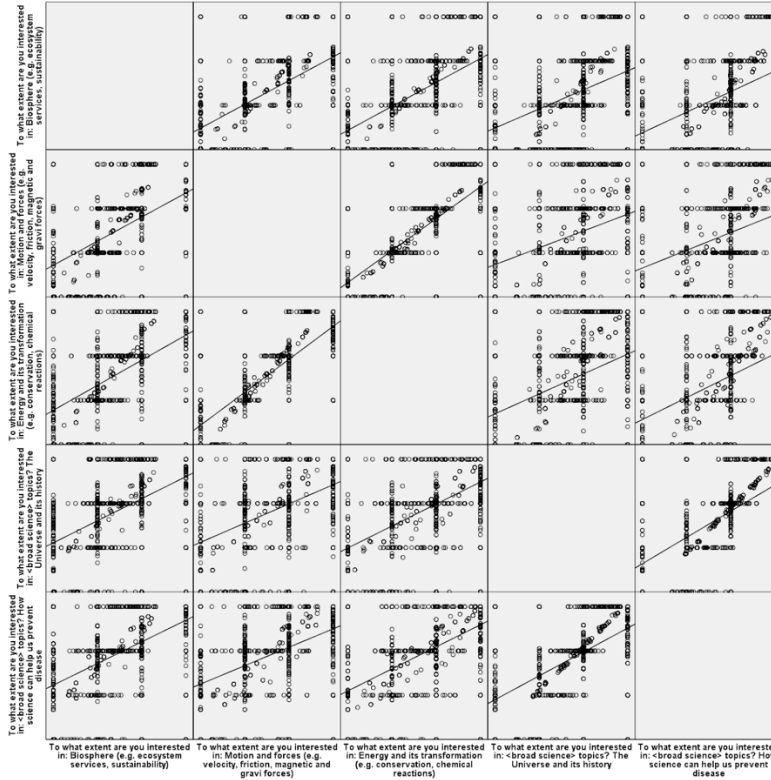
How much do you agree with the statements below?				
(Please select one response in each row.)				
	Strongly agree	Agree	Disagree	Strongly disagree
Making an effort in my science class(es) is worth it because this will help me in the work I want to do later on	ST113Q01TA01 ☐	ST113Q01TA02 ☐	ST113Q01TA03 ☐	ST113Q01TA04 ☐
What I learn in my science classes is important for me because I need this for what I want to do later on.	ST113Q02TA01 ☐	ST113Q02TA02 ☐	ST113Q02TA03 ☐	ST113Q02TA04 ☐
Studying science is worthwhile for me because what I learn will improve my career prospects	ST113Q03TA01 ☐	ST113Q03TA02 ☐	ST113Q03TA03 ☐	ST113Q03TA04 ☐
Many things I learn in my science subject(s) will help me to get a job.	ST113Q04TA01 ☐	ST113Q04TA02 ☐	ST113Q04TA03 ☐	ST113Q04TA04 ☐

Ek 3. Doğrusallık varsayımı için Saçılım Diyagramı Testleri

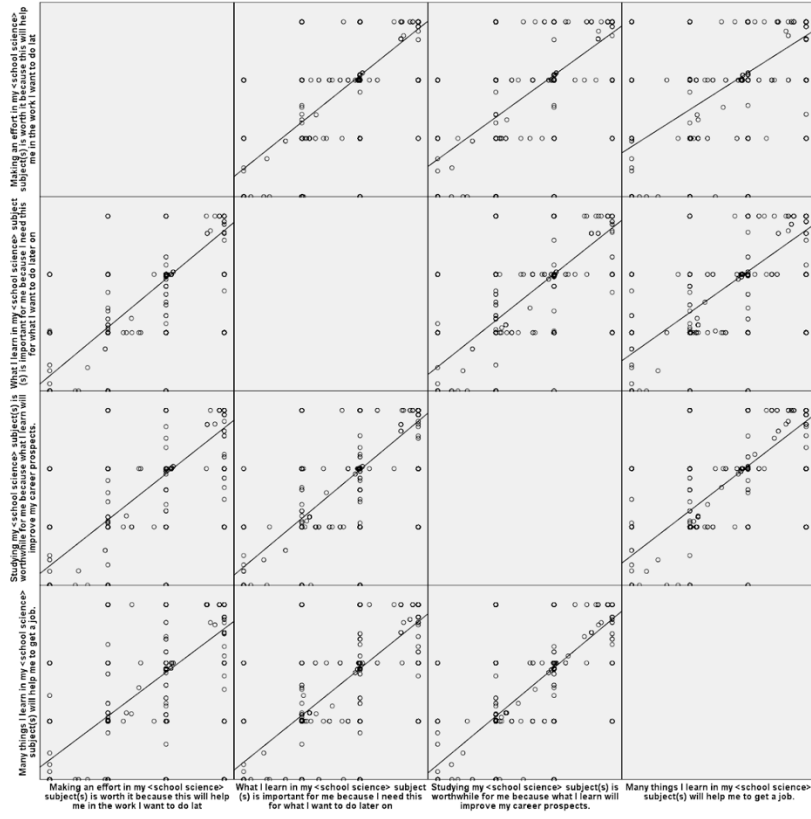
St094 Doğrusallık Varsayımı İçin Saçılım Diyagramı Matrisi



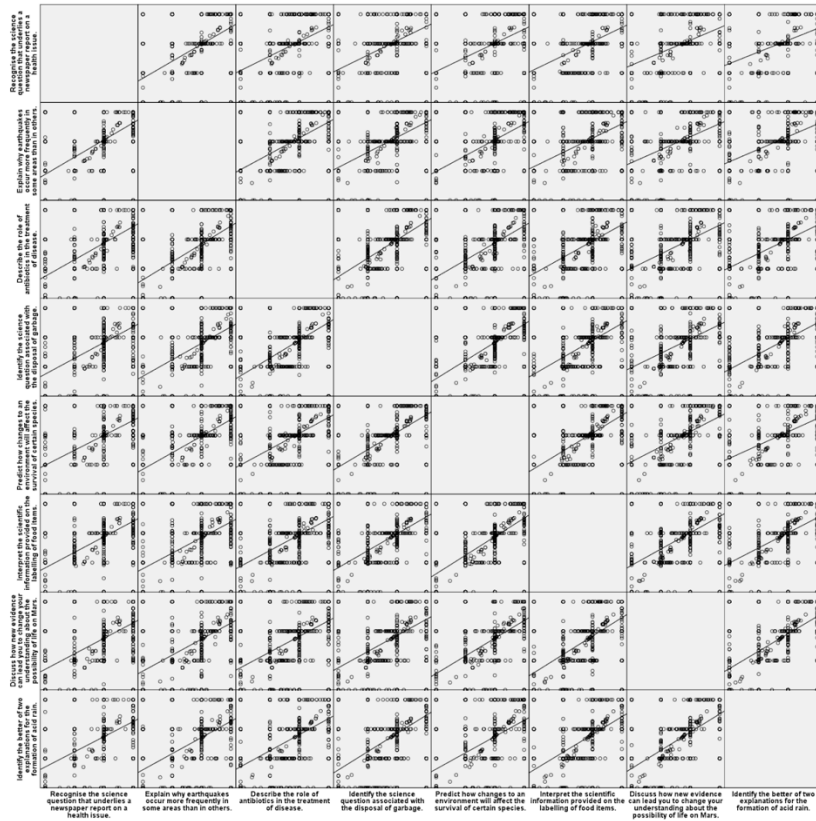
St095 Doğrusallık Varsayımı İçin Saçılım Diyagramı



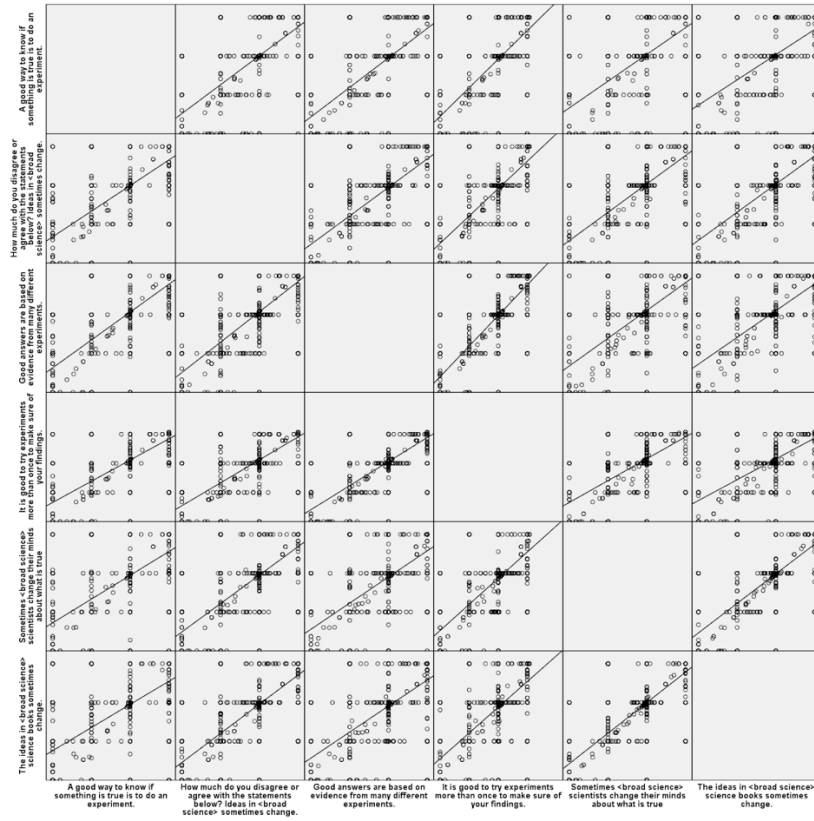
St113 Doğrusallık Varsayımı İçin Saçılım Diyagramı Matrisi



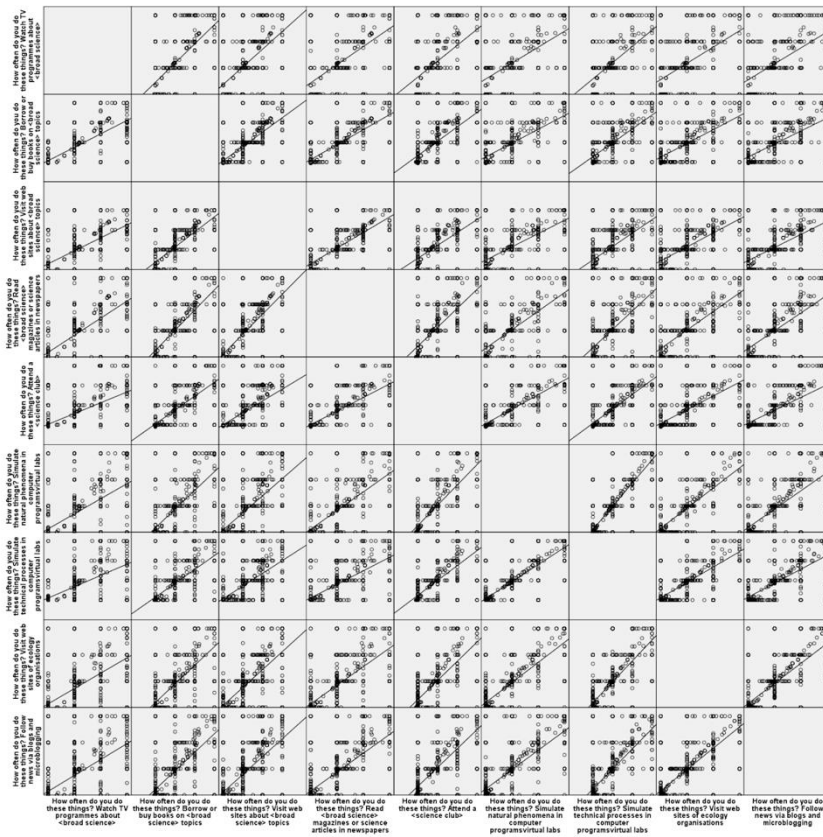
St129 Doğrusallık Varsayımı İçin Saçılım Diyagramı Matrisi



St131 Doğrusallık Varsayımı İçin Saçılım Diyagramı Matrisi



St146 Doğrusallık Varsayımı İçin Saçılım Diyagramı Matrisi



Ek 4. MTK-OO Örnek Çıktısı

ST094 Cinsiyete Göre Türkiye MTK-OO Çıktısı

	g2	df	a	b1	b2	b3	a	b1	b2	b3	μ	σ
1 All equal	34.5	4	3.55	-0.95	-0.23	1.05	3.87	-1.06	-0.32	1.03	0.01	1.00
1 a equal	4.0	1	3.71	-0.94	-0.22	1.03	3.71	-1.08	-0.32	1.05	0.01	1.01
1 b equal	30.5	3	3.69	-1.01	-0.27	1.04	3.69	-1.01	-0.27	1.04	0.01	1.00
2 All equal	2.5	4	4.31	-1.12	-0.25	0.97	4.56	-1.13	-0.26	0.95	0.01	1.00
3 All equal	29.8	4	4.96	-1.09	-0.26	0.94	5.64	-1.03	-0.17	0.99	0.02	1.00
3 a equal	6.4	1	5.28	-1.08	-0.26	0.93	5.28	-1.04	-0.17	1.00	0.02	1.00
3 b equal	23.5	3	5.25	-1.06	-0.21	0.97	5.25	-1.06	-0.21	0.97	0.01	1.00
4 All equal	25.3	4	4.90	-1.17	-0.42	0.79	6.27	-1.12	-0.44	0.77	0.01	1.00
4 a equal	21.2	1	5.48	-1.14	-0.41	0.77	5.48	-1.14	-0.45	0.79	0.01	1.00
4 b equal	4.1	3	5.47	-1.14	-0.43	0.78	5.47	-1.14	-0.43	0.78	0.01	1.00
5 All equal	12.5	4	4.63	-1.17	-0.44	0.76	5.25	-1.14	-0.45	0.79	0.01	1.00
5 a equal	6.7	1	4.92	-1.15	-0.43	0.74	4.92	-1.16	-0.45	0.80	0.01	1.00
5 b equal	5.9	3	4.90	-1.15	-0.44	0.77	4.90	-1.15	-0.44	0.77	0.01	1.00

Ek 5. Etik Kurul Kararı

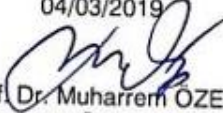
ANKARA ÜNİVERSİTESİ SOSYAL BİLİMLER ALT ETİK KURULU KARAR ÖRNEĞİ

Karar Tarihi : 04/03/2019
Toplantı Sayısı : 03
Karar Sayısı : 55

55- Üniversitemiz Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Anabilim Dalı yüksek lisans öğrencisi **Furkan Tiryaki**'nin "PISA 2015 Öğrenci Tutum Anketlerinin Değişen Madde Fonksiyonu ve Ölçme Değişmezliğinin İncelenmesi" başlıklı tezi ile ilgili 28/01/2019 tarihli "İnsan Üzerinde Yapılan Klinik Dışı Araştırmalar Başvuru Formu" Etik Kurulumuzca incelendi.

Üniversitemiz Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Anabilim Dalı yüksek lisans öğrencisi **Furkan Tiryaki**'nin "PISA 2015 Öğrenci Tutum Anketlerinin Değişen Madde Fonksiyonu ve Ölçme Değişmezliğinin İncelenmesi" başlıklı tezinin, araştırma protokolüne uyulması ve etik onay tarihinden itibaren geçerli olması koşuluyla uygulanmasının etik açıdan uygun olduğuna oy birliği ile karar verildi.

ASLININ AYNI DİR
04/03/2019


Prof. Dr. Muharrem ÖZEN
Ankara Üniversitesi
Etik Kurulu Başkanı

BENZERLİK BİLDİRİMİ

“PISA 2015 Öğrenci Tutum Anketlerinin Değişen Madde Fonksiyonu ve Ölçme Değişmezliğinin İncelenmesi” başlıklı tezimin ana bölümü (ön bölüm, kaynaklar ve ekler hariç) Turnitin İntihali Engelleme Programı aracılığıyla incelenmiş ve ilgili rapor danışmanım tarafından da kontrol edilmiştir. Kontrol sırasında (1) “Kapak, Önsöz, Özet, İçindekiler, Kısaltmalar/ Simgeler Sayfası, Tablolar Dizini, Şekiller Dizini, Dipnotlar, Kaynaklar, Ekler, Özgeçmiş” (2) “Tırnak içinde ve blok olarak verilen (doğrudan) alıntılar” (3) “Yedi sözcüğe kadar olan benzerlik/örtüşme içeren metin kısımları” (4) “Öğrencinin lisansüstü tezi ile ilgili yapmış olduğu yayınlar ile yararlandığı mevzuat metinleri” dışarıda tutulmuştur. Benzerlik kontrolüne ilişkin rapordan elde edilen sonuçlar aşağıda sunulmuştur.

Rapor Tarihi	: 11.06.2019
Gönderim Numarası	: 1142526820
Sayfa Sayısı	: 74
Sözcük Sayısı	: 11255
Karakter Sayısı	: 74286
Benzerlik Oranı	: %7
Savunma Tarihi	: 19.06.2019

Yukarıda belirtilen sonuçları gösteren Turnitin İntihali Engelleme Programı’na ilişkin orijinal raporu, sonuçlarda herhangi bir değişiklik yapmaksızın bu bildirimin ekinde Enstitüye teslim ettiğimi, tezimin %10’dan fazla benzerlik oranı içerdiğinin ve tek bir kaynakla eşleşme oranının %2’yi geçtiğinin belirlenmesi durumunda, bundan doğabilecek tüm yasal sorumluluğu kabul ettiğimi bildirir, saygılarımı sunarım.

Öğrencinin Adı Soyadı: Furkan TIRYAKI

Tarih: 12.06.2019

İmza: 

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı ve Soyadı : Furkan TİRYAKİ
E-Posta Adresi : furkanmvo@gmail.com

İş Deneyimi :

Unvan	Görev Yeri	Yıl
Ölçme ve Değerlendirme Uzmanı	Zafer Koleji	2017

Akademik Bilgiler

Öğrenim Durumu:

Derece	Bölüm/Program	Üniversite	Yıl
3.21	Rehberlik ve Psikolojik Danışmanlık	Başkent Üniversitesi	2015

Yayınlar: