



HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

PISA 2015 UYGULAMASINDAKİ MADDELERİN KÜLTÜRE GÖRE DEĞİŞEN
MADDE FONKSİYONU AÇISINDAN İNCELENMESİ

Esengül GÜR

Yüksek Lisans Tezi

Ankara, 2019

Liderlik, arařtırma, inovasyon, kaliteli eęitim ve deęiřim ile

Daha ileriye ... En İyiyeye ...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

PISA 2015 UYGULAMASINDAKİ MADDELERİN KÜLTÜRE GÖRE DEĞİŞEN
MADDE FONKSİYONU AÇISINDAN İNCELENMESİ

AN INVESTIGATION OF THE PISA 2015 IN TERMS OF DIFFERENTIAL ITEM
FUNCTIONING BASED ON CULTURE




Esengül GÜR

Yüksek Lisans Tezi

Ankara, 2019

Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,
Eseng¼l G¼R¼n hazırladıđı “PISA 2015 Uygulamasındaki Maddelerinin K¼lt¼re
G¼re Deđiřen Madde Fonksiyonu Açıřından İncelenmesi” bařlıklı bu alıřma
j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, ¼lme ve Deđerlendirme Bilim
Dalında Y¼ksek Lisans Tezi** olarak kabul edilmiřtir.

J¼ri Bařkanı	Do. Dr. Dilara BAKAN KALAYCIOĐLU	 İmza
J¼ri Üyesi (Danıřman)	Dr. ¼đr. Üyesi K¼bra ATALAY KABASAKAL	 İmza
J¼ri Üyesi	Do. Dr. Burcu ATAR	 İmza

Bu tez Hacettepe ¼niversitesi Lisans¼st¼ Eđitim, ¼đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından 30/05/2019 tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca / / tarihinde kabul edilmiřtir.

Prof. Dr. Ali Ekber řAHİN
Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

Öz

Bu çalışmanın amacı, PISA 2015 öğrenci anketinde yer alan fen bilimlerine karşı tutum maddelerinin dil ve kültüre göre değişen madde fonksiyonu (DMF) gösterip göstermediğini incelemektir. DMF analizleri için geliştirilmiş Mantel Haenszel (GMH), ordinal lojistik regresyon (OLR) ve poly-SIBTEST yöntemleri kullanılmıştır. GMH yöntemi için GMHDIF programı, poly-SIBTEST ve OLR yöntemleri için R yazılımından faydalanılmıştır. Çalışmanın veri setini PISA 2015 uygulamasına katılan 5089 öğrenci ile Türkiye, 5157 öğrenci ile İrlanda, 5000 öğrenci ile Amerika Birleşik Devletleri (ABD) ve 4417 öğrenci ile İngiltere örneklemi oluşturmaktadır. Yapılan analizler sonucunda dokuz tutum maddesinden, İngiltere-İrlanda karşılaştırmasında GMH ve OLR yöntemleriyle altı maddede ve poly-SIBTEST yöntemiyle yedi maddede DMF belirlenmiştir. İngiltere-ABD karşılaştırmasında GMH ve OLR yöntemleriyle yedi maddede ve poly-SIBTEST yöntemiyle altı maddede DMF belirlenmiştir. İngiltere-Türkiye karşılaştırmasında ise GMH yöntemiyle sekiz maddede, OLR yöntemiyle dokuz maddede ve poly-SIBTEST yöntemiyle yedi maddede DMF tespit edilmiştir. Her üç yönteme göre İngiltere-İrlanda ile İngiltere-ABD karşılaştırmalarında beş maddede, İngiltere-Türkiye karşılaştırmasında ise yedi maddede DMF içeren tutum maddeleri saptanmıştır. Araştırma sonucunda analizde kullanılan OLR, GMH ve poly-SIBTEST yöntemlerinin tespit ettikleri DMF'li madde sayısında uyum gözlenirken, OLR ve poly-SIBTEST yöntemlerinde DMF'li maddelerin düzeylerinde farklılıklar olduğu gözlenmiştir. Bununla birlikte, OLR ve poly-SIBTEST yöntemleriyle DMF içeren maddeler incelendiğinde DMF türleri ve maddelerin avantaj sağladığı gruplar arasında uyumluluk gözlenmiştir.

Anahtar sözcükler: PISA, değişen madde fonksiyonu, ordinal lojistik regresyon, geliştirilmiş Mantel Haenszel, poly-SIBTEST

Abstract

The aim of this study is to research whether or not the attitude items towards sciences in PISA 2015 student questionnaire including differential item functioning (DIF) in terms of language and culture. Generalized Mantel Haenszel (GMH), ordinal logistic regression (OLR) and poly-SIBTEST methods are applied for DIF analyzes. The GMHDIF program was used for the GMH method and the R software was used for the poly-SIBTEST and OLR methods. The study works with PISA 2015 dataset. This dataset encompasses 5089 students in Turkey, 5157 student in Ireland, 5000 students in USA and 4417 students in England. As a result of the analysis carried out on 9 items, DIF is determined in six items by using GMH and OLR methods and four items by using poly-SIBTEST method in the England-Ireland data. In the England-USA data, DIF is determined in eight items by GMH method, nine items by OLR method and three items by poly-SIBTEST method. In the England-Turkey data, DIF is determined eight items in the GMH method, nine items by OLR method and five items by poly-SIBTEST method. Based on three methods, five items in comparison with England-Ireland and England-USA, and also seven items in comparison England-Turkey are detected attitude items including DIF. As a result, it is found that there was accordance the number of DIF items determined by OLR, GMH and poly-SIBTEST methods. On the other hand there are differences in the levels of DIF according to OLR and poly-SIBTEST methods. Furthermore, when the items containing DIF are examined by OLR and poly-SIBTEST methods, it is observed that there was accordance between the groups with the advantages.

Keywords: PISA, differential item functioning, ordinal logistic regression, generalized Mantel Haenszel, poly-SIBTEST

Teşekkür

İlk olarak tezimin her aşamasında yanımda olan, iş disiplini örnek aldığım, tezimde sistematik bir şekilde adım adım ilerlememi sağlayan, güler yüzü ve samimiyeti ile desteğini hep hissettiğim değerli danışmanım Dr. Öğr. Üyesi Kübra ATALAY KABASAKAL'a,

Tezimin araştırma sürecinde birçok yabancı kaynağa ve programa erişimimi sağlayan, araştırmamın analiz bölümünde tıkanıp yerlerde tüm içtenliği ve yardımseverliği ile tezime katkıda bulunan değerli hocam Terry ACKERMAN'e,

Tez savunma jürimde yapıcı önerileri ile tezime destek olan değerli hocalarım Doç. Dr. Burcu Atar ve Doç. Dr. Dilara BAKAN KALAYCIOĞLU'na,

Tezimin sonlanmasında büyük katkıları bulunan değerli hocalarım Selahattin GELBAL ve Nuri DOĞAN'a,

Hayatımın en güzel ve en zor zamanlarında hep yanımda olan, ilk sevdiğim insanlar canım annem, babam ve en iyi arkadaşım ablama,

Bu sürece adım atmamı sağlayan, ilerlemem için bana her koşulu oluşturan ve her konuda destek olan, bu zaman diliminde daha fazla baba olmak zorunda kalan olan sevgili eşime,

Bu uzun süreç boyunca varlıkları ile güç aldığım zaman zaman ihmal ettiğim ilk göz ağrım canım oğlum Ömer Tuna'ya ve bana zamanı verimli kullanmayı öğreten, evimizin neşesi canım kızım Zeynep'e,

Teşekkürlerimi sunarım.

İçindekiler

Öz	ii
Abstract	iii
Teşekkür	iv
İçindekiler	v
Tablolar Dizini	viii
Şekiller Dizini	ix
Simgeler ve Kısaltmalar Dizini	x
Bölüm 1 Giriş	1
Problem Durumu	1
Araştırmanın Amacı ve Önemi	3
Araştırma Problemi.....	7
Sayıtlılar	8
Sınırlılıklar	8
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar	9
Yanlılık	9
Değişen Madde Fonksiyonu (DMF) Belirleme Yöntemleri	11
İlgili Araştırmalar.....	21
Bölüm 3 Yöntem	28
Araştırmanın Türü	28
Araştırmanın Evreni ve Örnekleme	28
Veri Toplama Süreci.....	30
Veri Toplama Araçları	30
Verilerin Analizi	30
Sayıtlıların İncelenmesi	35
Faktör Analizi	37
Betimsel İstatistikler.....	43

Bölüm 4 Bulgular ve Yorumlar	45
Alt Problem 1a'ya İlişkin Bulgular ve Yorumlar	45
Alt Problem 1 b'ye İlişkin Bulgular ve Yorumlar	46
Alt Problem 1c'ye İlişkin Bulgular ve Yorumlar	47
Alt Problem 1d'ye İlişkin Bulgular ve Yorumlar	48
Alt Problem 1e'ye İlişkin Bulgular ve Yorumlar	49
Alt Problem 2a'ya İlişkin Bulgular ve Yorumlar	49
Alt Problem 2b'ye İlişkin Bulgular ve Yorumlar	50
Alt Problem 2c'ye İlişkin Bulgular ve Yorumlar	51
Alt Problem 2e'ye İlişkin Bulgular ve Yorumlar	53
Alt Problem 3a'ya İlişkin Bulgular ve Yorumlar	54
Alt Problem 3b'ye İlişkin Bulgular ve Yorumlar	55
Alt Problem 3c'ye İlişkin Bulgular ve Yorumlar	56
Alt Problem 3d'ye İlişkin Bulgular ve Yorumlar	56
Alt Problem 3e'ye İlişkin Bulgular ve Yorumlar	58
Bölüm 5 Sonuç, Tartışma ve Öneriler	59
Sonuçlar	59
Tartışma	60
Öneriler	61
Kaynaklar	64
EK-A: PISA 2015 Öğrenci Anketinden Seçilen Tutum Maddeleri	76
EK-B: OLR ve poly SIBTEST Yöntemleri İçin R Programında Yazılan Kodlar	78
EK-C: OLR Yöntemine Göre Yapılan Analiz Çıktıları	79
EK-Ç: Poly SIBTEST Yöntemine Göre Yapılan Analiz Çıktıları	87
EK-D: GMH Yöntemine Göre Yapılan Analiz Çıktıları	91
EK-E: Etik Kurul İzin Muafiyet Formu	111
EK-F: Etik Beyanı	112

EK-G: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu.....	113
EK-H: Thesis/Dissertation Originality Report.....	114
EK-I: Yayımlama ve Fikri Mülkiyet Hakları Beyanı	115

Tablolar Dizini

Tablo 1 DMF belirleme yöntemlerinin sınıflandırılması	11
Tablo 2 GMH yöntemi olasılık tablosu.....	15
Tablo 3 Zumbo ve Thomas tarafından ΔR^2 değerlerinin yorumlanması.....	18
Tablo 4 Gierl ve Jodoin (2001) tarafından ΔR^2 değerlerinin yorumlanması	18
Tablo 5 Rousses ve Stout (1996) tarafından yorumlanan $ \beta $ değer ölçütleri	20
Tablo 6 Cinsiyete ve ülkelere göre Frekans ve yüzde dağılımları	28
Tablo 7 Ülkeler bazında maddelerdeki ortalama değerler.....	29
Tablo 8 ST 94 ve ST113 kodlu öğrenci anketine ait maddeler.....	31
Tablo 9 Ülkelere göre KMO ve Bartlett değerleri.....	38
Tablo 10 Ülkelere ait açıklanan özdeğerler ve toplam varyans yüzdeleri.....	39
Tablo 11 Ülkeler bazında maddelere ait faktör yük değerleri	40
Tablo 12 Ülkelere ait Cronbach Alpha (α) güvenilirlik katsayıları.....	43
Tablo 13 Ülkelere göre betimsel istatistik analizi.....	44
Tablo 14 İngiltere-İrlanda verisine ait OLR analiz sonuçları	45
Tablo 15 İngiltere-İrlanda verisine ait poly-SIBTEST analizi sonuçları.....	46
Tablo 16 İngiltere-İrlanda verisine ait GMH analizi sonuçları	47
Tablo 17 İngiltere-İrlanda verisinde OLR, poly-SIBTEST ve GMH yöntemlerinin karşılaştırılması	48
Tablo 18 İngiltere-ABD verisine ait OLR analiz sonuçları.....	50
Tablo 19 İngiltere-ABD verisine ait poly-SIBTEST analizi sonuçları	51
Tablo 20 İngiltere-ABD verisine ait GMH analizi sonuçları.....	52
Tablo 21 İngiltere-ABD verisinde OLR, poly-SIBTEST ve GMH yöntemlerinin karşılaştırılması	53
Tablo 22 İngiltere-Türkiye verisine ait OLR analiz sonuçları	54
Tablo 23 İngiltere-Türkiye verisine ait poly-SIBTEST analizi sonuçları	55
Tablo 24 İngiltere-Türkiye verisine ait GMH analizi sonuçları	56
Tablo 25 İngiltere-Türkiye verisinde OLR, poly-SIBTEST ve GMH yöntemlerinin karşılaştırılması	57

Şekiller Dizini

Şekil 1. Dört ülke için yamaç birikinti grafikleri.....	42
---	----

Simgeler ve Kısaltmalar Dizini

- ABD:** Amerika Birleşik Devletleri
DMF: Değişen Madde Fonksiyonu
GMH: Genelleştirilmiş Mantel-Haenszel
KTK: Klasik Test Kuramı
LDFA: Lojistik Diskriminant Fonksiyon Analizi
LR: Lojistik Regresyon
KPM: Kısmi Puan Modeli
MH: Mantel Haenszel
MTK: Madde Tepki Kuramı
OECD: Ekonomik İşbirliği ve Kalkınma Örgütü
OLR: Ordinal Lojistik Regresyon
PIRLS: Uluslararası Okuma Becerileri Gelişim Projesi
PISA: Uluslararası Öğrenci Değerlendirme Programı
SD: Standartlaştırma Yöntemi
SMD: Standart Ortalama Farkları
TIMMS: Uluslararası Matematik ve Fen Eğilimleri Araştırması

Bölüm 1

Giriş

Araştırmanın bu bölümünde problem durumu, araştırma problemi, araştırmanın amacı ve önemi, sayılılar, sınırlılıklar ve tanımlara yer verilmiştir.

Problem Durumu

Günümüzde üzerinde durulması gereken önemli konulardan biri de kaliteli eğitimidir. Kaliteli bir eğitim için, öncelikle eşitlikçi bir eğitim sisteminin var olması gerekir. Bu bağlamda eğitimin cinsiyet, dil, ırk, bölge farkı gözetmeksizin toplum bütününe ayrımsız şekilde hitap edebilir nitelikte olması beklenir. Toplumun her kesimine eşit eğitim imkânı sunmak devletin asli görevleri arasındadır. Sosyal devlet ilkesi gereği her vatandaş eşit şartlar altında aynı kalitede ve nitelikte eğitim almalıdır. Öngörülen eğitim vatandaşlara aynı imkanları sunmadığı takdirde fırsat eşitsizliği ortaya çıkmaktadır. Bu eşitsizlikler toplumlarda farklı nedenlerden birçok kaynaklanmaktadır. Türkiye’de coğrafi, ekonomik ve toplumsal yapıdan kaynaklanan birçok sorun eğitimde fırsat eşitsizliğine neden olmaktadır. Coğrafi açıdan, kırsal alandaki okulların yetersizliği; ekonomik açıdan ailenin gelir düzeyinin düşük olması, devletin istihdam ve imkan sağlama konusundaki yetersizliği; toplumsal açıdan ise cinsiyet, ailenin eğitim durumu ve ailenin sunduğu imkânların yetersizliği eğitimde fırsat eşitliğine gölge düşürmektedir (Cleary, 1968).

Birleşmiş Milletler Çocuklara Yardım Fonu (UNICEF, 2003) yaptığı araştırmada Türkiye’de eğitimi toplumsal açıdan ele alarak, okula erişim ve okula devam konusunda kız öğrencilerin engellendiği ve erkek öğrencilere göre dezavantajlı olduğunu belirterek eşit şartlar altında eğitim alamadıklarına vurgulamaktadır. Bu incelemeye göre, eğitimin her kademesinde kadın-erkek arasında ayrımcılığa rastlandığı ve kız çocuklarının eğitim hizmetlerinden gerekli şekilde yararlanamadığı sonucuna ulaşılmıştır. Bu cinsiyet farkının en fazla Doğu ve Güneydoğu’nun kırsal kesimlerinde ve şehir merkezlerinin gecekondu yapılaşmasının yoğun olduğu bölgelerde rastlanmakta olduğu belirtilmektedir. Ancak, alan yazın incelendiğinde akademik başarıyı etkileyen faktörler içerisinde cinsiyet, eğitim durumları ve öğrenci tutumlarının yanı sıra sosyoekonomik ve sosyokültürel değişkenlerin daha etkin rol aldığı görülmektedir (Dinçer ve Kolaşın, 2009). Yıldırım (2012) araştırmasında, Türkiye’de eğitim kalitesinin yordayıcıları arasında ailenin sosyoekonomik durumu,

eđitim dzeyi ve yařadığı blge olduđu belirtilmektedir. Sosyoekonomik yetersizliklere ek olarak, hedeflenen temel eđitim becerilerinin toplumun tm kesimine eřit derecede kazandırılmaması, okullar arasında kalite farkının olması ve đrenim hayatının tamamlanamaması eđitim alanında ciddi eřitsizliklere neden olmaktadır (Polat, 2009).

zetle, alan yazın alıřmalardan da anlařılacađı zere eđitimde fırsat eřitsizliđi ok farklı nedenlere dayanmaktadır. Sorunun ozm adına tm bu nedenlerin sistematik alıřmalarla dzenli bir řekilde ortaya konulması gerekmektedir. Bu amala, eđitimde fırsat eřitliđine engel olan durumları nlemek adına ncelikle eřitsizliđe sebep olan unsurların kaynađının tespit edilmesi gerekmektedir.

Eđitimde eřitsizliđi ve aksayan noktaları belirleyebilmek iin ncelikle lkelerin eđitim durumları ve dzeyleri, eđitimdeki aksaklıkların nedenleri ve yrrlkteki mfredat programının etkinliđi hakkında n bilgi edinmek gerekmektedir. Bu dođrultuda, dnya apında geerliđi olan TIMSS (Uluslararası Matematik ve Fen Eđilimleri Arařtırması), PISA (Uluslararası đrenci Deđerlendirme Programı) ve PIRLS (Uluslararası Okuma Becerileri Geliřim Projesi) gibi uluslararası bir veri setine ihtiya duyulmaktadır (Ercikan ve Koh, 2005; Ersoy, 2007). PISA, 2000 yılından itibaren Ekonomik İř Birliđi ve Kalkınma Teřkilatı (OECD) tarafından dnyanın en kapsamlı eđitim arařtırması niteliđinde olan bir alıřmadır. Toplam 72 lkeden 15 yař grubu đrencilerin katıldığı bu alıřmada, katılımcıların matematik, fen ve okuma beceri seviyelerini lmektedir (OECD, 2016). Aynı zamanda, bu deđerlendirme đrenci, veli, okul anketleri kullanılarak đrenmeye etki eden đrenme ortamları, okul trleri, đrencilerin sosyo-ekonomik durumları hakkında bilgi toplamaktadır.

PISA 2015 arařtırmasında ilk defa bilgisayar tabanlı deđerlendirme kullanılmıřtır. Katılımcı 72 lkeden; 57'si 2015 uygulamasını bilgisayar tabanlı deđerlendirirken, 15'i kâđıt-kalem tabanlı deđerlendirme olarak geerleřtirmiřtir. Yapılan pilot uygulamadan elde edilen veriler iřıđında kullanılan ynteme gre anlamlı bir farklılık gstermediđi tespit edilmiřtir (OECD, 2016). Fakat lkeler zeline her deđerřkeni incelemek mmkn olmamakla birlikte lkeler arası dil, kltr, cinsiyet farklılıklarından etkilenip etkilenmediđi tartıřma konusu oluřturmakta ve her geen gn yeni arařtırmaların gerekliliđini ortaya koymaktadır.

Her  yılda bir yapılan PISA arařtırması, temel olarak fen, matematik ve okuma becerileri alanında đrenci durumlarını deđerlendirmektedir. Bu  yıllık PISA dngsnde temel alanlardan birine ađırlık verilmektedir. PISA 2015'te seilen temel

alan fen okuryazarlığıdır. Burada okuryazarlık ile anlatılmak istenen, öğrencilerin temel konu alanlarındaki çeşitli durumlarda karşılaşılan problemleri tanımlaması, yorumlaması, becerilerini kullanabilmesi, analiz edebilmesi ve analizlerine dayalı çıkarımlar yapabilmesidir (OECD, 2016). PISA fen okuryazarlığının yanı sıra öğrencilerin akademik başarılarını etkileyen duyuşsal özelliklere de yer verilmektedir. Bu amaçla, fen konularına ilgisi ve araçsal motivasyon başlığı altında, öğrencilerin fen derslerinin ileride eğitim hayatında ve kariyer planlarına katkı sağlayıp sağlamayacağı algısına ilişkin anket soruları sorulmuştur. Fen öğrenme motivasyonu başlığı altında öğrencilerin fen öğrenmeyi ne derece ilginç ve eğlenceli buldukları ile fen özyeterliliği başlığı altında ise öğrencilerin kendi becerilerini gerçek yaşamda etkin bir şekilde kullanabilme ve karşılaşılan zorluklarla baş edebilme düzeyleri hakkında öğrencilere anket soruları yöneltilmiştir (OECD, 2016). Temel alanların yanında uygulanan bağlamsal anketler aracılığı ile öğrencilerin motivasyonu, kendileri hakkında tutumları, öğrenme süreçlerine dayalı psikolojik durumları, aileleri ve okul ortamlarına ilişkin bilgilerde sunulmaktadır (OECD, 2016).

Geleceğe yön vermede katkı sağlayacağı düşünölen PISA gibi uluslararası sınavların, güvenilirlik ilkesi geređi; ölçme sonuçlarının tesadüfi hatalardan arınık olması (Turgut ve Baykul, 2011) ve geçerlik ilkesi geređi; istenilen özelliđi ölçmesi ve bunu yaparken diđer özelliklerin etkilerini ölçölere yansıtmadan yapması beklenir (Özçelik, 2010). Yani ölçme işlemine karışan hataların en az olması ve ölçölmek istenen amaca hizmet etmesi iyi geliştirilmiş ölçme araçlarında bulunması gereken niteliklerdendir. Ne var ki sosyal bilimlerde olduđu gibi davranışsal özelliklerin ağır bastığı alanlarda hatanın karışma olasılığı diđer alanlara göre her zaman daha fazladır (Clauser ve Mazor, 1998). Hatanın katılımcıların yetenek düzeyinden mi, yoksa testin yapısından mı kaynakladığını saptamak için yanlılık çalışmalarının yapılması gerekir. Test içeriğinde yanlı maddelerin varlığı bozucu etkiye sahip olup bu maddelerden oluşan testin sonuçları ise tartışmaya açıktır.

Araştırmanın Amacı ve Önemi

Sınavların amaçları arasında durum tespiti, sıralama, kıyaslama, seviye belirleme, yerleřtirme veya karar verme vardır. PISA uygulamasında uluslararası alanda öлке kıyaslamalarına veya sıralamasını gözlemlemeye olanak verirken içerdii farklı tutum maddeleri ile sadece öğrencilerin temel bilgi ve becerilere ne ölçüde sahip

oldukları değil aynı zamanda öğrenci başarısı ve eğitim çıktılarına etkileyen faktörlerin de araştırılmasına kapı açmaktadır. Yalnız, sınavların bu denli geniş kitlelere hitap etmesi birtakım sorunlara yol açabilir. PISA gibi uluslararası sınavların katılımcıları farklı kültürlere mensup olduklarından, cevaplayıcıların ait oldukları kültüre göre soruları algılayış biçimleri ve buna bağlı olarak yanıtlamalarında farklılıklar olabileceği göz ardı edilmemelidir. Yapılan sınavlarda, gruplar arasında gerçekleşen bu farklılıkların testin yapısından mı yoksa bireysel farklılıklardan mı kaynaklandığının araştırılması yanlılık çalışmaları ile mümkündür.

Yanlılık, ölçmeye gelişi güzel karışan tesadüfi hatalardan değil, ölçme işlemine veya koşullarına bağlı olarak karışan sistematik hatalardan kaynaklanmaktadır (Reynolds, Livingston ve Willson, 2009). Buna göre bir maddenin, ölçmeyi amaçladığı değişken ile ilgisiz bir özellik nedeniyle bireyler arasında belli bir gruba avantaj veya dezavantaj sağlaması durumunda söz konusu maddenin yanlı olduğu söylenebilir. Madde yanlılığı olarak adlandırılan bu durum, test koşullarından veya maddenin özelliklerinden dolayı bir grubun doğru yanıt olma olasılığının diğer gruba göre daha fazla veya az olmasıdır (Zumbo, 1999). Diğer bir ifade ile benzer yetenek düzeyinde sahip olan bireyler farklı alt grupta bulunmaları nedeniyle cevaplarının değişebileceği sonucu çıkmaktadır. Ölçme sonuçlarına hata karışmasına neden olan bu durum aynı zamanda testin geçerliğini de zedelemektedir (Osterlind, 1983; Zumbo, 1999). Yanlılık çalışmaları için öncelikle Değişen Madde Fonksiyonu (DMF) gösteren maddelerin gerekli istatistiksel analizler ile belirlenmesi gerekir. DMF, benzer yetenek düzeyine sahip fakat farklı grupta yer alan cevaplayıcıların maddeyi yanıt olma olasılıklarının farklılaşması olarak tanımlanabilir. DMF birçok nedenden kaynaklanabilir. Bunlardan biri madde etkisidir (Camilli ve Shepard, 1994). Madde etkisi, farklı gruplardaki cevaplayıcıların yetenekleri arasında gerçek farklılıklarının olması durumudur (Zumbo, 2007). Eğer DMF, grupların yetenek düzeyleri arasındaki farklılıktan değil de madde yapısından kaynaklanıyorsa bu durumda yanlılıktan söz edilir. DMF, madde etkisi ve madde yanlılığını belirlemek için bir ön koşuldur. Bu nedenle DMF içerdiği tespit edilen maddelerin gerekli içerik analizi incelemeleri sonucunda maddenin yanlı olup olmadığına karar verilir.

DMF içeren maddelerin varlığı ölçme araçlarının geçerliğini tehdit eden bir unsurdur. Bu durum geleceğe yön vermesi planlanan sınavların adil olmasına gölge düşürmektedir. DMF araştırmalarının yapılması, geçerliği ve güvenilirliği daha yüksek

sınavlar hazırlanabilmesi bakımından önemlidir. Diğer taraftan, PISA gibi belli aralıklarla tekrarlanan uygulamalarda bilgilerin güncelliğini sağlamak adına uygulamaların yakından takip edilmesi ve düzenli DMF analizlerinin yapılması gerekmektedir. Bu araştırmada PISA uygulamasındaki tutum maddelerinin DMF içeren maddelerin belirlenmesi ve olası nedenlerinin belirlenmesi açısından önemlidir.

Ulusal ve uluslararası alanlarda yapılan performans değerlendirmelerinde DMF çalışmalarına sıklıkla rastlanmaktadır. Bireyler hakkında önemli kararların alındığı bu tarz kültürler arası sınavların ölçme ve değerlendirme işlemlerinin yansız olması beklenir. Fakat alan yazına bakıldığında gruplar arasında farklılaşmalar arttıkça DMF gösteren madde sayılarının da arttığı sonucuna varılmıştır (Gök, Atalay Kabasakal ve Kelecioğlu, 2014; Uzun ve Gelbal, 2017). Çalışmaların yanlılık nedenlerine bakıldığında bu farklılaşmanın nedeni olarak dil, kültür, cinsiyet gibi etmenler gösterilmektedir. Bu bağlamda, bu araştırmada dilsel ve kültürel açıdan çeşitlilik olabileceği düşünülen sınavlar arasından PISA uygulaması seçilmiştir. PISA 2015 fen bilimlerine yönelik tutum maddeleri karşılaştırılırken, dil ve kültür bakımından benzerlik ve farklılıklar gösterebilecek ülkeler tercih edilmiştir.

Bu çalışmanın amacı, PISA 2015 öğrenci anketlerinde yer alan tutum maddelerinin, Türkiye, İrlanda, İngiltere ve Amerika Birleşik Devletleri (ABD) ülkelerine göre DMF gösterip göstermediğini poly-SIBTEST, ordinal lojistik regresyon (OLR) ve genelleştirilmiş Mantel Haenszel (GMH) yöntemlerini kullanarak tespit etmektir. Ülke seçimlerinde; aynı dil benzer kültür (İngiltere-İrlanda), aynı dil farklı kültür (İngiltere-ABD) ve farklı dil ile farklı kültür (Türkiye-İngiltere) yapılarıyla birbirine yakın ve birbirinden bağımsız kültür örneklemi seçilmiştir.

Alan yazın taraması yapıldığında iki kategorili maddeler için geliştirilen modeller üzerine çok fazla araştırma (Abedalaziz, 2010; Akın Arıkan, Uğurlu ve Atar, 2016; Butakor, 2015; Büyükköse, 2018; Cuevas ve Cervantes, 2012; Demir, 2013; Demir ve Köse, 2014; Demirus, 2015; Gök, Kelecioğlu ve Doğan, 2010; Koyuncu, Aksu ve Kelecioğlu, 2018; Terzi ve Yakar, 2018; Yıldırım, 2008) bulunurken, çok kategorili verilerde ulusal yazında çok sayıda araştırma (Atalay, 2010; Başokçu ve Öğretmen, 2013; Gök Atalay, Kelecioğlu, 2014; Salı ve Sırgancı, 2018; Uyar ve Uyanık, 2016) bulunmamaktadır. Bunun nedenlerinden birisi çok kategorili maddeler için geliştirilen modellerin daha karmaşık yapıda olması gösterilmektedir (Maydeu-Olivares, Drasgow, ve Mead, 1996).

Bu araştırma için çok kategorili verilerde farklı yöntem ve programlar kullanılarak yöntemler arasında uyumlara bakılmıştır. Araştırmada kullanılan GMH, OLR ve poly-SIBTEST yöntemlerine ilişkin alan yazın incelendiğinde ayrı ayrı çok sayıda araştırmaya (Arıkan, 2015; Atalay, 2010; Atar ve Kamata, 2011; Büyükköse, 2018; Paul K. Crane, Gibbons, Jolley, ve Belle, 2006; Kristjansson, Aylesworth, Mcdowell, ve Zumbo, 2005; Özberk ve Koç, 2017; Sandilands, 2008; Wang ve Su, 2004) rastlamak mümkün iken her üç yöntemin bir arada kullanıldığı DMF çalışmalarının azlığı bu alanla ilgili yapılacak çalışmalara yardımcı olması açısından önem taşımaktadır.

Yapılan çalışmalarda DMF belirlemede kullanılan yöntemlerin birbirine göre üstünlük ve sınırlılıkları bulunmaktadır. Her yöntemin dayandığı istatistiksel modeller veya varsayımlar açısından diğer yöntemlere göre avantajları veya dezavantajları olabilmektedir. Bu amaçla bir yönteme göre yapılan çalışmalar yerine birden çok yönteme dayalı çalışmalar yapılması önerilmektedir (Holland ve Wainer, 1993; Osterlind ve Everson, 2009). Bu bağlamda bu çalışmada DMF belirlemede üç farklı yöntem (GMH, OLR, poly-SIBTEST) kullanılarak analiz sonuçları karşılaştırılmıştır. Bu araştırma için seçilen yöntemlerin avantajları ise şu şekilde özetlenebilir. Bazı araştırmacılar OLR yöntemini, tek biçimli ve tek biçimli olmayan DMF'yi belirleyebilmesi kullanım ve yorumlama açısından daha uygulanabilir bir yöntem olması nedeniyle istatistiksel analizlerde kullanımını tavsiye etmektedir (Clauser ve Mazor, 1998; Gök ve diğerleri, 2010; Kaya, Leite, ve Miller, 2015; Swaminathan ve Rogers, 1990). Ayrıca OLR yönteminin DMF'nin doğasına ilişkin detaylı bilgi vermesi ve tek biçimli olmayan DMF'yi belirlemede diğer yöntemlere göre daha güçlü olması bakımından bu çalışmada tercih edilmiştir. GMH yöntemi ise madde ortalamaları yerine maddeye verilen tüm cevapların dağılımına göre ilişki belirleyebilmektedir. Bu nedenle daha karmaşık eşleşme modeli oluşturup daha gerçekçi sonuçlar vermesi amacıyla analize dahil edilmiştir. Diğer bir yöntem olan SIBTEST yöntemi ise, gruptan bağımsız olma özelliği nedeniyle daha az hatalı bilgiler sunması bu çalışmada tercih nedenlerinden biri olmuştur (Kelecioğlu, Karabay ve Karabay, 2014). Bununla birlikte, poly-SIBTEST yönteminin en fazla DMF belirleyen yöntem olduğu ve bu DMF'li maddelerin diğer yöntemler ile uyumluluk gösterdiği yapılan araştırmalarla desteklenmektedir (Henderson, 2001). Yapılan bir çalışmada, poly-SIBTEST ve GMH yöntemlerinin diğer yöntemlere göre daha hassas sonuçlar verdiği gözlenmiştir

(Mellor, 1995). Aynı zamanda GMH yönteminin daha karmaşık yapılarda iyi çalışabilmesi ve diğer yöntemlere göre 1. tip hata kontrolünü daha iyi sağlayabilmesi (Penfield, 2001) bakımından bu araştırma için tercih edilmiştir.

Araştırma Problemi

PISA 2015 öğrenci anketinde yer alan tutum maddeleri dile ve kültüre göre değişen madde fonksiyonu içermekte midir?

Alt problemler.

1. PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-İrlanda uygulamasında
 - a. OLR yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - b. Poly-SIBTEST ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - c. GMH ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - d. OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizlerin sonuçları birbiriyle uyumlu mudur?
 - e. OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucunda her üç yönetime göre DMF gösteren maddeler nelerdir?
2. PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-ABD uygulamasında
 - a. OLR yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - b. Poly-SIBTEST ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - c. GMH ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - d. OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizlerin sonuçları birbiriyle uyumlu mudur?

- e. OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucunda her üç yönleme göre DMF gösteren maddeler nelerdir?
3. PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-Türkiye uygulamasında
 - a. OLR yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - b. Poly-SIBTEST ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - c. GMH ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
 - d. OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizlerin sonuçları birbiriyle uyumlu mudur?
 - e. OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucunda her üç yönleme göre DMF gösteren maddeler nelerdir?

Sayıtlılar

1. PISA 2015 uygulamasına katılan öğrencilerin, test maddelerine içtenlikle ve duyarlı bir şekilde cevap verdikleri
2. PISA 2015 uygulamaya katılan öğrencilerin test maddelerini kendi bilgi ve yetenek düzeyleri doğrultusunda kopya çekmeden yanıtladıkları
3. PISA 2015 uygulamasına katılan öğrencilere uygulanan sınavlar aynı koşullar altında gerçekleştirildiği varsayılmıştır.

Sınırlılıklar

1. Bu araştırmada, PISA 2015 uygulamasında yer alan öğrenci anketinde tutum maddelerine (özyeterlik ve araçsal motivasyon) verilen cevaplar ile sınırlıdır.
2. Bu çalışmadaki karşılaştırmalar PISA 2015 uygulamasına katılan ABD, Türkiye, İrlanda ve İngiltere ülkelerinin örneklemi ile sınırlıdır.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Yanlılık

Yanlılık kavramı ilk olarak, zekâ testleri çalışmalarında ortaya çıkmıştır. Tarihsel gelişimine bakıldığında, 1900'lü yılların başında Alfred Binet'in düşük ekonomi düzeyine sahip çocuklar üzerinde uygulamış olduğu çalışmaya rastlanmaktadır. Binet bu araştırmada bazı test maddelerine verilen cevaplarda, çocukların zihinsel kapasitelerinden ziyade kültürel eğilimlerinin etkisi altında olabileceği sonucuna varmıştır (Camilli ve Shepard, 1994). 1916 yılında Terman tarafından geliştirilen Stanford-Binet zeka testinde bazı maddelere verilen cevaplarda cinsiyete ve ülkenin belli bölgesinde yaşayan insanlara göre değiştiğini bazı maddelerin ise yaş farkı gözetmeksizin hemen herkes tarafından cevaplandığını gözlemlemiştir (Camilli ve Shepard, 1994). Stern (1914) ise zekâyı kronolojik yaşa göre tanımlayıp, karşılaşılan farklılıkların kaynağını belirlemek için, kullanılan testlerin tek bir gruba dönük hazırlanmasını önermiştir (Camilli ve Shepard, 1994). Bu bağlamda, farklı yıllarda yapılan araştırmalara bakıldığında zekâyı tanımlarken zekâ üzerindeki değişkenlerin varlığına rastlanmış hatta bunun yaş, cinsiyet, kültür ve ülke gibi faktörlerin etkisi altında olabileceği çıkarımı yapılmıştır. Zekâ testlerinin sonuçlarının ırka, dile ve kültüre göre farklılık gösterdiğinin fark edilmesi 1950 yıllarında yanlılık çalışmalarının önemini artmasına neden olmuştur. Gelineen durumda belirlenen değişkenlere ek olarak Eells ve diğerleri (1951), öğrencilerin zekâlarındaki farklılıkların nedeninin sadece bireysel farklılıklardan değil ayrıca test maddelerinin yapısından kaynaklandığını ve bu yüzden öğrencilerin yeteneklerini tam olarak yansıtamadığını düşünerek aslında yanlılıktan bahsetmiştir (Camilli ve Shepard, 1994). 1960'lı yılların sonu ve 70'li yıllarda ise, özel eğitim gereksinimleri, vatandaşlık hakları, eğitim eşitliği gibi kavramların ortaya çıkması ile eğitim alanında seçme işlemlerinde ve işe alımlarda bireyler hakkında verilen kararlar için uygulanan zekâ testleri sorgulanmaya başlanmıştır. Çünkü o döneme kadar eleman seçmek için Griggs Şirketi tarafından uygulanan testlerin artık yetersiz kaldığı gözlenmiştir (Camilli ve Shepard, 1994). Bu sebeple, değişen ve gelişen toplumlara ayak uydurabilecek geçerliği yüksek sınavlara ve bunları objektif şekilde değerlendirebilmek için de daha kapsamlı analizlere ihtiyaç duyulmuştur. Bu gereksinimler sebebiyle yapılan çalışmalar neticesinde 70'li yılların sonlarında ilk kez madde yanlılığı kavramı ortaya atılmış ve 80'li yıllarda testteki yanlı

maddelerin ölçme işlemi için bir tehdit oluşturduğu varsayılarak bu çalışmaların önemi ve gerekliliği savunulmuştur (Camilli ve Shepard, 1994; Kristjansson, Aylesworth, Mcdowell ve Zumbo, 2005).

Yanlılığın tarihsel sürecine bakıldığında, günümüze kadar önemini yitirmeden araştırma konusu olmuştur. Her bir araştırmacı da madde yanlılığını farklı açılardan ele alıp tanımlamıştır. Cleary (1968) çalışmasında, regresyon doğrusu üzerinden yordanan ölçüt puanlarının alt örneklemelerde çok düşük veya çok yüksek düzeyde olduğunu fark etmiştir. Bu tanım yanlılık kavramının ortaya çıktığı ilk çalışmalardandır (Lee, 2003). Madde yanlılığını test geçerliği içinde ele alan araştırmacılardan; Kelderman (1989), yanlılığın farklı gruplardan gelen aynı yetenek düzeyine sahip bireylerin, doğru cevabı farklı yanıtlama olasılığı olarak açıklamaktadır. Hambleton ve Rogers (1989) madde yanlılığını, farklı alt gruplardan gelen eşit yetenek düzeyindeki bireylerin, bir maddedeki farklı performansları olarak tanımlarlar. Diğer bir ifadeyle; maddelerin farklı alt gruptaki bireyler tarafından farklı zorlukta algılanması olarak görülmektedir. Madde yanlılığı kavramını ölçme hataları kapsamında ele alan araştırmacılar ise madde yanlılığını, test maddelerinin gözlenen değerleri ile gerçek değerleri arasındaki farklılık yani sistematik hata içermesi olarak tanımlarlar (Camilli ve Shepard, 1994). Bu sistematik hatanın ölçme sonuçlarına karıştığı ve testin yapı ve yordama geçerliliğini olumsuz yönde etkilediğini belirtilmiştir (Childs, 1990). Yanlılık, bir gruba avantaj veya dezavantaj sağlamasına yol açan sistematik hataların varlığı durumu olarak da tanımlanabilir (Cromwell, 2002; Reynolds ve diğerleri 2009).

Ölçme araçlarında bulunan maddeler tüm cevaplayıcılara kendi yetenek ve bilgilerini aktaracak şekilde eşit imkân sağlamalıdır. Bununla birlikte, bir grubun maddeyi doğru cevaplama olasılığı diğerlerine göre manidar bir fark göstermemelidir. Söz konusu durumda bahsedildiği gibi farklı bir üst/alt grup oluşmasına neden olan yanlı maddeleri belirleme çalışmaları, DMF analizleri ile gerçekleştirilmektedir. Ancak, bir maddenin yanlı olması maddenin DMF içerdiğini gösterirken, her DMF gösteren madde için kesin yanlıdır demek mümkün değildir (Kamata ve Vaughn, 2004).

Madde yanlılığı çalışmalarında, DMF analizi tek başına yeterli değildir. Ayrıca uzmanlar tarafından da yorumlanması gereken yargısal süreci de kapsamaktadır (Clauser ve Mazor, 1998). Zumbo (1999) bu görüşü destekler nitelikte, istatistiksel süreçlerin madde yanlılığını belirlemede yeterli olmadığını, ayrıca psikolojik ve eğitime dayanan bulgularla desteklenmesi gerektiğini savunmaktadır. Dolayısıyla DMF içeren

maddelerin mantıksal analizi sonucunda yanlı olup olmadığı kanısına varılabilir. Fakat, gruplar arası farklılaşma başka nedenlerden dolayı kaynaklanabileceği gibi her DMF içeren maddenin yanlı olduğu anlamına gelmemektedir (Camilli ve Shepard, 1994; Zumbo, 1999).

Değişen Madde Fonksiyonu (DMF) Belirleme Yöntemleri

DMF, aynı yetenek düzeyine sahip fakat farklı gruplara ait bireylerin bir maddeyi doğru cevaplama ihtimallerinin farklı olması olarak tanımlanabilir. Diğer bir deyişle, aynı yetenek düzeyindeki iki bireyin doğru cevap verme olasılığının aynı olmaması durumu o maddenin DMF'ye sahip olduğunun göstergesidir (Hambleton ve Rogers, 1989). Tanımdan da anlaşılacağı üzere, DMF analizleri her bir maddenin aynı örneklemden elde edilen farklı alt gruplarda aynı şekilde işleyip işlemediği sorununu inceler.

DMF'yi belirlemek için çok sayıda yöntem geliştirilmiştir. Bu yöntemler için birçok sınıflandırma mevcuttur. Bu sınıflandırmaların dört tanesine Tablo 1'de yer verilmiştir.

Tablo 1

DMF belirleme yöntemlerinin sınıflandırılması

Camilli ve Shepard (1994)	Benito ve Navas-Ara (2000)	Ellis ve Raju (2003)	Zumbo (2009)
MTK	MTK	MTK	MTK
Varyans Analizi	KTK	KTK	Çok Boyutlu Yöntemler
Olasılık Tablosu	Faktör Analizi Ki Kare		Olasılık Tablosu

İki kategorili verilerde KTK'ya dayalı yöntemler içerisinde Mantel-Haenszel (MH), lojistik regresyon (LR), standartlaştırma yöntemi ve ki-kare yer alırken, MTK'ya dayalı yöntemler içerisinde Lord'un ki-karesi ve olabilirlik oranı yöntemleri yer almaktadır (Camilli ve Shepard, 1994; Ellis ve Raju, 2003; Holland ve Wainer, 1993). Diğer bir sınıflandırma ise örtük puan ile gözlenen puana dayalı olarak yapılmaktadır (Potenza ve Dorans, 1995). Burada iki kategorili verilerde örtük puana göre SIBTEST,

olabilirlik oranı ve maddenin ve testlerin farklı fonksiyonlaşması yöntemleri kullanılabilirken; gözlenen puana göre standartlaştırma, MH ve LR yöntemleri kullanılabilir (Potenza ve Dorans, 1995).

Çok kategorili veriler için kullanılan yöntemler iki kategorili veriler için uygulanan yöntemlerin bir uzantısı şeklindedir. Bu yöntemlerin çok kategorili veriler için uygulamasında ise;

- MH yönteminin uzantısı olarak Genelleştirilmiş Mantel-Haenszel (GMH) (Zwick, Donoghue, ve Grima, 1993)
- Standartlaştırma yönteminin uzantısı olarak Standart Ortalama Farkları (SMD) (Dorans ve Schmitt, 1991)
- Lojistik regresyon yönteminin uzantısı olarak Çok Kategorili Lojistik Regresyon (French ve Miller, 1996), Ordinal Lojistik Regresyon (Zumbo, 1999) ve Lojistik Diskriminant Fonksiyon Analizi (Miller ve Spray, 1994)
- SIBTEST'in uzantısı olarak poly-SIBTEST (Chang, Mazzeo ve Roussos, 1996) yöntemleri örnek verilebilir (Atar ve Kamata, 2011).

Potenza ve Dorans (1995) DMF inceleme çalışmalarını eşleştirme değişkeni olarak gözlenen puanların mı yoksa örtük değişkenin mi kullanıldığı ve modellerin parametrik bir teknik olup olmamasına göre ele almışlardır. DMF belirlemede çok kategorili veriler için parametrik (OLR, MTK) ve parametrik olmayan (GMH, SD, poly-SIBTEST) yöntemler olacak şekilde sınıflama yapmışlardır. Burada ifade edilen parametrik analizlerin yapılabilmesi için bazı varsayımların (verilerin normal dağılımı, deneklerin yeterli sayıda olması gibi) karşılanması beklenirken aynı durum parametrik olmayan analizlerden beklenemez. Parametrik olmayan yöntemler istatistiksel güç bakımından parametrik yöntemler kadar güçlü değilken parametrik yöntemlerin uygulanmadığı koşullarda rahatlıkla kullanılabilir. Diğer taraftan, parametrik yöntemlerin yanlış model tanımlanması hususunda sınırlılıkları vardır. Çünkü çok küçük uyumsuzluklar bile ciddi I. tip hata sorunu oluşturabilmektedir (Bolt, 2002). Buna karşın, parametrik olmayan yöntemler için aynı durum söz konusu değildir.

Tek Biçimli (TB) ve Tek Biçimli Olmayan (TBO) DMF. DMF tek biçimli (uniform) ve tek biçimli olmayan (nonuniform) şeklinde ikiye farklı şekilde görülebilir (Mellenbergh, 1982). Tek biçimli DMF, yetenek düzeyi ile grup arasında bir etkileşimin olmadığı durumlarda gözlenir. Diğer bir ifade ile, maddenin doğru cevaplanma olasılığı,

tüm yetenek düzeyleri içinde bir gruba avantaj sağlıyorsa tek biçimli DMF den söz edilebilir. DMF belirleme yöntemleri uygulanırken, bireyler odak ve referans grubu olmak üzere ikiye ayrılır. Daha sonra bu gruplardaki bireyler ölçülen yetenek düzeyine göre eşleştirilir. Ardından maddeler üzerindeki grup farklılıkları belirlenir. Tek biçimli DMF, bireyin yetenek düzeyi ile maddeyi doğru cevaplama olasılığının aynı düzeyde değişmesi durumudur.

Tek biçimli olmayan DMF ise, yetenek düzeyi ile grup arasında etkileşimin olduğu durumlarda rastlanmaktadır. Yani maddenin doğru cevaplanma olasılığı, farklı yetenek düzeylerinde farklı gruplara avantaj sağlıyorsa tek biçimli olmayan DMF'den söz edilir (Hambleton ve Rogers, 1989; Mellenbergh, 1982; Swaminathan ve Rogers, 1990). Bireylerin yetenek düzeyleri değiştikçe maddeyi doğru cevaplama olasılığı da değişmektedir.

Tek biçimli DMF'yi belirlemek için birçok yöntem mevcutken, tek biçimli olmayan DMF belirleme çalışmalarında tekniklerin birçoğu yetersiz kalmaktadır (Gierl ve Jodoin, 2001). Swaminathan ve Rogers (1990), DMF belirleme çalışmasında, MH ve LR yöntemlerinin tek biçimli DMF'yi belirlemede eşit güçte olduğunu ancak LR'nin tek biçimli olmayan DMF'yi belirlemede daha etkili olduğu sonucuna varmışlardır.

Çok kategorili verilerde DMF. Puanlama modelleri iki kategorili ve çok kategorili puanlama şeklinde ele alınmaktadır. Yanıtların iki kategorili puanlanması (Doğru-yanlış, çoktan seçmeli testler gibi) doğru yanıtlara 1, yanlış veya boş bırakılan yanıtlara 0 puan vererek yanıtlayıcıları sadece tek bir seçenek üzerinden değerlendirilmesi üzerinden olmaktadır. Bu puanlama modeli, kısmi bilgi ve yanlış bilgiyi ayırt etmeyi göz ardı eden bir durumdur. Üstelik, iki kategorili puanlamada cevaplayıcılar tam bilgi veya beklenen yetenek düzeyine sahip olmadığı sürece puan alamayacaklardır (Kurz, 1999; Wongwiwatthanakul, Popovich ve Bennett, 2000). Hatta, yanlış öğrenme veya şans faktörüyle elde edilen cevapların puanlamaya karışmasına fırsat verecektir. Bu sebeple yetenek düzeylerinin düzgün bir şekilde belirlenebilmesi için puanlamaya karışan şans, kısmi bilgi ve yanlış bilgi ayrımlarının iyi bir şekilde yapılması gerekmektedir (Tunç, 2017). Bu sorunların üstesinden gelebilmek için çok kategorili verilerde puanlama modelleri geliştirilmiştir. Özellikle sosyal ve davranış bilimleri alanlarında tutumların ölçülmesinde sıkça kullanılan likert tipi ölçeklerde, kısmi puanlama gerektiren maddelerde veya geniş kitlelerle hitap eden

başarı testlerinde çok kategorili yapılandırılmış maddelere de yer verilmesi çok kategorili puanlama yöntemlerinin gerekliliğini ortaya koymaktadır.

Bu araştırmada PISA 2015 tutum maddelerinde GMH, OLR ve poly-SIBTEST yöntemleri kullanılarak DMF analizleri yapılmıştır. Çok kategorili veriler için kullanılan bu yöntemler aşağıda açıklanmaktadır.

Genelleştirilmiş Mantel Haenszel (GMH). Mantel ve Haenszel (1959) tarafından bulunan MH yönteminde bireyler yeteneklerine göre odak ve referans grup olmak üzere kategorilere ayrılır. MH yöntemi ile tabakalara ayrılmış grupları karşılaştırabilmek için olasılık tablosundan yararlanılır (Zwick, Donoghue ve Grima, 1993). Holland ve Thayer (1988) çalışmalarında bir adım öteye giderek MH yöntemi ile elde edilen olasılık tablosunun DMF belirlemede kullanılabileceğini ileri sürmüşlerdir.

Yaygın olarak tek biçimli DMF belirlemede kullanılan MH yöntemi eşleştirilmiş gruplarda ki kare istatistiğine dayalı bir yöntemdir (Agresti, 1984). Olasılık oranlarının (α) elde edildiği MH yönteminde odak ve referans gruptaki performans farklılığının derecesi ölçülür. Analizin kolay yorumlanabilmesi için de logaritmik bir dönüşümden yararlanılarak delta katsayısı ($\Delta_{MH} = -2,35 \ln(\alpha_{MH})$) elde edilir. Elde edilen değer pozitif olması odak grubu lehine, negatif olması ise referans grubu lehine katkı sağladığını gösterir (Holland ve Wainer, 1993; Osterlind ve Everson, 2009; Swaminathan ve Rogers, 1990).

MH yönteminin bir uzantısı olarak Genelleştirilmiş Mantel Haenszel (GMH) yöntemi çok kategorili veriler için alternatif olarak geliştirilmiştir. MH yönteminde kategoriler sıralı şekilde ele alınırken, GMH yönteminde kategoriler sınıflamalı olacak şekilde ele alınır (Wang ve Su, 2004). Fakat bu yöntemlerin sınırlılıklarından birisi tek biçimli ve tek biçimli olmayan DMF'yi ayırt edememesidir (Wang ve Su, 2004).

MH yönteminde, GMH yönteminin aksine bireylerin verdikleri cevaplara göre kategorilendirmesini dikkate almak yerine iki grubun kendi cevap dağılımlarına göre bir eşleşme değişkenine bağlı olarak karşılaştırılmasını sağlar (Zwick, Donoghue ve Grima, 1993). Benzer bir ifade ile GMH yöntemi tüm maddelere verilen yanıtların dağılımını esas alırken, MH yöntemi sadece odak ve referans gruptan beklenen madde puanlarını esas almaktadır (Potenza ve Dorans, 1995; Wang ve Su, 2004).

Welch ve Hoover (1993) I. tip hata kontrolünde ve t istatistiğinde GMH'nin gücünü araştırmışlardır. Yapılan çalışmada kullanılan yöntemlerin I. tip hatanın kontrolünü iyi sağladığı fakat MH yönteminin GMH yöntemine göre daha güçlü olduğu gözlenmiştir. Zwick, Donoghue ve Grima (1993) ise Mantel ve GMH yöntemlerinin benzer şekilde performans gösterdiğini ve eşleştirme kriteri dahil edilmediği zaman her iki yöntemde tatmin edici sonuçlar verdiğini belirtmişlerdir. Chang, Mazzeo ve Roussos (1996) çalışmalarında Mantel, modifiye edilmiş SIBTEST (şimdiki adı ile poly-SIBTEST) ve standart ortama farkları yöntemlerini karşılaştırmışlar ve analiz neticesinde her üç yöntemden benzer sonuçlar elde etmişlerdir. Ankenmann, Witt ve Dunbar (1999) ise olasılık oranı ve Mantel yöntemlerinin gücünü MTK'ya dayalı incelemişlerdir. Mantel I. tip hata kontrolünü sağlarken DMF belirlemede olasılık oranı testinden daha güçlü olduğu kanısına varılmıştır. Fidalgo ve Bartram (2010) çalışmalarında Mantel testini hesaplamak için kullanılan puanlama sisteminin DMF'yi tespit etme gücünü etkilediği ve daha karmaşık yapıları tespit etmek için GMH'nin en iyi seçenek olabileceğini vurgulamışlardır. Bu bağlamda, bu araştırmada DMF belirlemek için parametrik olmayan yöntemlerden GMH yöntemi tercih edilmiştir.

Analizin yapılabilmesi için farklı gruptaki bireylerin verdikleri cevaplara göre, $2 \times T \times k$ büyüklüğünde üç boyutlu bir matris oluşturulur. Tablo 2'de olasılık tablosuna yer verilmiştir.

Tablo 2

GMH yöntemi olasılık tablosu

Grup	Maddeler					Toplam
	y_1	y_2	y_3	..	y_r	
Referans	n_{R1k}	n_{R2k}	n_{R3k}	...	n_{RTk}	n_{R+k}
Odak	n_{F1k}	n_{F2k}	n_{F3k}	...	n_{FTk}	n_{F+k}
Toplam	n_{+1k}	n_{+2k}	n_{+3k}	...	n_{+Tk}	n_{++k}

Tablo 2'de geçen T değeri çok kategorili verilerde maddedeki yanıt kategori sayısını, k değeri de eşleşen değişkenin düzeylerin sayısını temsil etmektedir. n_{RTk} ve n_{FTk} değerleri ise k düzeyinde y_T maddesini alan referans ve odak grup üyelerinin sayısını belirtmektedir. GMH analizine ait denklem (1) şu şekilde ifade edilebilir;

$$X^2_{GMH} = [\sum Ak - \sum E(Ak)]' [\sum V(Ak)]^{-1} [\sum Ak - \sum E(Ak)] \quad (1)$$

A_k : (T-1) x (T-1) matrisi

$E(A_k)$: (T-1) matris vektör uzunluğu

$V(A_k)$: (T-1) x (T-1) kovaryans matrisi

Gruplar ve cevap kategorisi arasında bir ilişkinin olmadığı varsayımına göre Denklem 1'deki test istatistiği T-1 serbestlik derecesinde ki kare dağılımına sahiptir. Eğer sıfır hipotezi (H_0) reddedilirse koşullu bir ilişkiye sahip olduğu ve böylece maddenin DMF gösterdiği söylenebilir (Wang ve Su, 2004).

Ordinal Lojistik Regresyon (OLR). Mantel-Haenszel (MH) yöntemi kullanımı kolay olmasına rağmen, tek biçimli ve tek biçimli olmayan DMF belirlemede yetersiz kalmaktadır. Bu amaçla, lojistik regresyon (LR) yöntemi ilk olarak Swaminathan ve Rogers (1990) tarafından MH yöntemine alternatif olarak geliştirilmiştir. Yapılan analizler sonucu, MH ve LR yöntemlerinin tek biçimli DMF'yi belirlemede eşit güçte olduğunu ancak LR'nin tek biçimli olmayan DMF'yi belirlemede daha etkili olduğu sonucuna varmışlardır (Swaminathan ve Rogers, 1990).

LR, DMF belirlemede iki kategorili veriler için iyi yapılandırılmıştır (Swaminathan ve Rogers, 1990; Zumbo, 1999). French ve Miller (1996) çok kategorili verilerde DMF belirlemek için çok kategorili lojistik regresyon yöntemini ileri sürmüşlerdir. Miller ve Spray (1994) lojistik regresyonun bir uzantısı olarak lojistik diskriminant fonksiyon analizi yöntemi ile DMF belirleme çalışmalarında bulunmuşlardır. Ayrıca bu yöntemin tek biçimli ve tek biçimli olmayan verilerde uygulanan diğer yöntemlere göre daha kolay uygulanabilir olduğu görüşü savunulmaktadır (Miller ve Spray, 1994). Zumbo (1999) çok kategorili veriler için etki büyüklüğünü incelemiştir. Ayrıca çalışmasında OLR yöntemiyle maddeye verilen cevapları bağımlı değişken, grup değişkeninin her birey için toplam puanla beraber grup toplam puan etkileşiminin bağımsız değişken olarak kullanmış ve ölçeğin tek boyutlu olduğu farz edilmiştir (Gelin ve Zumbo, 2003).

Bu çalışmada yürütülen OLR analizleri diğer çalışmalardan farklı olarak MTK'ya dayalı olarak gerçekleştirilmiştir. MTK ve yetenek puanlarına dayalı üç aşamalı modelden oluşan bu analizde, her madde için bir sabit terim modeli ile üç tane iç içe modelden oluşan ek bağımsız değişkenlerin etkileşimi aşağıdaki denklemler (2) ile ifade edilebilir (Choi, Gibbons ve Crane, 2011).

$$\text{Model 0 : } \text{logit}P(u_j \geq k) = \alpha_k \quad (2)$$

$$\text{Model 1 : } \text{logit} P (u_j \geq k) = \alpha_k + \beta_1 * \text{yetenek}$$

$$\text{Model 2 : } \text{logit} P (u_j \geq k) = \alpha_k + \beta_1 * \text{yetenek} + \beta_2 * \text{grup}$$

$$\text{Model 3 : } \text{logit} P (u_j \geq k) = \alpha_k + \beta_1 * \text{yetenek} + \beta_2 * \text{grup} + \beta_3 * \text{yetenek} * \text{grup}$$

Denkleimde geçen,

u_j : madde i için verilen cevapları,

α_k : kesişim terimini (kategori k 'ya bağlıdır),

$P (u_j \geq k)$: kategori k ve üstünde yanıt olasılığını temsil eder.

Ayrıca denkleimde geçen *yetenek*; gizil değişken veya gözlenen toplam puanı temsil eder. Model üç aşamadan oluşmaktadır (Denklemler 1-3). İlk aşamada madde performansını kestirebilmek için yalnızca yeteneği (Model 1), ikinci aşamada madde performansını kestirebilmek için yetenek ve grup değişkenini (Model 2), üçüncü aşamada ise yetenek, grup değişkeni ile grup ve yetenek etkileşim değişkenlerinin (Model 3) modellenmesiyle oluşur (Choi, Gibbons ve Crane, 2011).

LR analizlerinde tek biçimli ve tek biçimli olmayan DMF'li madde belirlemede ki kare değerinden yararlanılır (Swaminathan ve Rogers, 1990). Benzer şekilde bu analizde, TB DMF için Model 1 ve Model 2 log olasılık oran değerlerinin kıyaslanarak log olasılık oran değerleri arasındaki fark (1 serbestlik derecesinde, $sd=1$) ki kare değerleri ile karşılaştırılır. TBO DMF için Model 2 ve Model 3 log olasılık oran değerleri arasındaki fark (1 serbestlik derecesinde, $sd=1$) ki kare dağılımları ile karşılaştırılır.

DMF analizi ve DMF'nin belirlenmesinin ardından etki büyüklüğüne bakmak için R^2 değerleri yorumlanır (Kristjansson ve diğerleri, 2005). Zumbo ve Thomas (1996) tarafından önerilen ölçütler Tablo 3'te verilmiştir.

Tablo 3

Zumbo ve Thomas tarafından ΔR^2 değerlerinin yorumlanması

DMF Düzeyi	Değer	DMF Yorumu
A	$\Delta R^2 < 0,13$	DMF yoktur ya da ihmal edilebilir
B	$0,13 \leq \Delta R^2 < 0,26$	Orta düzeyde
C	$\Delta R^2 \geq 0,26$	Yüksek düzeyde

İstatiksel güç, örneklem büyüklüğünden etkilendiğinden (Cohen, 1988) popülasyon parametrelerinde yeterince büyük örneklem olduğunda istatistiksel olarak anlamlı olacaktır (Choi, Gibbons ve Crane 2011, Gierl ve Jodoin, 2001; Kim, Gibbons ve Crane, 2007; Zumbo, 1999). Bu bağlamda LORDIF analizi birden fazla etki büyüklüğü ölçümlerine yer vermektedir. Yapılan son araştırmalarda (Gierl ve Jodoin, 2001; Kim, Gibbons ve Crane, 2007), Zumbo (1999) tarafından önerilen ölçütlerin (Tablo 3) aralıklarının çok büyük olması nedeniyle DMF belirlemede yetersiz kaldığı kanısına varılmıştır. Bu bağlamda, Gierl ve Jodoin (2001) tarafından önerilen ölçütlere Tablo 3'te yer verilmiştir.

Tablo 4

Gierl ve Jodoin (2001) tarafından ΔR^2 değerlerinin yorumlanması

DMF Düzeyi	Değer	DMF Yorumu
A	$\Delta R^2 < 0,035$	DMF yoktur ya da ihmal edilebilir
B	$0,035 \leq \Delta R^2 < 0,070$	Orta düzeyde
C	$\Delta R^2 \geq 0,070$	Yüksek düzeyde

Tablo 4 incelendiğinde Gierl ve Jodoin (2001) tarafından önerilen R^2 ölçüt değerlerinin Zumbo ve Thomas (1996) tarafından önerilen R^2 ölçüt değerlerinden daha hassas olduğu görülmektedir. Cohen'e (1988) göre analiz sonucu elde edilen etki büyüklüğü sonuçları ise: 0,019 küçük; 0,130 orta; 0,260 ise yüksek etki değeri şeklinde yorumlanır. Seçilen yöntemlerin örneklem büyüklüğünden etkilenmesi, farklı

kategorilendirme sistemlerinden farklı sonuçlar elde edilmesi ve kullanılan ölçüt değerlerde ortak bir görüş sağlanamaması nedeniyle bu çalışma için daha hassas ölçüt değer aralıkları (Gierl ve Jodoin, 2001; Hidalgo ve López-Pina, 2004, Choi, Gibbons ve Crane, 2011) olarak Cohen ölçütleri tercih edilmiştir.

DMF tespit edilen her maddenin ölçekten çıkarılması yanlış bir karar olabilir (Gelin ve Zumbo, 2003). Zwick (1990) çalışmasında DMF içeren maddelerin testten çıkarılmasının yanlılığa sebep olduğunu göstermiştir. Yapılması gereken öncelikle analizler sonucu DMF'li olduğuna karar verilen maddeleri A (ihmal edilebilir düzeyde), B (orta düzeyde) ve C (yüksek düzeyde) şeklinde Tablo 4'teki gibi düzey aralığı belirlenmelidir. Daha sonra, hangi maddelerin düzeltilebilir hangilerinin üzerinde çalışılabilir olabileceğine karar verilmelidir.

İlgili alan yazında bahsedildiği gibi, OLR yönteminin ikili ve sıralı veriler için kullanılabilir olması, bağımlı değişkenin sürekli olma zorunluluğunun bulunmaması, etki büyüklüğünü hesaplayabilmesi ve uygulamanın kolay olması diğer yöntemlere göre üstünlük sağlamaktadır (Swaminathan ve Rogers, 1990; Zumbo, 1999). Bunun yanı sıra, OLR yönteminin yalnız DMF belirlemekle kalmayıp aynı zamanda DMF'nin doğasına ilişkin daha detaylı bilgi sağladığı ve tek biçimli olmayan DMF'yi belirlemede diğer yöntemlere daha güçlü olduğu düşünüldüğünden bu araştırma için tercih edilmiştir (Gök ve diğerleri, 2010).

Poly-SIBTEST. SIBTEST, iki kategorili veri setlerinde DMF belirlemek için Shealy ve Stout (1993) tarafından geliştirilmiş bir metottur. DMF belirlemede standartlaştırma yöntemi ile benzerlik gösterse de I. tip hatayı kontrol etmede regresyon temelli düzeltme kullanması yönünden farklılık göstermektedir (Clauser ve Mazor, 1998). SIBTEST yönteminin, yetenekleri karşılaştırmada gözlenen puanlar yerine gerçek puanlar kestiriminin kullanılması ve madde analizinin yapılması istatistiksel test sayısını azalttığı için istatistiksel gücü artırdığı düşünülmektedir (Abbott, 2007). Aynı zamanda küçük bir soru öbeği veya birkaç soru içeren test maddeleri üzerinden DMF çalışmaları yapılabilmesine olanak vermesi bakımından diğer yöntemlerden ayrılmaktadır (Clauser ve Mazor, 1998).

Poly-SIBTEST yöntemi ise SIBTEST yönteminin bir uzantısı olarak çok kategorili verilerde DMF belirlemek için geliştirilmiştir (Chang ve diğerleri, 1996). DMF miktarına ait denklem (3) şu şekildedir;

$$\beta = \frac{\hat{\beta}}{\hat{\sigma}(\hat{\beta})} \quad (3)$$

Denklemden geçen;

$$\hat{\sigma}(\hat{\beta}) = \left[\sum_{k=0}^{n_H} P_k^2 \left(\frac{\hat{\sigma}^2(Y|k, R)}{N_{Rk}} + \frac{\hat{\sigma}^2(Y|k, F)}{N_{Fk}} \right) \right]^{1/2}$$

$\hat{\sigma}^2(Y|k, R)$: referans grubun puan varyansı

$\hat{\sigma}^2(Y|k, F)$: odak grubun puan varyansı

Elde edilen sonuçları yorumlamak için Roussos ve Stout (1996) tarafından belirlenen β değerlerine ait ölçütler Tablo 5'te verilmiştir. Elde edilen β değerinin pozitif çıkması referans grubu lehine, negatif çıkması ise odak grubu lehine DMF gösterdiğini ifade eder (Holland ve Thayer, 1988; Holland ve Wainer, 1993; Osterlind ve Everson, 2009).

Tablo 5

Roussos ve Stout (1996) tarafından yorumlanan $|\beta|$ değer ölçütleri

DMF Düzeyi	Değer	DMF Yorumu
A	$ \beta < 0,059$	DMF yoktur veya ihmal edilebilir
B	$0,059 \leq \beta < 0,088$	Orta düzeyde
C	$ \beta \geq 0,088$	Yüksek düzeyde

Li ve Stout (1996), çalışmalarında TBO DMF tespit etmede SIBTEST analizinden daha etkili olduğunu düşündükleri Crossing-SIBTEST (CSIBTEST) analizini geliştirmişlerdir. Chalmers (2012) ise CSIBTEST hipotez test yaklaşımının yetersiz olduğunu öne sürerek, CSIBTEST istatistiğinde bazı basit değişiklikler yapmıştır. Geliştirilen CSIBTEST istatistiğinin bu yeni versiyonu ile SIBTEST istatistiği (Shealy ve Stout, 1991) arasında daha yakın bir ilişki olduğu gözlenmiştir. Özellikle etki büyüklüğü tahminlerinde SIBTEST ve yeni CSIBTEST istatistikleri ile aynı olduğu kanısına varılmıştır (Chalmers, 2012).

Her iki istatistiğin bir arada kullanıldığı analizlerde β_{uni} ve β_{cro} değerleri hesaplanır. Elde edilen β_{uni} değeri SIBTEST analizinden, β_{cro} değeri ise TBO DMF tespit etmede daha iyi olduğu tahmin edilen CSIBTEST analizinden gelir. DMF türünü belirlemek için elde edilen β değerleri şu şekilde yorumlanmaktadır: Eğer, β_{cro} ve β_{uni} değerleri özdeş ise TB DMF; $|\beta_{cro}| \geq |\beta_{uni}|$ eşitliği var ise TBO DMF varlığından söz edilebilir. Bununla birlikte, TBO DMF iki standart normal dağılımın karesinin toplamını temsil etmekte ve $\beta_{cro} = 0$ hipotezi 2 serbestlik derecesindeki ki kare dağılımı ile hesaplanmaktadır. Kesişme olmaması halinde ise β_{uni} 1 serbestlik derecesindeki ki kare dağılımı ile test edilmektedir.

Tek bir teknik kullanılarak yapılan analizler ile sonuçların tek bir tekniğe dayandırılması yanlılık çalışmalarında yanlış kararlar alınmasına neden olabilir. Bu nedenle birden fazla yöntemle göre yapılan DMF analizlerinde maddenin DMF gösterdiğine dair daha fazla gerekçe elde edilmiş olur (Akın ve diğerleri, 2016; Asil ve Gelbal, 2012; Camilli ve Shepard, 1994; Gök ve diğerleri, 2010; Karami ve Nodoushan, 2011). Bu görüşü destekler nitelikte olan araştırmalarda, birden çok DMF belirleme yöntemlerinin birlikte kullanılması yanlılık çalışmaları için uzmanlara daha fazla deliller sağlayacağını belirtmektedir (Akın ve diğerleri, 2016; Camilli ve Shepard, 1994; Gök ve diğerleri, 2010; Hambleton ve Patsula, 1999; Hidalgo ve López-Pina, 2004; Karami ve Nodoushan, 2011; Koyuncu, Aksu ve Kelecioğlu, 2018). Bu doğrultuda çok kategorili verilerde kullanılan bu üç yöntemle literatürde sık rastlanmaması, yöntemlere erişimin ve uygulamasının kolay olması nedeniyle bu araştırma için tercih edilmiştir.

İlgili Araştırmalar

Bu bölümde, Türkiye’de ve çeşitli ülkelerde farklı yöntemlere göre yapılan çok kategorili verilerde DMF analizleri ile ilgili araştırmalara yer verilmiştir.

Mellor (1995) çok kategorili veri setinde DMF analizini GMH, poly-SIBTEST, OLR ve Lojistik Diskriminant Fonksiyon Analizi (LDFA) yöntemleriyle yapmıştır. Araştırma sonucunda dört yöntemin farklı yetenek dağılımlarında TB ve TBO DMF tespit etme gücü ve 1. tip hata oranlarını incelemiştir. Bu çalışmada simülasyon çalışmasında üretilen veriler ile Texas Eğitim Bölümünden edinilen gerçek veriler üzerinde analizler yapılmıştır. Araştırma neticesinde iki grup arasında yetenek dağılımlarının aynı olduğu şartlarda, dört yöntemde TB DMF belirleyebildiği ancak poly-SIBTEST ve GMH yöntemlerinin diğer yöntemlerden daha hassas sonuçlar verdiği kanısına varılmıştır. Yetenek dağılımlarının eşit olmadığı şartlarda ise; GMH

yönteminin diğer yöntemlere göre en iyi çalışan yöntem olduğu ve 1. tip hatanın kontrolünde de mükemmel olduğu sonucuna varılmıştır. Aynı zamanda gerçek veri setinde etnik kökene göre yürütülen DMF analizinde, hispanic ve beyaz ırka göre yapılan kıyaslamada dört yöntemde de aynı düzeyde TB DMF'ye rastlanırken, beyaz ve siyah ırka göre yapılan kıyaslamada LDFA yönteminde TBO DMF'ye, OLR ve GMH yöntemlerinde ise TB DMF'ye rastlanmıştır.

Henderson (2001) İngilizce, Sosyal Bilimler, Matematik ve Biyoloji alanları üzerinde gerçekleşen lise sınavlarında cinsiyete göre DMF gösteren maddelerin nedenlerini araştırmıştır. Araştırma için iki kategorili maddeler için MH ve SIBTEST çok kategorili maddeler için ise GMH ve poly-SIBTEST yöntemleri kullanılarak elde edilen sonuçların uyumu incelenmiştir. GMH yönteminin etki büyüklüğünü belirlemek için standartlaştırılmış ortalama farkından yararlanılmıştır. Araştırma neticesinde iki kategorili ve çok kategorili maddeler için kullanılan yöntemlerin karşılaştırılabilir olduğu ve aralarında manidar bir fark olmadığı belirtilmiştir. Poly-SIBTEST yönteminin daha fazla DMF'li madde saptadığı ve GMH yönteminde DMF'li çıkan tüm maddeleri de kapsadığı sonucuna varılmıştır. Ayrıca iki kategorili maddelerin erkek öğrenciler lehine, çok kategorili maddelerin ise kız öğrenciler lehine çalıştığı gözlenmiştir. Sonuç olarak, cinsiyet ile madde yapısı arasında bir ilişki olduğu ifade edilmektedir.

Sandilands (2008) PIRLS 2001 uygulamasından elde edilen veriler ile DMF analizinde bulunmuştur. Örneklemini Arjantin, Kolombiya, İngiltere ve ABD ülkelerinden katılan cevaplayıcıların oluşturduğu bu çalışmada iki kategorili ve çok kategorili veriler için OLR ve poly-SIBTEST yöntemleri tercih edilmiştir. Elde edilen analizler neticesinde DMF belirlemede seçilen iki yöntem arasında yüksek oranda benzerlik olduğu fakat etki büyüklük düzeyleri bakımından uyum gözlenmediği tespit edilmiştir. Poly SIBTEST yöntemindeki etki büyüklüğü (Rousso ve Stout, 1996) ile OLR yönteminde etki büyüklüğü (Gierl ve Jodoin, 2001) karşılaştırıldığında poly-SIBTEST yöntemi her düzeyde DMF tespit ederken, OLR yönteminde ihmal edilebilir düzeyde DMF'li madde sayısının daha fazla olduğu aktarılmaktadır.

Atalay (2010) PISA 2006 öğrenci anketinde ABD ve Türkiye örneklemelerinde cinsiyete ve kültüre göre poly-SIBTEST ve Ordinal Lojistik Regresyon tekniklerini kullanarak DMF analizleri yapmıştır. Elde edilen bulgulara göre Türkiye de; OLR yöntemi 12 madde, poly-SIBTEST yöntemi 13 maddede DMF saptarken, ABD de uygulanan iki yöntemde 15 maddede DMF saptanmıştır. Her iki yöntemde ortak DMF içeren madde sayısı ABD örneğinde 13, Türkiye örneğinde 7'dir. Cinsiyet ve

kültüre göre yapılan bu arařtırmada yöntemlerin DMF belirlemede uyumluluk gösterdiđi; ancak düzeyler arasında farklılıklar olabileceđi sonucuna varılmıřtır.

Huang (2010) arařtırmasında PISA uygulamasına katılan ABD, Kanada, Hong Kong ve Çin verilerini karřılařtırarak tutum maddelerinde DMF analizinde yapmıřtır. Rash modelinin tercih edildiđi bu alıřmada DMF ieren madde sayısına en az Kanada ve ABD rneklemleri arasında rastlanırken, en fazla Çin ve ABD rneklemleri arasında rastlanmıřtır. Yapılan ierik analizinde DMF'nin nemli bir kısmına Hong Kong-in ve ABD-in grupları arasında rastlanmıř ve DMF kaynađı olarak farklı mfredat ierikleri gsterilmiřtir. Beklenenin aksine ABD ve in đrencileri arasındaki dil farkı ise manidar düzeyde DMF iermediđi sonucuna varılmıřtır.

Yıldırım ve Yıldırım (2011) alıřmalarında madde glklerdeki ortak varyansla iliřkili deđiřkenlerin, farklı iřleyen madde analizlerinde ek eřleme deđiřkeni olarak incelemiřlerdir. Bu dođrultuda, 2003 ve 2006 PISA matematik okuryazarlıđı testinin verilerinden yararlanılmıřtır. Sonu olarak, đrencilerin okul dıřında đrenmek iin harcadıkları zaman deđiřkeninin ikinci bir eřleřtirme deđiřkeni olarak kullanılabileceđi tespit edilmiřtir. Ayrıca, bu deđiřkenin farklı iřlemesinin sebebi olarak ok deđiřkenli eřleme ynteminin tercih edilmesi dayanak gsterilmiřtir.

Asil ve Gelbal (2012) arařtırmalarında PISA đrenci anketinin kltrler arası eřdeđerliđini Avustralya, Yeni Zelanda, Amerika Birleřik Devletleri ve Trkiye rneklemleri zerinde karřılařtırmalı olarak incelemiřlerdir. Dođrulayıcı Faktr Analizi ve oklu Grup Dođrulayıcı Faktr Analizi kullanarak maddelerin farklı dil ve kltre gre DMF ierip iermediđine bakılmıřtır. Sonu olarak lkeler arasında dil ve kltr farklılıkları ile orantılı olarak DMF ieren maddelerin sayısında da artıř gzlemlediđi bildirilmiřtir. Ayrıca, DMF'nin asıl nedeni olarak da eviri problemleri ve kltrel farklılıklar olduđu sonucuna varılmıřtır. alıřmanın sonu blmnde ilgili arařtırmada analizler ile yapıldıđı iin DMF belirlemede Madde Tepki Kuramına (MTK) dayalı diđer yntemlerinde kullanılarak incelenebileceđi nerilmiřtir.

Taylor ve Lee (2012) arařtırmalarında poly-SIBTEST ve Rash modelinin yntemleri ile cinsiyete dayalı yanlılık alıřmasında bulunmuřlardır. Arařtırma sonucunda Rasch modeli SIBTEST ynteminden daha fazla sayıda DMF'li madde tespit ettiđi kanısına varmıřlardır. Ayrıca oktan semeli maddeler erkekler lehine iřlerken, yorumlama, akıl yrtme, metin analizi gibi yapılandırılmıř maddeler kadınlar lehine alıřtıđı gzlenmiřtir.

Fidalgo, Pons ve Quintanilla (2013) çocukların duygularını anlama ve bu doğrultuda yeteneklerini değerlendirmek için farklı dil ve farklı kültürlerde uygulanan testte psikometrik özelliklerini keşfedici analizler yapmışlardır. DMF analizi için GMHDIF programından yararlanılmıştır. BILOG-MG 3.0 parametre kestirimi kullanılarak 3-11 yaşları arasında 100 İngiliz çocuğun teste verdikleri yanıtlar incelenmiştir. DMF'li maddeleri değerlendirmek için X^2_{MH} ($\alpha = 0,20$) istatistiği kullanılmıştır. Analiz neticesinde yalnızca üç maddenin kız ve erkek öğrenciler için farklı işlediği saptanmıştır. Ayrıca daha büyük örneklem büyüklüğü ile analizin tekrar test edilmesi gerektiğini vurgulamışlardır.

Başokçu ve Öğretmen (2013) araştırmalarında Ege Üniversitesi Pedagojik formasyon programında öğretmen öz yeterlik ölçeğinin geçerlik çalışması yapmışlardır. Öğretmenlik mesleğine ilişkin yeterlik algısını ölçmek için ölçme aracının alt popülasyonlar için madde bazında farklılaşmasını DMF analizleri ile incelenmiştir. Likert tipi puanlama nedeniyle analizlerde Samejima'nın ağırlıklandırılmış cevap modeli seçilmiştir. 2012-2013 öğretim yılları arasında Ege üniversitesi pedagojik formasyon programına katılan 626 öğretmen adayı araştırmanın örneklemi oluşturmaktadır. Hem ölçek toplam puan hem de ölçeğin alt boyutlarının toplam puanı ayrı ayrı ele alınarak yapılan analizler sonucu cinsiyete göre 7 maddede DMF belirlenmiştir.

Lisova ve Kovalchuk (2016) TIMSS 2011 verisini inceleyerek DMF analizi yapmışlardır. 249 Ukrayna, 742 ABD ve 354 Rusya Federasyonu katılımcılarından oluşan örnekleme, matematik alanından 32 madde (0-1 veya 2 puanlama) ele alınmıştır. Tek biçimli ve tek biçimli olmayan DMF'yi belirlemek için poly-SIBTEST ve crossing-SIBTEST yöntemleri kullanılmıştır. Analizde 0.05 anlamlılık değerini dikkate alınarak, R^2 etki büyüklüğü hesaplanmıştır. Sonuçta, DMF belirlemede MH ve SIBTEST yöntem çifti arasında en yüksek (%88) uyumluluk tespit edilirken, en düşük (%70) uyumluluk ise SIBTEST-LR çifti arasında gözlenmiştir.

Köse (2015) PISA 2009 öğrenci anketi alt ölçeklerinde (Q32-okul öğrenmelerine ilişkin algı ve Q33-öğretmenlere yönelik algı) bulunan maddelerin değişen madde fonksiyonu açısından incelenmesi adlı araştırmasında okula ve öğretmenlere ilişkin algı alt testlerinde yer alan maddeler üzerinden DMF analizleri yapmıştır. ABD, Türkiye, İngiltere ve İrlanda ülkeleri seçilerek oluşturulan veri setinde cinsiyete değişkeni ele alınmıştır. Alt ölçeklerdeki maddeler çoklu puanlandığı için Aşamalı Tepki Modeli

tercih edilerek MULTILOG programı kullanılmıştır. Öğretmenlere yönelik algı alt ölçeğinde cinsiyete göre 1; İrlanda-İngiltere’de 1, ABD-İngiltere’de 2 ve ABD-Türkiye’de 4 maddede DMF bulunurken okul öğrenmelerine ilişkin algı alt ölçeğinde cinsiyete göre 2, İrlanda-İngiltere’de 1, ABD-İngiltere’de 2 ve ABD-Türkiye’de 4 maddede DMF bulunduğu analizler sonucu ortaya konmuştur.

Akın Arıkan (2015) çalışmasında PISA 2012 matematik çalışma etiği tutum maddelerini cinsiyete göre DMF belirlemede MTK-Olabilirlik Oranı (MTK-OO), Ordinal Lojistik Regresyon ve poly-SIBTEST yöntemlerini karşılaştırmıştır. Çalışmanın örneklemini, matematik çalışma etiği tutum maddelerini cevaplayan 3217 bireyden rastgele seçilen 500 kadın 500 erkek bireyler oluşturmaktadır. MTK ile yapılan madde analizinde, 3. maddenin poly-SIBTEST yöntemine göre yüksek düzeyde erkekler lehine, MTK-OO yöntemine göre orta düzeyde ve OLR yöntemine göre yüksek düzeyde tek biçimli olmayan DMF içeren maddelere ulaşılmıştır.

Mohorić ve Takšić (2016) on iki farklı ülkede uygulanan duygusal yeterlilik anketinde GMHDIF programı kullanılarak DMF analizi yapmıştır. Analizde ilk olarak GMHDIF analizi, üç ülkeden (Hırvatistan, Slovenya ve Çin) rastgele iki örnekle karşılaştırarak test edilmiştir. Ardından on iki ülkeden (Hırvatistan, Slovenya, Sırbistan, İspanya, Finlandiya, İsveç, Türkiye, ABD, Arjantin, Çin, Japonya ve Hindistan) kadın ve erkek katılımcı cevapları üzerinde DMF analizi yapılmıştır. Daha sonra, Slav (Hırvatistan, Sırbistan ve Slovenya) ve İspanyolca (İspanya ve Arjantin) dil grup grupları arasında DMF analizi yapılmıştır. Analiz sonucunda, cinsiyete göre ve üç ülke arasından rastgele seçilen örnek karşılaştırması en az sayıda DMF tespit ederken, dil grubuna göre yapılan analizde farklı dil grupları arasında en fazla DMF tespit edilmiştir.

Özberk ve Koç (2017), Wechsler Çocuklar için Zekâ Ölçeği IV Türkçe formundaki maddelerin sosyo ekonomik düzey ve cinsiyete göre yanlılık çalışmasını birden fazla yöntemle inceleyerek yöntemler arası uyumluluğa bakmışlardır. Çalışmada çoklu puanlanan maddelerden oluşan (küplerle desen, benzerlikler, sayı dizisi, sözcük dağarcığı, harf-rakam dizisi ve kavrama) ve ikili puanlanan maddelerinden oluşan (resim kavramları, mantık yürütme kareleri, resim tanımlama, genel bilgi, aritmetik ve sözcük bulma) 12 alt test ve bu alt testlerde yer alan 315 madde incelenmiştir. Maddelerin farklı işleyip işlemediğini belirlemek amacıyla ikili puanlama modeli için Rasch Modeli, Mantel-Haenszel, SIBTEST yöntemleri; çoklu puanlanan maddeler için de kısmi puan modeli, Mantel test ve poly-SIBTEST yöntemleri

kullanılmıştır. DMF belirleme yöntemleri incelendiğinde, genel olarak yöntemlerin birbiriyle tutarlı sonuçlar verdiği ancak cinsiyet açısından DMF içeren maddelerin tespitinde MH, SIBTEST ve Mantel Test, poly-SIBTEST istatistiklerinin daha tutarlı sonuçlar verirken sosyoekonomik düzeye göre DMF içeren maddelerin tespitinde MH, Rasch Modeli ve Mantel Test, Kısmi Puan Modelinin (KPM) daha tutarlı sonuçlar verdiği analizler sonucu tespit edilmiştir.

Tunç ve Kutlu (2017) çalışmasında iki ve çok kategorili verilerde değişen madde fonksiyonunu karşılaştırmak amacıyla bir simülasyon çalışması gerçekleştirmişlerdir. I. tip hata ve istatistiksel güç oranları üzerinde çalışılan bu araştırmada 20 madde üzerinden hem iki kategorili hem de çok kategorili puanlama yapılarak iki ayrı veri seti oluşturulmuştur. İki kategorili veriler için Rasch, çok kategorili veriler için ise KPM tercih edilmiştir. R programında veriler türetilmiş ve WinGen programıyla madde parametreleri hesaplanmıştır. Araştırma sonucunda iki kategorili ve çok kategorili puanlama modelleri kapsamında DMF kıyaslaması yapıldığında, çok kategorili puanlama yapılması durumunda I. tip hata oranlarının daha düşük, istatistiksel güç oranlarının ise daha yüksek olduğu saptanmıştır. Bu bağlamda puanlama modellerinin DMF sonuçlarına etkisi olduğu ve kısmi puana göre de DMF sonuçlarının değişebileceği aktarılmıştır.

Pasca, Coello, Aragonés ve Frantz (2018) İspanya ve Kuzey ABD’de uygulanan 13 likert tip maddeden oluşan doğaya bağlanma ölçeği üzerine DMF araştırması yapmışlardır. Bu analiz için GMH ($Q_{GMH(2)}$) değeri için GMHDIF programı ve OLR (SAS 9.4 paketi) yöntemleri test edilmiştir. GMH istatistiği sonucunda 13 maddenin 7’sinde DMF tespit edilmiştir. Bu yöntemin sınırlılığı olarak TB ve TBO DMF hakkında bilgi vermemesi gösterilmektedir. OLR istatistiğiyle ise 13 maddeden 11’inde DMF’li madde bulunmuş ve bu maddelerin hepsi TB DMF içerdiği belirtilmiştir. Farklı kültürlerde uygulanan bu testin içeriğine dair yayınlar ABD grubunda 2004 yılında yayınlanırken, İspanyol grubunda 2014 yılına kadar yayınlanan çalışmaları kapsamaktadır. Bu nedenle, kavramlardaki farklılıkların gruplar arasında zaman içinde farklı anlarda ortaya çıkıp çıkmadığının araştırılması gerekmektedir. Diğer bir ifade ile, farklı ülkelerden elde edilen sonuçlar karşılaştırıldığında ölçeğe farklı verilen cevapların nedeni farklı toplumsal yapıdan mı yoksa farklı anlarda bilgi sahibi oluşundan mı kaynaklandığının dikkatli yorumlanması gerektiği aktarılmaktadır.

Sonuç olarak alandaki DMF çalışmaları incelendiğinde kültüre, cinsiyete, yöneme ve yıllara göre yapılan karşılaştırmaların kapsamlı bir şekilde ele alındığını söylemek mümkündür. Genel olarak araştırma sonuçları kültürel arası farklılıklar arttıkça testlerdeki DMF'li madde sayısının arttığını destekler niteliktedir. Bunun yanı sıra yöntem kıyaslamaları çalışmalarına bakıldığında çoklu yöntemlerle yapılan karşılaştırmaların tek bir yöneme göre yapılan DMF analizlerinden daha sağlıklı sonuçlar verdiği gözlenmiştir.

Mevcut çalışma PISA 2015 fen bilimleri alt ölçeğinde yer alan maddelerin çok kategorili veriler için geliştirilmiş farklı yöntemler (GMH, OLR ve poly-SIBTEST) kullanılarak dil ve kültüre göre DMF gösterip göstermediğini incelemeye yönelik bir çalışma olması nedeniyle diğer çalışmalardan özgünlük göstermektedir. Üç yöntemden elde edilen analizler sonucu yöntemlerin birbiri ile uyumu da ayrıca incelenmiştir.

Araştırmaya konu olan PISA 2015 alt ölçeğinde fene yönelik tutumları içeren maddeler seçilmiştir. Bu kapsama dahil olan fen öğrenmekten zevk alma adlı ST 94 kodlu ve fen öğretiminde araçsal güdülenme adlı ST 113 kodlu maddeler ele alınmıştır.

Bölüm 3

Yöntem

Bu bölümde araştırmanın türüne, araştırmanın evreni ve örnekleme, veri toplama sürecine, veri toplama araçlarına ve verilerin analizine yer verilmiştir.

Araştırmanın Türü

Araştırmada PISA 2015 fen bilimleri uygulamasında yer alan anket maddelerinin farklı kültürlerle ve farklı dil gruplarına göre DMF gösterip göstermediği incelenmesi bakımından betimsel bir araştırma niteliği taşımaktadır.

Araştırmanın Evreni ve Örnekleme

Araştırmanın evrenini PISA 2015 uygulamasına katılan 72 ülkeden yaklaşık 540 bine yakın öğrenci oluşturmaktadır (PISA, 2015). Bu uygulamaya katılan ABD, İrlanda, İngiltere ve Türkiye örneklemi seçilmiştir. Maddenin DMF göstermesinin olası nedenleri arasında kültürel farklılıklar ve çeviri sorunları yer aldığı düşünüldüğünden bu ülkeler seçilmiştir. PISA 2015 uygulamasına katılan örneklemeindeki öğrencilerin cinsiyete ve ülkelere göre frekans ve yüzde dağılımları Tablo 6'da verilmiştir.

Tablo 6

Cinsiyete ve ülkelere göre Frekans ve yüzde dağılımları

	Kadın		Erkek		Toplam
	<i>f</i>	%	<i>f</i>	%	
İngiltere	2151	48,7	2266	51,3	5000
İrlanda	2597	50,4	2560	49,6	5157
ABD	2528	49,8	2472	50,2	5000
Türkiye	2617	51,4	2472	48,6	5089

Tablo 6 incelendiğinde Türkiye'den 5089, İrlanda'dan 5157, ABD'den 5000 ve İngiltere'den toplam 4417 katılımcının yer aldığı görülmektedir. Cinsiyet dağılımları incelendiğinde dört ülkenin de cinsiyet dağılım yüzdeleri birbirine yakın olduğu gözlenmektedir. Ayrıca, araştırmada kullanılacak 2 farklı alt boyutta yer alan 9 madde için ülkeler bazında ortalama değerlere Tablo 7'de yer verilmiştir.

Tablo 7

Ülkeler bazında maddelerdeki ortalama değerler

	Maddeler	ING	IRL	ABD	TR
ST94	Fen bilimleri ile ilgili bir şeyler öğrenmeye ilgi duyarım.	2,77	2,69	2,83	2,63
	Fen bilimleri hakkında yeni bilgiler edinmekten zevk alırım.	2,56	2,59	2,61	2,67
	Fen bilimleri ile ilgili konularda çalışmak beni mutlu ediyor.	2,81	2,78	2,78	2,64
	Fen bilimleri ile ilgili yazıları okumaktan zevk alırım.	2,83	2,91	2,90	2,77
	Fen bilimleri ile ilgili konuları öğrenirken genellikle eğlenirim.	2,81	2,88	2,88	2,78
ST113	Fen dersi konularında (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrendiğim birçok şey, iş bulmamda yardımcı olacak.	1,91	1,96	1,96	1,95
	Öğreneceklerim geleceğe yönelik mesleki beklentilerime yardımcı olacağından, okuldaki fen dersi konularını (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrenmek benim için önemlidir.	2,10	2,10	2,08	1,99
	Sonradan yapmak isteyeceğim şeyler için ihtiyaç duyacağımdan, okuldaki fen dersi konularında (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrendiklerim benim için önemlidir.	1,96	1,99	2,06	2,03
	Daha sonra yapacağım çalışmalarda bana yardımcı olacağından, okuldaki fen dersi konularını (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrenmek için çaba göstermeye değer.	2,07	2,07	2,13	2,12

Tablo 7 incelendiğinde, ülkelere ait ortalama değerlerin yaklaşık olarak birbirleri ile tutarlılık gösterdiği saptanmıştır. Fen öğrenmeden zevk alma (ST94) alt boyutuna ait ortalama değerlerin fene yönelik motivasyona (ST113) ait ortalama değerlere göre daha yüksek olduğu görülmektedir. Ayrıca 1,91 ile en düşük ortalama “fen dersi konularında (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) öğrendiğim birçok şey, iş bulmamda yardımcı olacak” maddesi ile İngiltere iken 2,91 ile en yüksek ortalama değere sahip “fen bilimleri ile ilgili yazıları okumaktan zevk alırım” maddesi ile İrlanda olduğu gözlenmiştir.

Veri Toplama Süreci

Araştırmada Ekonomik İşbirliği ve Kalkınma Örgütü'nün 2015 yılında uyguladığı PISA öğrenci anket verisi kullanılmıştır. Kullanılan veriler 2019 yılı Ocak ayında OECD PISA (<http://www.oecd.org/pisa/data/2015database/>) internet sitesinden indirilmiştir.

Veri Toplama Araçları

Araştırmada OECD PISA sayfasından alınan SPSS verileri ve öğrenci anketinde yer alan tutum maddelerinden yararlanılmıştır. Anket maddelerine EK-A'da yer verilmiştir.

Verilerin Analizi

Bu bölümde veri analizi aşamaları anlatılmaktadır.

1. Aşama. PISA 2015 uygulamasında fen okuryazarlığının yanında başarılarını etkileyen unsurların belirlenmesi amacıyla duyuşsal özelliklere ilişkin tutum maddelerine yer verilmiştir. Bu uygulama kapsamında öğrencilere fen öğrenme motivasyonları (fen öğrenmeyi öğrencilerin ne kadar ilginç ve eğlenceli buldukları), fen konularına ilgileri ve fen öğreniminde araçsal güdülenmeleri (öğrendikleri fen dersini gelecekteki eğitimlerine ve kariyerlerine katkı sağlayacağı algısına ilişkin tutumları) ve fen öz yeterliliği (kendi yetenekleri doğrultusunda fen ödevlerinin ve zorlukların üstesinden gelebilme düzeyleri) başlıkları altında sorular sorulmuştur (OECD, 2016). Bu bağlamda PISA 2015 öğrenci anketi maddeleri incelenmiş ve araştırmaya fene yönelik tutum maddeleri dahil edilmiştir. Ankette yer alan fen öğrenmekten zevk alma ST94 kodlu soru (ST94Q01NA, ST94Q02NA, ST94Q03NA, ST94Q04NA, ST94Q05NA) ve fen öğreniminde araçsal güdülenme ST113 kodlu soru (ST113Q01TA, ST113Q02TA, ST113Q03TA, ST113Q04NA) maddeleri incelenmiştir.

Bu arařtırmada iki alt boyut iin seilen 9 tutum maddelerine ait kodlar ve Trke evirileri Tablo 8'de verilmiřtir.

Tablo 8

ST 94 ve ST113 kodlu ğrenci anketine ait maddeler

Boyut	Madde	Soruların orijinal hali	Soruların Trke'ye evrilmiř hali*
ST 94	Q01	I generally have fun when I am learning science topics.	Fen bilimleri ile ilgili bir řeyler ğrenmeye ilgi duyarım.
	Q02	I like reading about science.	Fen bilimleri hakkında yeni bilgiler edinmekten zevk alırım.
	Q03	I am happy working on science topics.	Fen bilimleri ile ilgili konularda alıřmak beni mutlu ediyor.
	Q04	I enjoy acquiring new knowledge in science.	Fen bilimleri ile ilgili yazıları okumaktan zevk alırım.
	Q05	I am interested in learning about science.	Fen bilimleri ile ilgili konuları ğrenirken genellikle eęlenirim.
ST113	Q01	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.	Fen dersi konularında (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) ğrendięim birok řey, iř bulmamda yardımcı olacak.
	Q02	What I learn in my science classes is important for me because I need this for what I want to do later on.	ğreneceklerim geleceęe ynelik mesleki beklentilerime yardımcı olacaęından, okuldaki fen dersi konularını (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) ğrenmek benim iin nemlidir.
	Q03	Studying science is worthwhile for me because what I learn will improve my career prospects.	Sonradan yapmak isteyeceęim řeyler iin ihtiya duyacaęımdan, okuldaki fen dersi konularında (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) ğrendiklerim benim iin nemlidir.
	Q04	Many things I learn in my science subject(s) will help me to get a job.	Daha sonra yapacaęım alıřmalarda bana yardımcı olacaęından, okuldaki fen dersi konularını (Fen ve Teknoloji, Fizik, Kimya, Biyoloji) ğrenmek iin aba gstermeye deęer.

*PISA 2015 ğrenci anketine ait ST 94 ve ST113 kodlu sorularının evirileri MEB'in PISA 2015 ulusal raporu sayfasından (http://pisa.meb.gov.tr/wp-content/uploads/2014/11/PISA2015_UlusalRapor.pdf) alınmıřtır.

2. Aşama. Analize başlamadan önce sayıtların kontrolü amacıyla kayıp veri analizi yapılmıştır. Yapılan kayıp veri analiziyle kayıp veri içeren maddeler silinmiştir. Ardından tek değişkenli uç değer analizi için z puanları hesaplanmıştır (± 3 aralığını aşan değerlerin olup olmadığı araştırılmıştır). Son olarak çok değişkenli uç değer belirleme yöntemlerinden mahalalanobis uzaklığı analizi sonucunda $p < ,001$ ölçütüne göre serbestlik derecesi 9'a karşılık gelen 27.877'den büyük veriler silinmiştir.

3. Aşama. Faktör yapıları belirlenen alt boyutlarda olup olmadığı kontrolü amacıyla açımlayıcı faktör analizi yapılmıştır. Bu bağlamda verilerin faktör analizi yapımına uygunluğunu incelemek için Barlett küresellik testinin anlamlılığına ($< 0,05$) ve Kaiser-Meyer-Olkin (KMO) ($> 0,50$) değerlerinin uygunluk derecesine bakılmıştır. KMO ölçütünün ise 0,50 altında değer alması halinde veri setinin faktör analizi için uygun olmadığı ve bu analize dayalı yorumlar yapılmaması gerektiği vurgulanmaktadır (Çokluk, Şekercioğlu, ve Büyüköztürk, 2012; Field, 2000; Şencan, 2005; Tabachnick ve Fidell, 2007). Bu amaçla elde edilen Barlett küresellik testi ($< 0,05$) ile KMO değerleri ($> 0,900$) analiz için uygunluğu test edilmiştir. Elde edilen bu sonuçlar verilerin çok değişkenli normal dağılımdan geldiğini göstererek faktör analizinin diğer sayıtlısını da karşılamaktadır.

4. Aşama. Bir önceki aşamadaki koşulların sağlanmasının ardından değişkenlerin boyutlarına bakılmıştır. SPSS programından yararlanılarak faktör sayısının belirlemek amacıyla özdeğer ve varyans analizi, faktörlerin rotasyonu, faktör yükleri ve yamaç birikinti grafiklerine bakılarak faktörlerin kaç grup altında toplandığı incelenmiştir. Özdeğer istatistiğinde 1'den büyük değerler anlamlı kabul edildiğinden (Çokluk ve diğerleri., 2012) her bir faktör için 1'den büyük değerler dikkate alınmıştır. Varyans yüzdeleri incelenerek faktör dağılımına bakılmıştır. Yamaç birikinti grafikleri ile kırılma noktaları incelenerek faktör sayıları belirlenmiştir.

5. Aşama. Bu çalışmada çok kategorili verilerde R programı kapsamında poly-SIBTEST analizleri için ise MIRT paketinden (Chalmers, 2012), OLR analizleri için LORDIF paketinden (Choi, Gibbons ve Crane, 2011) yararlanılmıştır. Bu analizler için R programında girilen örnek komutlara EK B'de yer verilmiştir. Poly-SIBTEST analizi için tercih edilen R programından SIBTEST ve CSIBTEST çıktıları elde edilmiştir. Bu çalışmada, iki istatistikten elde edilen β değerlerinin yorumlanması ile DMF türüne bakılmıştır. Aynı zamanda, SIBTEST analizi sonucunda DMF varlığı, düzeyi ve avantaj sağladığı gruplarda ayrıca incelenmiştir.

OLR analizi için LORDIF paketi kullanılmıştır. LORDIF, LR ve MTK özelliklerini birlikte kullanarak geliştirilmiş bir paket programdır. Crane ve diğerleri (2006), yaptıkları çalışmada LORDIF paketi OLR analizinde daha hassas sonuçlar verdiği için TBO DMF belirlemede yeni bir teknik geliştirmişlerdir. Bu amaçla, Bonferroni düzeltmesi (α/n) kullanmışlardır. Bu hesaplama göre analizde kullanılan alfa değeri madde sayısına bölünür. Bu çalışmada iki alt boyutta ayrı ayrı analizler yapıldığı için iki farklı Bonferroni değeri ($ST94:0,002$ ve $ST113:0,0025$) hesaplanmıştır. TBO DMF belirlemek için Model3-Model2 R^2 , TB DMF belirlemek için Model2-Model1 R^2 ve toplam DMF etkisini belirlemek için ise Model1-Model3 R^2 log olasılık oran değerleri arasındaki fark ki kare dağılımları ile karşılaştırılmış ve analiz sonuçları tablolaştırılmıştır (Choi, Gibbons ve Crane, 2011). Elde edilen bulgular sonucunda, McFadden Model3-Model2 farkları ile hesaplanan Bonferroni değerleri karşılaştırılarak TB ve TBO DMF türleri belirlenmiştir.

LORDIF paketi ile OLR analizi. R yazılımında yer alan (Choi, Gibbons ve Crane, 2011) LORDIF paketinde LR analizi tablosunda öncelikle her bir LR modelini (Model 1'e karşı Model 3, Model 1'e karşı Model 2 ve Model 2'ye karşı Model 3) karşılaştırmak için X^2 oran istatistikleri oluşturulur. Ardından, etki büyüklüğü kestiricisi (R^2) analizi yaparken diğer programlardan farklı olarak her bir modele ilişkin Nagelkerke (Nagelkerke, 1991) değerlerinin yanında Mc Fadden (Menard, 2000) ve Cox ve Snell (Cox ve Snell, 1989) değerleri hesaplanır.

LORDIF paketi ile DMF'nin belirlenmesi. Bu paket program çerçevesinde yapılan DMF belirleme kriterlerinden (ki-kare, beta veya R^2) biri kullanıcı tarafından seçilerek (örneğin, kriter ki kare için alfa değeri 0,001) DMF içeren maddeler tespit edilir. Üç olasılık oranından X^2 istatistiğinden birinin anlamlı olması durumunda o maddenin DMF'li olduğu söylenebilir. Fakat seçilen tek bir kriter TB DMF'ye atfediliyorsa, TBO DMF'yi belirlemede zayıf kalabileceği de eleştiri olarak sunulmaktadır (serbestlik derecesinin 2 olması halinde) (Choi, Gibbons ve Crane, 2011).

Bu çalışmada GMH yöntemi için Fidalgo (2011) tarafından geliştirilen GMHDIF programı kullanılmıştır. Bu programın tercih nedenlerinden birisi kullanımının kolay olmasıdır. Diğer avantajları ise; iki aşamalı DMF analizlerini aynı anda çoklu gruplarda yapabilmesi, (eğer H_0 reddedilirse) ikili karşılaştırmalar yapabilmesi ve son olarak iki ve çok kategorili veriler için uygulanabilir olmasıdır (Fidalgo, 2011). Bu

çalışmada GMHDIF programı kapsamında GMH istatistiğinde $Q_{GMH(1)}$ ve $Q_{GMH(2)}$ hipotezleri analiz edilmiştir. $Q_{GMH(1)}$ istatistiği (*General Nominal MH Statistic GNMH*) H_0 hipotezi altında serbestlik derecesi $df=(R-1)(C-1)$ için yaklaşık olarak ki kare dağılımına sahip ve $R=C=2$ için $X^2_{(MH)}$ istatistiğine benzerlik gösterir. $Q_{GMH(2)}$ (*Generalized Ordinal MH Statistic GOMH*) H_0 hipotezi altında serbestlik derecesi $df=R-1$ ki kare dağılımına sahiptir. $Q_{GMH(1)}$ Genelleştirilmiş Mantel Haenszel (Mantel ve Haenszel, 1959) yöntemine benzerlik gösterirken $Q_{GMH(2)}$ geliştirilmiş MH yöntemine benzerlik gösterir (Mantel, 1963). Fidalgo ve Cobian (2018), tüm simülasyon koşullarında örneklem büyüklüğünün artması MH istatistiklerinin güçlerini artırdığı yönünde olduğunu bildirir. Fakat benzer çalışmada gruplar arası yetenek dağılımlarının farklı olduğu durumlarda DMF tespit oranlarında düşüşe neden olduğu aktarılmaktadır. Bu bağlamda simülasyon çalışmalarında tek yöntem tercih edilecekse $Q_{GMH(2)}$, simülasyon olmayan yani maddelerin ne tür DMF'ye sahip olduğu öngörülemeyen testlerde ise $Q_{GMH(1)}$ 'nin daha karmaşık ilişkilendirme modelini saptayabildiği ve daha gerçekçi sonuçlar verdiği gözlenmiştir (Fidalgo ve Bartram, 2010; Fidalgo ve Scalon, 2012). Bu kapsamda bu araştırmada $Q_{GMH(1)}$ analiz sonuçları dikkate alınmıştır. Bununla birlikte, Fidalgo (2011) GMHDIF programını üç aşamada şu şekilde açıklamıştır: Birinci aşamada eşleştirme değişkeninden elimine edilmiş bir ilk analiz yapılır. İkinci aşamada birinci aşamadan elde edilen saflaştırılmış eşleşme değişkenini kullanarak maddeler analiz edilir. Üçüncü aşamada ise ikinci aşamada DMF tespit edilmiş maddelere GMH istatistikleri uygulanır. Fakat bu son aşamada grupları karşılaştırmak için Bonferroni hesaplaması (α/n) ile bir alfa değeri belirlenir.

GMHDIF programında örneklem büyüklüğü sınırlaması ($n:10.000$) bulunsa da araştırma için eşleştirilen ülkelerde bu limit aşılmamıştır. Diğer taraftan, GMH yönteminin birçok avantajı bulunmasına rağmen DMF türü, etki büyüklüğü veya yönü hakkında detaylı bilgiler vermemektedir (Alavi, Rezaee, ve Amirian, 2011; Fidalgo ve Scalon, 2012). Bu amaçla, DMF gösteren maddelerin hangi gruba avantaj veya dezavantaj sağladığını belirleyebilmek için çoklu puanlanan maddeler için DIFAS programında (Penfield, 2013) Liu-Agresti-Lojistik Odds Oranı (*Liu-Agresti Cumulative Common Log-Odds Ratio L-A LOR*) istatistiğinden (Liu ve Agresti, 1996; Penfield ve Algina, 2003) yararlanılmıştır. Bu program iki kategorili veriler için BD algoritması (Breslow-Day Chi-Square) ile etki büyüklüğü hakkında detaylı bilgi vermektedir. Fakat çok kategorili veriler için L-A LOR algoritması (Liu-Agresti Cumulative Common Log-

Odds Ratio) sadece hangi gruba avantaj sağladığı (değerin pozitif olması referans, negatif olması odak grup lehine) hakkında bilgi vermektedir (Penfield, 2013). Bu sebeple çok kategorili veriler için GMH yönteminin etki büyüklüğü hesaplanamamıştır.

Sayıtların İncelenmesi

Bu aşamada öncelikle veri setinin analiz yapılmasına uygun olup olmadığını test etmek amacıyla kayıp veri analizlerine ve uç değerlere bakılacaktır.

Kayıp verilerin kontrolü. Bazı cevaplayıcılardan veya gözlemlerden hiçbir bilgi elde edilememesi durumunda kayıp veri ortaya çıkmaktadır. Veri matrisine “gözlenemeyen veri” olarak girilen ve kayıp veriler içeren bir veri seti ortaya çıkar (Demir, 2013). Kayıp verilerin varlığı ise veri setinin daralmasına ve buna göre yapılacak kestirim gücünün de azalması neden olmaktadır. Kayıp veriler için kullanılan yöntemler genel hatlarıyla iki grupta sınıflandırmak mümkündür. Bunlar,

1. Silmeye ve basit atamaya dayalı yöntemler
2. Olasılıklı ve ötemeli veri atama yöntemleridir.

İstatiksel analizlerde yazılımların özelliklerine göre kullanılacak farklı kayıp veri yöntemi bulunmaktadır. Bu yöntemler dizin silme yöntemi ile silmeye göre “0 atama”, seri ortalamaları ataması, gözlem birimi ortalama ataması, yakın noktalar medyan ataması, doğrusal interpolasyon ve doğrusal eğim noktası ataması yöntemleri; basit atamaya göre regresyon atama, beklenti-maksimizasyon algoritması ve veri çoğaltma yöntemleri ve son olarak en çok olabilirlik yaklaşımına göre Markov zincirleri Monte Carlo ise çoklu atamaya yaklaşımına göre yapılan yöntemler arasında yer almaktadır.

Kayıp verilerin veri setinden çıkarma işlemi yapabilmek için kayıp verilerin tamamen rastlantısal olarak dağılıyor olması gerekir. Ayrıca silinen kayıp verilerin analizde var olan başka değişkenler ile bağlantılı olması durumunda yapılacak işlemin yanlı olmasına sebep olacaktır (Osborne, 2013). Bu bağlamda, kayıp veri silme işlemi yerine yaklaşık değer atama yöntemleri verilerin korunmasına katkı sağlayacaktır.

Verilerin seçkisiz kayıp (missing at random) içerdiği varsayıldığı ve büyük örneklerde %5'den az olan kayıp verilerin analiz dışı bırakılması gerektiği savunulmaktadır (Demir, 2013; Kline, 2011). Acock'e (2005) göre bu oran %20 ye kadar çıkabilmektedir. Bu araştırmada her bir ülke örnekleme için SPSS programında

kayıp veri analizi yapıldığında Türkiye örnekleminde 551, İngiltere'den 615, İrlanda'dan 412 ve ABD'den 502 veri silinmiştir.

Uç değerler analizi. Uç değerler bazı özelliklerden dolayı diğer gözlemlere zarar verebilen gözlemler olarak tanımlanmaktadır (Elashoff, 1972). Başka bir tanımda ise veri setinde diğer gözlemlere aykırı düşen, onlarla uyuşmayan gözlemler olarak ifade edilmektedir (Barnett ve Lewis, 1995). Barnett ve Lewis'e göre uç değerler üç nedenden kaynaklanabilir;

- Ölçme hatası (*Okuma, kaydetme veya hesaplama hatası gibi nedenden kaynaklanması*)
- Uygulama hatası (*İncelenen evrenin içinde bulunmayan bireylerin örnekleme görülmesi*)
- Doğal çeşitlilik (*Modelden kaynaklanan ve verilerin doğal çeşitlilik göstermesi*)

Barnett ve Lewis'e (1995) göre bütün hatalı veriler uç değeri göstermediği gibi bütün uç değerler de hatalı değildir. Bu nedenle veri setindeki uç değerlerin kaynağı iyi araştırılmalıdır.

Uç değerler için üç yöntem uygulanabilir. Bunlar; silme, görmezden gelme (dokunmama) ve veri dönüşüm işlemleridir. Silme işlemi örneklem büyüklüğünün değişmesine neden olacağından dağılımın normalliğini etkileyecektir. Üstelik uç değerlerin silinmesi, o uç değerlere ait gözlemleri de ortadan kaldıracacağı için değişiklerdeki uç değerde bulunan adayların sonuçta yer almamasına neden olur (Johnson ve Wichern, 2002). Bu durumda örnekleme zarar vermesi olası bir durumdur.

Uç değer belirleme yöntemleri tek değişkenli uç belirleme yöntemleri olarak z puanına dönüştürme ve grafiksel teknikler ile çok değişkenli uç belirleme yöntemleri olarak Mahalanobis uzaklığı (Mahalanobis, 1937), Cook uzaklığı (Cook, 1977) ve kaldıraç noktası (Wisnowski, 2001) olmak üzere iki temel gruba ayrılır.

Bu çalışmada tek değişkenli uç belirleme yöntemlerinden z puanına dönüştürme yolu ile analiz yapılmıştır. Z puanı, standart sapma kullanılarak gözlenen değişkenin ortalama değere kaç standart sapma uzakta olduğunu gösterir. Verinin uç değer farz edilebilmesi için z puanı değerinin -3'den küçük veya +3'den büyük olması beklenir. Yapılan analizler sonucunda bu çalışmada z puanı değerinin ± 3 aralığı aşmadığı gözlenmiştir.

Çok değişkenli uç değer belirlemede kullanılan bir yöntem olan Mahalanobis uzaklığı da bu araştırmada test edilmiştir. *Mahalanobis uzaklığı*, geniş örneklerde değişken sayısına bağlı olarak serbestlik derecesi ile ki kare dağılım gösterir (Johnson ve Wichern, 2002). Mahalanobis uzaklık değeri, puan setleri ile bütün değişkenlerin örneklem ortalaması arasındaki standart sapma uzaklığını gösterir. Analizin yorumlanmasında değişken sayısına eşit serbestlik derecesi ile ki kare dağılımına bakılır (Çokluk ve diğerleri, 2012; Kline, 2011). Tablo değerini ($p < 0,01$ ve serbestlik derecesi=değişken sayısı) aşan veriler uç değer kabul edilir ve veri setinden çıkartılabilir. Bu amaçla Mahalanobis uzaklığı analizi sonucunda $p < ,001$ ölçütüne göre serbestlik derecesi 9'a karşılık gelen 27,877'den büyük (İngiltere örnekleminde 162, İrlanda'dan 172, Türkiye'den 255 ve ABD'den 210) veriler silinmiştir.

Faktör Analizi

Faktör analizi yapılarının uygunluğunu belirlemek amacıyla aşamalı faktör analizi yapılmıştır. Bu amaçla verilerin için Bartlett küresellik ve Kaiser-Meyer- Olkin (KMO) değerlerine bakılmıştır. Bartlett ve KMO analizleri değişkenler arası ilişkinin gücünü ölçer. KMO örneklem yeterliği ölçütü, gözlenen korelasyon katsayıları büyüklüğü ile kısmi korelasyon katsayılarının büyüklüğünü karşılaştırır. KMO testi, faktör çıkarmak için örneklem verilerinin uygunluğunu gösterir (Şencan, 2005). Korelasyon ve kısmi korelasyon istatistiklerine dayalı olarak 0 ile 1 aralığında değer alır. KMO değerinin yüksek çıkması ölçekte yer alan her bir değişkenin diğer değişkenler tarafından mükemmel bir şekilde tahmin edilebileceğini gösterirken (1'e yakın olması); değerlerin sıfır ya da sıfıra yakın olması ise korelasyon dağılımında dağınıklık olduğunu ve bu değerlere dayalı yorumlar yapılmaması gerektiği bildirilmektedir. Kaiser (1974) KMO değerlerine ilişkin 0,90 ve üzeri değerler mükemmel; 0,81-0,90 değerleri iyi; 0,80-0,71 değerleri orta; 0,70-0,61 değerleri zayıf; 0,60-0,50 değerleri kötü iken; 0,50 ve altındaki değerlerin ise faktör analizi için uygun olmadığı ve buna dayalı yorumlar yapılmaması gerektiği belirtmektedir. (Akt: Şencan, 1998; Şencan, 2005; Çokluk Şekercioğlu ve Büyükoztürk, 2012). Benzer şekilde, Field (2000) KMO alt sınır değerinin 0,50 olması gerektiğini aksi durumda veri setinin faktörlenemeyeceğini belirtmiştir.

Elde edilen KMO sonuçlarına göre (KMO değerleri 0,90'dan büyük) analiz için veri setinin uygun olduğu belirlenmiştir. Bu bilgilere ait değerlere Tablo 9'da yer verilmiştir.

Tablo 9

Ülkelere göre KMO ve Bartlett değerleri

KMO and Bartlett's Test		İngiltere	İrlanda	ABD	Türkiye
KMO değerleri		0,910	0,916	0,904	0,961
Bartlett küresellik testi	<i>Ki-kare değeri</i>	38561,571	47072,948	45476,135	45341,377
	<i>Anlamlılık değeri</i>	,000	,000	,000	,000

Tablo 9'da KMO değerlerinin yanında Bartlett küresellik değerlerine de yer verilmiştir. Bu araştırmada Bartlett küresellik testi sonuçları incelendiğinde ($X^2_{(İRL)}=47072,948$, $X^2_{(İNG)}=38561,571$, $X^2_{(TÜR)}=45341,377$, $X^2_{(ABD)}=45476,135$; $p=0,000<0,05$) ki-kare istatistiğinin her bir örneklem için anlamlı çıktığını söylemek mümkündür. Bu bağlamda, Bartlett küresellik testinin anlamlı olması verilerin çok değişkenli normal dağılımdan geldiğini ispatlayarak normallik ve doğrusallık sayıltılarının karşılanması (Çokluk Şekercioğlu ve Büyüköztürk 2012) ile faktör analizi yapmaya yeterli bir ilişkinin olduğu görülmektedir.

Faktör sayısının belirlenmesi. Bahsi geçen sayıltıların karşılanması nedeniyle maddeler arasında ilişkiyi ortaya koyabilecek faktör sayısını tespit etmek amacıyla yamaç birikinti grafiği, özdeğer ve varyans yüzdelerinden yararlanılmıştır (Çokluk, Şekercioğlu ve Büyüköztürk, 2012).

Özdeğer (Eigen Value), faktör yüklerinin karelerinin toplamı ile ifade edilmektedir. Faktörler tarafından açıklanan varyansların oranlarının hesaplanmasında ve faktör sayısına karar vermede kullanılan katsayıdır (Büyüköztürk, 2002). Özdeğer arttıkça açıklanan faktör varyansı da artar. Aksi durumda, yani açıklanan varyansın düşük çıkması o ölçekten elde edilen bilginin de o denli az olduğunu gösterir (Tabachnick ve Fidell, 2007). Özdeğer istatistiğinde 1 ve 1 den büyük olan faktörler anlamlı kabul edilir (Çokluk, Şekercioğlu ve Büyüköztürk 2012).

Faktör analizinde üç çeşit varyans açıklamasından söz edilmektedir. Ortak faktörler tarafından açıklanabilen *Ortak Faktör Varyansı* değişkende gözlenen varyansı ifade eden özgün varyans (specific variance) ve veri setindeki varyansın

açıklanmayan bölümünü tanımlayan hata varyansıdır (error variance) (Kline, 2011). Testin güvenilirliğini yorumlarken ortak varyans ile özgün varyansın toplamından yararlanılır. Ortak faktör varyansı, maddelerin çoklu regresyonunun karesi ile hesaplanmaktadır ve bu değer yüksek çıkması açıklanan toplam varyansı artırmaktadır (Büyüköztürk, 2002). Her bir faktör tek başına toplam varyans üzerinden maddenin en az %5'ini açıklayabilmelidir. Analiz neticesinde varyans oranlarının yüksek olması ölçeğin faktör yapısının gücünü göstermektedir. Bu bağlamda tek faktörlü ölçeklerde açıklanan varyans yüzdesi 30 ve üzeri yeterli görülürken (Büyüköztürk, 2013), çok faktörlü ölçeklerde açıklanan varyans yüzdesinin 40 ile 60 arasında olması beklenir (Tavşancıl, 2010). Genel kanı açıklanan varyansın, toplam varyans üzerinden %50'sini aşabilen bir analiz geçerli bir analiz kabul edilir.

Bu çalışma için ülkelere göre yapılan faktör analizine ait özdeğer ve toplam açıklanan varyans yüzdelerine dair değerlere Tablo 10'da yer verilmiştir.

Tablo 10

Ülkelere ait açıklanan özdeğerler ve toplam varyans yüzdeleri

Boyut	İngiltere		İrlanda		ABD		Türkiye	
	Özdeğer	Toplam Açıklanan % Varyans	Özdeğer	Toplam Açıklanan % Varyans	Özdeğer	Toplam Açıklanan % Varyans	Özdeğer	Toplam Açıklanan % Varyans
1	5,5	61,11	5,75	63,89	5,152	57,241	4,468	49,643
2	1,994	83,262	1,83	84,225	2,469	84,43	3,101	84,102
3	0,344	87,089	0,299	87,55	0,29	87,898	0,364	88,143
4	0,267	90,061	0,249	90,322	0,266	90,851	0,235	90,755
5	0,23	92,62	0,244	93,037	0,21	93,187	0,216	93,156
6	0,221	95,071	0,198	95,241	0,178	95,169	0,177	95,123
7	0,173	96,99	0,163	97,056	0,154	96,876	0,174	97,059
8	0,152	98,683	0,148	98,702	0,15	98,547	0,156	98,791
9	0,119	100	0,117	100	0,131	100	0,109	100

Tablo 10 incelendiğinde Türkiye, ABD, İngiltere ve İrlanda örneklemlerinde özdeğeri 1'den büyük 2 boyuta rastlanmaktadır. Ayrıca her bir açıklanan varyansın tek başına %5'i aştığı ve iki faktörün varyansa olan katkısının Türkiye örnekleminin

%84,102, ABD örnekleminin %84,430, İngiltere örneklemini %83,262 ve İrlanda örneklemini için %84,225 olduğu gözlenmektedir.

Faktörlerin Rotasyonu, faktör yapılarını daha kolay yorumlayabilmek adına faktörlere döndürme işlemi uygulanmaktadır. Döndürme araştırmanın amacına göre belirlenmelidir. İki temel döndürme yöntemi bulunmaktadır. *Dik açılı döndürme (Orthogonal rotation)*; bu tip döndürme işlemlerinde gizli yapıların ilişkisiz olduğu farz edilir. Faktör çıkarma işleminde eksenler dik açı ile döndürülerek birbiriyle ilişkisiz olacak şekilde farklı yerlerde konumlandırılır. Faktörler arasında korelasyonun sıfır olduğu bu yöntemde faktörler arasındaki gerçek ilişkileri tam olarak açıklayamaz (Şencan, 2005). *Eğik açılı döndürme (Oblique rotation)*; bu yöntemde faktörler arasında ilişki olduğu varsayılır. Faktörlerin beklenenden çok daha fazla ilişkili hale getirmek amaçlanır. Aynı zamanda araştırmacı faktörler arasında bir ilişki olup olmadığından emin değilse de bu yöntem tercih edilebilir (Akt:Şencan, 2005). Bu amaçla tutum ölçeklerinde faktörlerin birbiriyle ilişkileri daha gerçekçi bir şekilde açıklayacağı düşünüldüğünden ayrıca kullanımı daha pratik olduğundan eğik döndürme yöntemlerinden promax yöntemi tercih edilmiştir (Büyüköztürk, 2002). Döndürülmüş faktör yükleri Tablo 11’de gösterilmektedir.

Tablo 11

Ülkeler bazında maddelere ait faktör yük değerleri

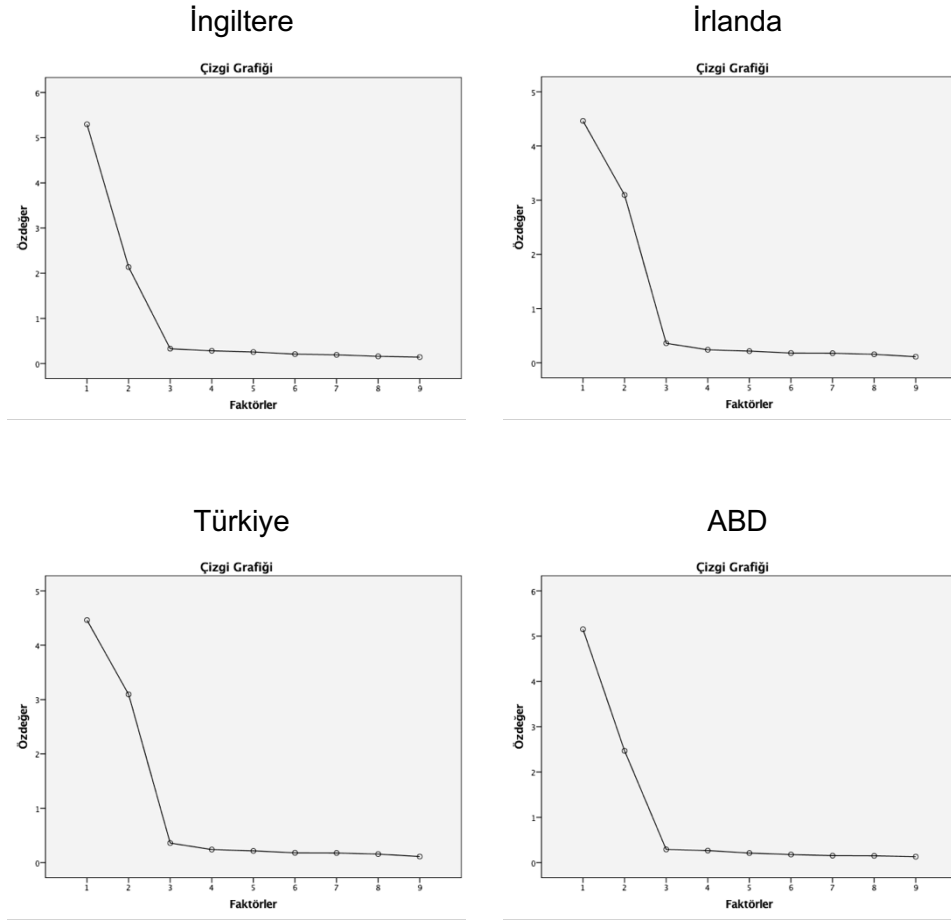
Boyut	Maddeler	İngiltere		İrlanda		ABD		Türkiye	
		Faktörler	Faktörler	Faktörler	Faktörler	Faktörler	Faktörler	Faktörler	Faktörler
		1	2	1	2	1	2	1	2
ST94	Q01	0,889	-	0,902	-	0,917	-	0,91	-
	Q02	0,882	-	0,906	-	0,888	-	0,927	-
	Q03	0,95	-	0,934	-	0,939	-	0,934	-
	Q04	0,94	-	0,939	-	0,937	-	0,941	-
	Q05	0,93	-	0,908	-	0,922	-	0,935	-
ST113	Q01	-	0,898	-	0,906	-	0,906	-	0,895
	Q02	-	0,935	-	0,938	-	0,94	-	0,922
	Q03	-	0,904	-	0,906	-	0,931	-	0,913
	Q04	-	0,866	-	0,893	-	0,899	-	0,872

* 0,30’un altındaki değerler (-) şeklinde gösterilmiştir.

Tablo 11 incelendiğinde, Türkiye, ABD, İngiltere ve İrlanda örneklerinin 2 boyutta faktörleştiği saptanmıştır. Fen öğrenmeden zevk alma boyutundaki maddelerin tamamı bir faktör altında toplanırken (Türkiye'ye de 0,941 ile 0,910; ABD'de 0,939 ile 0,888, İngiltere'de 0,950 ile 0,882 ve İrlanda'da 0,939 ile 0,902 arasında faktör yüklerine sahipken), fen öğreniminde araçsal güdülenme boyutundaki maddelerin tamamı ise başka bir faktör altında (Türkiye'ye de 0,922 ile 0,872; ABD'de 0,940 ile 0,899, İngiltere'de 0,935 ile 0,866 ve İrlanda'da 0,938 ile 0,893 arasında faktör yüklerinde) toplandığı gözlenmiştir. Her iki boyuta ait faktör yük değerlerinin kabul düzeyi olan 0,32'nin üzerinde olduğu görülmektedir (Tabachnick ve Fidell, 2007).

Açımlayıcı faktör analizlerinde kullanılan ve Cattell (1966) tarafından geliştirilmiş olan *Yamaç Birikinti Grafiği (Scree Plot)* faktör sayısı kestirimi için kullanılmaktadır. Grafikte bulunan y eksenini özdeğer büyüklüğü ve x eksenini de bileşen sayısını vermektedir. Bu grafikler faktörlerin varyansa olan katkıları doğrultusunda kırılmaları hakkında bilgi verir. Ciddi bir düşüşten sonra düzleşme gözlenen kısım, kırılma noktasından sonraki faktörlerin varyansa olan etkisinin önemsiz olmayacak büyüklükte olduğunu ifade eder. Kırılma noktasına kadar olan aralık sayıları ise faktör sayısını verir (Koçak, Çokluk ve Kayri, 2016). İlk faktör en yüksek varyansı açıklarken, son faktör en düşük varyansı açıklamaktadır. Kırılma noktaları arasındaki değişimler gözlemlenerek yorumlar yapılır. Bu durumda faktör yapısını belirlemede araştırmacıların kararlarında sübjektif bir bakış açısı sağlamaktadır (Velicer ve Fava, 2000).

Yapılan analiz sonucunda ülkeler bazında maddelere ait yamaç birikinti grafikleri Şekil 1'de gösterilmiştir.



Şekil 1. Dört ülke için yamaç birikinti grafikleri

Şekil 1 incelendiğinde İngiltere, İrlanda, ABD ve Türkiye örneklerinde yüksek ivmeli düşüşün üçüncü noktadan sonra azaldığı görülmektedir. Üçüncü noktadan sonra bileşenlerin varyansa yaptığı katkının da azaldığı da ayrıca gözlenmektedir. Burada Türkiye, ABD, İngiltere ve İrlanda verilerinin kırılması üçüncü noktada (2 faktör) gerçekleşmiştir. Bu durum, faktör yük değerlerini gösteren Tablo 11 ile uyum göstermektedir.

Bileşenlerin arttıkça özdeğerin azaldığını gösteren yamaç birikinti grafiğinde, eğimin artık kaybolmaya başladığı noktaya kadar olan bileşen sayısı faktör sayısını vermektedir. Tüm bu bilgiler ışığında, özdeğer ve varyans yüzdeleri ile yamaç birikinti grafiklerinden elde edilen veriler doğrultusunda analizin iki faktör için yapılmasına karar verilmiştir.

Betimsel İstatistikler

Kişilik, ilgi ve tutum gibi veya envanterlerin cevapları derecelendirilmiş ölçeklerin güvenilirliğini hesaplamasında tercih edilen yöntemlerden birisi *Cronbach Alfa güvenilirlik katsayısıdır*. Cronbach (1951) tarafından geliştirilen bu yöntem, maddelerin doğru yanlış olarak puanlanmadığı yani 1-3, 1-4, 1-5 gibi sıralı puanlamaların olduğu durumlarda kullanılması uygun olan iç tutarlılık tahmin yöntemidir (Akt:Kılıç, 2016).

Cronbach Alfa (α) güvenilirlik katsayısı 0 ile +1 arasında değer alabilir. Değerin 1'e yakın olması testin güvenilirliğini göstermektedir. Fakat maddeler arası korelasyonun negatif olması durumunda α değeri de negatif değer alır. Bu durumda testin güvenilirliğini olumsuz etkiler

Bu çalışmada güvenilirlik düzeyini hesaplamak için Cronbach Alfa (α) güvenilirlik yönteminden yararlanılmıştır. Hesaplanan Cronbach Alfa güvenilirlik katsayılarına ilişkin sonuçlara Tablo 12'de yer verilmiştir.

Tablo 12

Ülkelere ait Cronbach Alpha (α) güvenilirlik katsayıları

	İngiltere	İrlanda	ABD	Türkiye
Fen öğrenmeden zevk alma	0,954	0,955	0,955	0,961
Fen öğreniminde araçsal güdülenme	0,924	0,934	0,939	0,921

Tablo 12'de ülkelere ait güvenilirlik katsayılarını yorumlamak için elde edilen α değerlerinin hangi aralıkta olduğunun belirlenmesi gerekir. Bu bağlamda, Özdamar (2002) tarafından geliştirilen güvenilirlik katsayılarına ilişkin ölçüt değerler şu şekildedir; $0,00 < \alpha < 0,40$ için ölçek güvenilir değildir, $0,41 < \alpha < 0,60$ için ölçek düşük düzeyde güvenilir, $0,61 < \alpha < 0,80$ için ölçek orta düzeyde güvenilir ve $0,81 < \alpha < 1,00$ için ise ölçek yüksek düzeyde güvenilirdir. Analizde kullanılacak tüm ülke verilerine göre Cronbach Alpha değerlerinin ($\alpha > 0,90$) yüksek olması, yüksek düzeyde güvenilir bir test olduğu şeklinde yorumlanabilir (Kalaycı, 2018).

Tablo 13

Ülkelere göre betimsel istatistik analizi

Boyut	Madde	İngiltere		İrlanda		ABD		Türkiye	
		S.s.	Ort.	S.s.	Ort.	S.s.	Ort.	S.s.*	Ort.**
Özyeterlik (ST94)	Q01	0,946	2,62	0,855	2,69	0,817	2,83	0,926	2,65
	Q02	0,915	2,65	0,882	2,59	0,834	2,61	0,905	2,67
	Q03	0,923	2,63	0,821	2,78	0,821	2,79	0,913	2,64
	Q04	0,92	2,78	0,829	2,91	0,795	2,9	0,91	2,77
	Q05	0,926	2,78	0,859	2,88	0,829	2,88	0,911	2,78
Motivasyon (ST113)	Q01	0,82	1,95	0,829	1,96	0,781	1,96	0,804	1,95
	Q02	0,799	1,99	0,894	2,1	0,826	2,08	0,789	1,99
	Q03	0,858	2,05	0,841	1,99	0,81	2,06	0,841	2,03
	Q04	0,898	2,13	0,867	2,07	0,834	2,13	0,876	2,11

*S.s.:Standart sapma, **Ort:Ortalama değer

Tablo 13'te her bir ülke ve madde için ortalama ve standart sapma değerleri verilmiştir. Türkiye verisinde ortalama değerler 1,95 ile 2,78; İngiltere verisinde 1,95 ile 2,78; İrlanda verisinde 1,96 ile 2,91 arasında değişirken ABD verisinde 1,96 ile 2,88 aralığında değişmektedir. Katılımcıların maddelere verdikleri cevapların ortalaması incelendiğinde değerlerin birbirine yaklaşık olduğu gözlenmektedir. Ayrıca maddelerin standart sapmalarına bakıldığında İngiltere'de 0,799 ile 0,946; Türkiye'de 0,789 ile 0,926; ABD'de 0,781 ile 0,837 ve İrlanda'da 0,821 ile 0,894 aralığında değer aldığı görülmüştür.

Bölüm 4

Bulgular ve Yorumlar

Bu bölümde araştırmanın amaçları doğrultusunda elde edilen bulgulara ve yorumlara yer verilmiştir. Alt problemler için GMH, OLR ve poly-SIBTEST yöntemleri kullanılarak tutum maddelerinde DMF varlığı araştırılmıştır. İngiltere örneklemini sabit tutularak İrlanda (alt problem 1: aynı dil benzer kültür), ABD (alt problem 2: aynı dil farklı kültür) ve Türkiye (alt problem 3: farklı dil farklı kültür) örneklemleriyle karşılaştırılarak analizler yapılmıştır.

Bu çalışmada OLR ve poly-SIBTEST analizleri için R programı (LORDIF, Haven, MIRT paketleri) ve GMH analizi için ise GMHDIF programları kullanılmıştır.

Alt Problem 1a'ya İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-İrlanda uygulamasında aynı dil benzer kültüre göre OLR yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

İngiltere ve İrlanda verilerine ait toplam 9 maddeye ait OLR analiz sonuçları Tablo 14'te verilmiştir. Analizlerde kullanılan OLR yöntemine ait tabloların çıktıları EK-C'de yer almaktadır.

Tablo 14

İngiltere-İrlanda verisine ait OLR analiz sonuçları

Boyut	Madde	X^2_{12}	X^2_{13}	X^2_{23}	β_1	Mc Fadden R^2_{12}	Mc Fadden R^2_{13}	Mc Fadden R^2_{23}	DMF Türü	DMF Düzeyi
ST94	Q01	0	0	0,814	0,02	0,007	0,007	0	TB	A
	Q02	0,208	0,04	0,028	0	0	0	0	-	-
	Q03	0	0	0,013	0,02	0,004	0,004	0	TB	A
	Q04	0	0	0,075	0,03	0,006	0,006	0	TB	A
	Q05	0	0	0,438	0,01	0,003	0,003	0	TB	A
ST113	Q01	0	0	0,89	0,003	0,001	0,001	0	TB	A
	Q02	0,508	0,138	0,06	0	0	0	0	-	-
	Q03	0,04	0,118	0,807	0,008	0	0	0	-	-
	Q04	0	0	0,89	0,003	0,001	0,001	0	TB	A

Tablo 14'te, her bir madde için McFadden Model3-Model1 R² farkları dikkate alınarak DMF incelemeleri yapılmıştır. Analiz sonucunda ST94 fen öğrenmeden zevk alma maddesine ait 1., 3., 4. ve 5. maddelerde DMF tespit edilirken, ST113 fen öğrenmede araçsal güdülenme maddesine ait 1. ve 4. maddelerde DMF tespit edilmiştir. Benzer çalışmalarda (Choi, Gibbons ve Crane, 2011; Crane, Belle ve Larson, 2004) DMF düzeyini kestirmede yararlanılan ölçütlere (Cohen, 1988) bakıldığında Tablo 14'teki tüm maddelerin (<0,020) A düzeyinde DMF olduğu gözlenmektedir. Ayrıca TBO DMF tespitinde kullanılan McFadden Model3-Model2 farkları ile elde edilen Bonferroni değerleri (ST94: 0,002; ST113: 0,0025) karşılaştırıldığında (Crane ve arkadaşları, 2006) İngiltere ve İrlanda verilerine ait tüm DMF'li maddelerin TB olduğu görülmektedir.

Alt Problem 1 b'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-İrlanda uygulamasında aynı dil benzer kültüre göre poly-SIBTEST ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

İngiltere ve İrlanda verisinde fen öğrenmekten zevk alma (ST 94) ve fen öğreniminde araçsal güdülenme (ST 113) maddelerinin poly-SIBTEST sonuçları 15'te gösterilmiştir. Analizlerde kullanılan poly-SIBTEST yöntemine ait tabloların çıktıkları EK-D'de yer almaktadır.

Tablo 15

İngiltere-İrlanda verisine ait poly-SIBTEST analizi sonuçları

	Madde	β_{uni}	SE	X ²	p	DMF Düzeyi	DMF Türü	Avantajlı Grup
ST 94	Q01	0,129	0,011	151	0,000	C	TB	İngiltere
	Q02	-0,028	0,011	6	0,015	A	TB	İrlanda
	Q03	0,062	0,008	56	0,000	B	TB	İngiltere
	Q04	-0,093	0,008	123	0,000	C	TB	İrlanda
	Q05	-0,075	0,009	68	0,000	B	TB	İrlanda
ST 113	Q01	-0,028	0,012	5,777	0,016	A	TB	İrlanda
	Q02	0,037	0,012	9,91	0,002	A	TB	İngiltere
	Q03	-0,007	0,012	0,315	0,058	-	-	-
	Q04	0,030	0,013	4,949	0,026	A	TB	İngiltere

R paketi MIRT programı ile hesaplanan Tablo 15'teki verilen 8 madde için β değerleri anlamlıdır (Chalmers, 2012). Fen öğrenmeden zevk alma (ST 94) testinde poly-SIBTEST sonucunda elde edilen β değerleri incelendiğinde 1 ve 4. maddelerde C, 3 ve 5. B ve 2. maddede ise A düzeyde DMF belirlenmiştir. Aynı zamanda DMF türlerine bakıldığında DMF içeren maddelerin tamamında TB DMF'ye rastlanmıştır. Fen öğrenmede araçsal güdülenme (ST113) boyutunda ise maddeler ihmal edilebilir düzeyde DMF göstermektedir. Ayrıca DMF'li maddeler incelendiğinde, fen öğrenmeden zevk alma 1. ve 3. maddeleri ve fen öğrenmede araçsal güdülenme 2. ve 4. maddeler İngiltere grubu, kalan DMF'li maddeler ise İrlanda lehine çalıştığı gözlenmektedir.

Alt Problem 1c'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-İrlanda uygulamasında aynı dil benzer kültüre göre GMH ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

Fen öğrenmeden zevk alma ve fen öğrenmede araçsal güdülenme alt boyutlarında İngiltere ve İrlanda örneklemelerine ait GMH analiz sonuçları Tablo 16'da gösterilmiştir. Analizlerde kullanılan GMH (GMHDIF ve DIFAS program çıktıları) yöntemine ait tabloların çıktıları EK-E'de yer almaktadır.

Tablo 16

İngiltere-İrlanda verisine ait GMH analizi sonuçları

Boyutlar	Madde	X^2_{GMH}	p^*	L-A LOR**	Avantajlı Grup
ST94	Q01	154,8080	0,0000	0,784	İngiltere
	Q02	30,4224	0,0000	-0,232	İrlanda
	Q03	63,4329	0,0000	0,592	İngiltere
	Q04	136,1824	0,0000	-0,825	İrlanda
	Q05	70,5869	0,0000	-0,604	İrlanda
ST113	Q01	15,7465	0,0013	-0,22	İrlanda
	Q02	9,2948	0,0256	0,121	-
	Q03	2,859	0,4139	-0,066	-
	Q04	6,5562	0,0875	0,139	-

* p ölçüt değeri olarak Bonferroni hesaplaması kullanılmıştır. **DIFAS programı ile hesaplanmıştır.

Benzer çalışmalarda yapıldığı gibi (Pasca, Coello, Arigones ve Frantz 2018) GMHDIF programı için önerilen Bonferroni değeri (Fidalgo, 2011) hesaplandığında ($ST94:0,01$; $ST113:0,015$), Tablo 16'da fen öğrenmeden zevk alma (ST94) maddelerinin tamamında DMF tespit edilirken, fen öğrenmede araçsal güdülenmede (ST113) 1. maddede DMF tespit edilmiştir. DMF içeren maddelerde L-ALOR değerine bakıldığında, fen öğrenmeden zevk alma boyutunda 1. ve 3. maddeler İngiltere, kalan DMF'li maddeler ise İrlanda lehine çalışmıştır.

Alt Problem 1d'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-İrlanda uygulamasında aynı dil benzer kültüre göre OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizlerin sonuçları birbiriyle uyumlu mudur?

İngiltere ve İrlanda verisinde OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucu DMF gösteren maddeler Tablo 17'de verilmiştir.

Tablo 17

İngiltere-İrlanda verisinde OLR, poly-SIBTEST ve GMH yöntemlerinin karşılaştırılması

Boyut	Madde	OLR DMF varlığı	OLR DMF düzeyi	OLR DMF Türü	poly-SIBTEST DMF Varlığı	poly-SIBTEST DMF Türü	poly-SIBTEST DMF Düzeyi	Avantajlı Grup (poly-SIB)	GMH DMF varlığı	Avantajlı Grup (GMH)
ST94	Q01	var	A	TB	var	TB	C	İngiltere	var	İngiltere
	Q02	-	-	-	var	TB	A	İrlanda	var	İrlanda
	Q03	var	A	TB	var	TB	B	İngiltere	var	İngiltere
	Q04	var	A	TB	var	TB	C	İrlanda	var	İrlanda
	Q05	var	A	TB	var	TB	B	İrlanda	var	İrlanda
ST113	Q01	var	A	TB	var	TB	A	İrlanda	var	İrlanda
	Q02	-	-	-	var	TB	A	İngiltere	-	-
	Q03	-	-	-	-	-	-	-	-	-
	Q04	var	A	TB	var	TB	A	İngiltere	-	-

Tablo 17 incelendiğinde 9 maddeden GMH ve OLR yöntemleri ile 6 maddede, poly-SIBTEST yöntemiyle (4'ü önemli düzeyde olmak üzere) 7 madde DMF

göstermektedir. Madde düzeyleri incelendiğinde fen öğrenmeden zevk alma boyutunda poly-SIBTEST yöntemi ile 1. ve 4. maddelerde C, 3. ve 5. maddelerde B düzeyinde DMF belirlenmiştir. Ayrıca önemli düzeyde DMF içeren bu maddeler diğer iki yöntemde de DMF içermekte olduğu gözlenmiştir. Maddelerin GMH ve poly-SIBTEST yöntemlerine göre avantaj sağladığı gruplara bakıldığında ise fen öğrenmeden zevk alma alt boyutunda 1. ve 3. maddelerin İngiltere; kalan ortak DMF içeren maddeler (fen öğrenmeden zevk alma 2., 4. ve 5. ile araçsal güdülenme boyutunda 1. maddeler) ise İrlanda lehine çalıştığı gözlenmiştir. Bu sonuçlar bakımından iki yöntem arasında uyumluluk gözlenmiştir. Ayrıca, DMF türleri incelendiğinde, OLR ve poly-SIBTEST yöntemleri arasında ortak DMF içeren maddelerde tutarlılık görülmektedir.

Alt Problem 1e'ye İlişkin Bulgular ve Yorumlar

OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucunda her üç yönteme göre DMF gösteren maddeler nelerdir?

Tablo 17'ye bakıldığında her üç yönteme göre DMF içerdiği belirlenen beş maddede DMF gösteren maddeler ortaktır. Bunlar, fen öğrenmeden zevk alma boyutunda 1.,3.,4. ve 5. maddeler ile araçsal güdülenme boyutunda 1. maddedir. Ayrıca fen öğrenmeden zevk alma boyutundaki ortak DMF'li maddeler poly-SIBTEST yöntemine göre önemli düzeyde DMF göstermektedir. İngiltere-İrlanda veri analizinde üç yönteme göre DMF içermeyen madde ise araçsal güdülenme boyutundaki 3. maddedir.

Alt Problem 2a'ya İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-ABD uygulamasında aynı dil farklı kültüre göre OLR yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

İngiltere ve ABD verisine ait 2 alt boyut için LORDIF paketi ile yapılan OLR sonuçları Tablo 18'de verilmiştir.

Tablo 18

İngiltere-ABD verisine ait OLR analiz sonuçları

Boyut	Madde	X^2_{12}	X^2_{13}	X^2_{23}	β_1	Mc Fadden R^2_{12}	Mc Fadden R^2_{13}	Mc Fadden R^2_{23}	DMF Türü	DMF Düzeyi
ST94	Q01	0,196	0,02	0,013	0	0	0	0	-	-
	Q02	0,938	0,196	0,071	0	0	0	0	-	-
	Q03	0	0	0	0,055	0,009	0,011	0,001	TB	A
	Q04	0	0	0,052	0,002	0,001	0,001	0	TB	A
	Q05	0,005	0,017	0,757	0	0	0,001	0	TB	A
ST113	Q01	0,53	0	0	0	0	0,001	0,001	TBO	A
	Q02	0	0	0	0,019	0,005	0,008	0,003	TB	A
	Q03	0	0	0	0,008	0,003	0,003	0,001	TB	A
	Q04	0,321	0	0	0	0	0,001	0,009	TBO	A

Tablo 18'de 9 madde için McFadden R^2 (Model1-Model3) farkı incelendiğinde fen öğrenmeden zevk alma (ST94) boyutunda 3., 4. ve 5. maddelerde DMF tespit edilirken, fen öğrenmede araçsal güdülenme (ST113) maddelerin tamamında DMF tespit edilmiştir. Ayrıca DMF türlerine (Model3-Model2) bakıldığında sadece fen öğrenmede araçsal güdülenme boyutunda 1. ve 4. maddelerde TBO DMF'ye rastlanırken kalan maddelerde TB DMF saptanmıştır. DMF etki büyüklükleri incelendiğinde ise tüm değerlerin çok küçük olması nedeniyle ($<0,020$) ihmal edilebilir düzeyde olduğu söylenebilir.

Alt Problem 2b'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-ABD uygulamasında aynı dil farklı kültüre göre poly-SIBTEST ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

İngiltere ve ABD verisine ait fen öğrenmeden zevk alma (ST 94) ve fen öğreniminde araçsal güdülenme (ST 113) testine ilişkin poly-SIBTEST analizi sonuçlarına Tablo 19'da yer verilmiştir.

Tablo 19

İngiltere-ABD verisine ait poly-SIBTEST analizi sonuçları

Boyut	Madde	β_{uni}	β_{cro}	SE	X^2	p	DMF Düzeyi	DMF türü	Avantajlı Grup
ST 94	Q01	-0,021	0,021	0,010	4,278	0,039	A	TB	ABD
	Q02	-0,015	0,015	0,011	1,853	0,173	-	-	-
	Q03	0,095	0,095	0,008	131,044	0,000	C	TB	İngiltere
	Q04	-0,039	0,039	0,008	21,645	0,000	A	TB	ABD
	Q05	-0,030	0,030	0,009	11,163	0,001	A	TB	ABD
ST 113	Q01	0,010	0,015	0,012	0,794	0,373	-	-	-
	Q02	0,116	0,116	0,011	106,470	0,000	C	TB	İngiltere
	Q03	-0,079	0,079	0,011	52,707	0,000	B	TB	ABD
	Q04	-0,015	0,029	0,013	1,296	0,255	-	-	-

Tablo 19'da fen öğrenmeden zevk alma (ST 94) boyutunda poly-SIBTEST sonucunda elde edilen β değerleri incelendiğinde fen öğrenmeden zevk alma 3. maddesi ve fen öğrenmede araçsal güdülenme 2. maddenin C DMF içerdiği, fen öğrenmede araçsal güdülenme 3. maddesinin B düzeyde DMF içerdiği gözlenmektedir. Ayrıca DMF tespit edilen, fen öğrenmeden zevk alma 3. ve araçsal güdülenmede 2. maddelerinin İngiltere grubuna avantaj sağladığı analizde görülmektedir. DMF türleri incelendiğinde ise DMF'li maddelerin tamamında TB DMF belirlenmiştir.

Alt Problem 2c'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-ABD uygulamasında aynı dil farklı kültüre göre GMH ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

Fen öğrenmeden zevk alma (ST94) ve fen öğrenmede araçsal güdülenme (ST113) alt boyutlarında İngiltere ve İrlanda örneklemlerine ait GMH analiz sonuçları Tablo 20'de gösterilmiştir.

Tablo 20

İngiltere-ABD verisine ait GMH analizi sonuçları

Boyut	Madde	X^2_{GMH}	p*	L-A LOR**	Avantaj Sağladığı Grup
ST94	Q01	15,3383	0,0015	-0,148	-
	Q02	17,6073	0,0005	-0,024	ABD
	Q03	167,7465	0,0000	0,894	İngiltere
	Q04	27,2388	0,0000	-0,355	ABD
	Q05	16,8652	0,0008	-0,288	ABD
ST113	Q01	13,9241	0,0030	0,089	İngiltere
	Q02	106,8313	0,0000	0,578	İngiltere
	Q03	60,123	0,0000	-0,515	ABD
	Q04	6,5779	0,0866	-0,103	-

*p ölçüt değeri olarak Bonferroni hesaplaması kullanılmıştır. **DIFAS programı ile hesaplanmıştır.

Tablo 20 incelendiğinde fen öğrenmeden zevk alma boyutuna ait 2.,3.,4. ve 5. maddelerde DMF varlığı söz konusu iken fen öğreniminde araçsal güdülenme (ST 113) 4. madde hariç tüm maddelerde DMF tespit edilmiştir. L-A LOR değerlerine bakıldığında ise fen öğrenmeden zevk alma 3. madde, araçsal güdülenmede 1. ve 2. maddelerinin İngiltere grubuna avantaj sağladığı görülmektedir.

Alt Problem 2d'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-ABD uygulamasında aynı dil farklı kültüre göre OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizlerin sonuçları birbiriyle uyumlu mudur?

İngiltere ve ABD verisinde OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucu DMF gösteren maddeler Tablo 21'de verilmiştir.

Tablo 21

İngiltere-ABD verisinde OLR, poly-SIBTEST ve GMH yöntemlerinin karşılaştırılması

Boyut	Madde	OLR DMF varlığı	OLR DMF düzeyi	OLR DMF Türü	poly-SIBTEST DMF Varlığı	poly-SIBTEST DMF Türü	poly-SIBTEST DMF Düzeyi	Avantajlı Grup (poly-SIB)	GMH DMF varlığı	Avantajlı Grup (GMH)
ST94	Q01	-	-	-	var	TB	A	ABD	-	
	Q02	-	-	-	-	-	-	-	var	ABD
	Q03	var	A	TB	var	TB	C	İngiltere	var	İngiltere
	Q04	var	A	TB	var	TB	A	ABD	var	ABD
	Q05	var	A	TB	var	TB	A	ABD	var	ABD
ST113	Q01	var	A	TBO	-	-	-	-	var	İngiltere
	Q02	var	A	TB	var	TB	C	İngiltere	var	İngiltere
	Q03	var	A	TB	var	TB	B	ABD	var	ABD
	Q04	var	A	TBO	-	-	-	-	-	-

Tablo 21 incelendiğinde GMH ve OLR yöntemleriyle 7 maddede, poly-SIBTEST yöntemiyle 6 maddede (3'ü önemli düzeyde olmak üzere) DMF belirlenmiştir. Ayrıca, GMH yönteminde fen öğrenmeden zevk alma 2. madde, OLR yönteminde ise araçsal güdülenme 4. madde diğer yöntemlerden farklı olarak DMF gözlenmiştir. DMF içeren maddelerin GMH ve poly-SIBTEST yöntemlerine göre avantaj sağladığı gruplara bakıldığında ise fen öğrenmeden zevk alma alt boyutunda 3. madde ile araçsal güdülenme alt boyutundaki 2. maddenin İngiltere, fen öğrenmeden zevk alma 4. ve 5. maddeler ile araçsal güdülenme 3. maddenin ise ABD lehine çalıştığı gözlenmiştir. Bu sonuçlar bakımından iki yöntem arasında uyumluluk gözlenmiştir. Aynı zamanda DMF türlerine bakıldığında, poly-SIBTEST ve OLR yöntemlerinde ortak DMF'li maddelerde de uyum gözlenmiştir.

Alt Problem 2e'ye İlişkin Bulgular ve Yorumlar

OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucunda her üç yönteme göre DMF gösteren maddeler nelerdir?

Tablo 21'e bakıldığında her üç yönteme göre fen öğrenmeden zevk alma boyutunda 3.,4. ve 5.maddeler ile araçsal güdülenme boyutunda 2. ve 3. maddeler tüm

yöntemler tarafından DMF tespit etmiştir. Bu ortak DMF'li maddeler içerisinde fen öğrenmeden zevk alma 2. ile araçsal güdülenme 2. ve 3. maddelerinde poly-SIBTEST tarafından önemli düzeyde DMF saptanmıştır. Ayrıca İngiltere-ABD veri analizinde her üç yönetime göre DMF içermeyen ortak maddelere rastlanmamıştır.

Alt Problem 3a'ya İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-Türkiye uygulaması baz alınarak farklı dil farklı kültüre göre OLR yöntemiyle yapılan analizlerde değişen madde fonksiyon göstermekte midir?

İngiltere ve Türkiye verisinde fen öğrenmeden zevk alma ve araçsal güdülenme boyutları için LORDIF paketinde yapılan OLR sonuçlarına Tablo 22'de yer verilmiştir.

Tablo 22

İngiltere-Türkiye verisine ait OLR analiz sonuçları

Boyut	Madde	X^2_{12}	X^2_{13}	X^2_{23}	β_1	Mc Fadden R^2_{12}	Mc Fadden R^2_{13}	Mc Fadden R^2_{23}	DMF Türü	DMF Düzeyi
ST94	Q01	0	0	0,009	0,008	0,004	0,004	0	TB	A
	Q02	0	0	0,031	0,052	0,016	0,016	0	TB	A
	Q03	0	0	0	0,067	0,018	0,021	0,003	TB	B
	Q04	0,883	0	0	0	0	0,001	0,001	TBO	A
	Q05	0	0	0,009	0,005	0,001	0,001	0	TB	A
ST113	Q01	0	0	0,979	0,005	0,001	0,001	0	TB	A
	Q02	0	0	0	0,033	0,001	0,015	0,004	TBO	A
	Q03	0	0	0	0,021	0,005	0,006	0,001	TB	A
	Q04	0	0	0	0,003	0,001	0,002	0,001	TB	A

Tablo 22'ye bakıldığında fen öğrenmeden zevk alma (ST94) ve fen öğrenmede araçsal güdülenme (ST113) maddelerin tamamında DMF'ye rastlanmıştır. DMF düzeyleri incelendiğinde ise yalnız 3. maddenin (>0,020) orta B düzeyinde kalan maddelerin A düzeyinde DMF içerdiği gözlenmektedir. DMF türlerine bakıldığında fen öğrenmeden zevk alma boyutunda 4.madde ile araçsal güdülenme boyutunda 2. maddede TBO DMF şeklinde fonksiyonlaştığı görülmektedir.

Alt Problem 3b'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-Türkiye uygulaması baz alınarak farklı dil farklı kültüre göre poly-SIBTEST ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

İngiltere Türkiye verisine ait fen öğrenmeden zevk alma (ST 94) ve fen öğreniminde araçsal güdülenme testine ait poly-SIBTEST analizi sonuçlarına Tablo 23'te yer verilmiştir.

Tablo 23

İngiltere-Türkiye verisine ait poly-SIBTEST analizi sonuçları

Boyut	Madde	β_{uni}	SE	X ²	p	DMF Düzeyi	DMF Türü	Avantajlı Grup
ST 94	Q01	0,108	0,011	95	0,000	C	TB	İngiltere
	Q02	-0,227	0,011	447	0,000	C	TB	Türkiye
	Q03	0,176	0,009	356	0,000	C	TB	İngiltere
	Q04	-0,015	0,009	3	0,103	-	-	-
	Q05	-0,051	0,010	29	0,000	A	TB	Türkiye
ST 113	Q01	-0,019	0,012	2,513	0,113	-	-	-
	Q02	0,214	0,011	356,208	0,000	C	TB	İngiltere
	Q03	-0,092	0,012	59,704	0,000	C	TB	Türkiye
	Q04	-0,045	0,014	10,602	0,001	A	TB	Türkiye

Tablo 23 incelendiğinde poly-SIBTEST sonucunda 9 madde içinden 7'sinde β değerleri anlamlıdır. Fen öğrenmeden zevk alma alt boyutuna ait 1.,2.,3. ve 5. maddelerde DMF görülürken, fen öğrenmede araçsal güdülenme boyutunda 2.,3. ve 4. maddelerde DMF görülmüştür. ST94 boyutunda 1., 2. ve 3. maddelerde ve ST113 boyutunda 2. ve 3. maddelerde C düzeyde DMF görülürken kalan maddelerde A düzeyde DMF görülmektedir. Ayrıca testin fen öğrenmeden zevk alma 1. ve 3., araçsal güdülenme boyutlarından 2. maddelerin İngiltere grubuna avantaj sağladığı görülmektedir. DMF türleri incelendiğinde ise DMF içeren tüm maddelerde TB DMF belirlendiği gözlenmiştir.

Alt Problem 3c'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-Türkiye uygulaması baz alınarak farklı dil farklı kültüre göre GMH ile yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

Fen öğrenmeden zevk alma (ST94) ve fen öğrenmede araçsal güdülenme (ST113) alt boyutlarında İngiltere ve Türkiye örneklemine ait GMH analiz sonuçları Tablo 24'te gösterilmiştir.

Tablo 24

İngiltere-Türkiye verisine ait GMH analizi sonuçları

Boyut	Madde	X^2_{GMH}	p*	L-A LOR**	Avantaj Sağladığı Grup
ST94	Q01	111,934	0,000	0,634	İngiltere
	Q02	434,785	0,000	-1,361	Türkiye
	Q03	458,524	0,000	1,304	İngiltere
	Q04	16,560	0,000	-0,123	Türkiye
	Q05	52,610	0,000	-0,429	Türkiye
ST113	Q01	7,676	0,053	-0,16	-
	Q02	394,494	0,000	1,086	İngiltere
	Q03	90,394	0,000	-0,535	Türkiye
	Q04	68,842	0,000	-0,259	Türkiye

*p ölçüt değeri olarak Bonferroni hesaplaması kullanılmıştır. **DIFAS programı ile hesaplanmıştır.

Tablo 24 incelendiğinde fen öğrenmede araçsal güdülenmesine ait ST113 1. maddesi dışındaki tüm maddelerin DMF gösterdiği söylenebilir. DIFAS sonucu elde edilen L-A LOR değerlerine bakıldığında fen öğrenmeden zevk alma boyutunda 1. ve 3. maddelerin araçsal güdülenme boyutunda yalnız 2. maddenin İngiltere grubuna avantaj sağladığı gözlenmektedir.

Alt Problem 3d'ye İlişkin Bulgular ve Yorumlar

PISA 2015 öğrenci anketlerinde yer alan tutum maddeleri İngiltere-Türkiye uygulaması baz alınarak farklı dil farklı kültüre göre OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizlerin sonuçları birbiriyle uyumlu mudur?

İngiltere ve Türkiye verisinde OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucu DMF gösteren maddelere Tablo 25'te yer verilmiştir.

Tablo 25

İngiltere-Türkiye verisinde OLR, poly-SIBTEST ve GMH yöntemlerinin karşılaştırılması

Boyut	Madde	OLR DMF varlığı	OLR DMF düzeyi	OLR DMF Türü	poly-SIBTEST DMF Varlığı	poly-SIBTEST DMF Türü	poly-SIBTEST DMF Düzeyi	Avantajlı Grup (poly-SIB)	GMH DMF varlığı	Avantajlı Grup (GMH)
ST94	Q01	var	A	TB	var	TB	C	İngiltere	var	İngiltere
	Q02	var	A	TB	var	TB	C	Türkiye	var	Türkiye
	Q03	var	B	TB	var	TB	C	İngiltere	var	İngiltere
	Q04	var	A	TBO	-	-	-	-	var	Türkiye
	Q05	var	A	TB	var	TB	A	Türkiye	var	Türkiye
ST113	Q01	var	A	TB	-	-	-	-	-	-
	Q02	var	A	TBO	var	TB	C	İngiltere	var	İngiltere
	Q03	var	A	TB	var	TB	C	Türkiye	var	Türkiye
	Q04	var	A	TB	var	TB	A	Türkiye	var	Türkiye

Tablo 25'e bakıldığında 9 maddenin tamamında DMF belirleyen sadece OLR yöntemidir. GMH yöntemiyle 8 ve poly-SIBTEST yöntemiyle ise 7 maddede (5'i yüksek düzeyde) DMF belirlenmiştir. Ayrıca poly-SIBTEST iki madde dışında yüksek düzeyde DMF belirlerken, çok hassas ölçümler yapması nedeniyle küçük değerler alan OLR yönteminde bir maddede (ST94 Q03) orta düzeyde DMF'li madde bulunmaktadır. Maddelerin GMH ve poly-SIBTEST yöntemlerine göre avantaj sağladığı gruplara bakıldığında ise fen öğrenmeden zevk alma 1.ve 3.; araçsal güdülenme 2. maddelerin İngiltere, fen öğrenmeden zevk alma 2. ve 5.; araçsal güdülenme 3. ve 4. maddelerin ise Türkiye lehine çalıştığı gözlenmektedir. Bu bakımdan DMF içeren maddelerin avantaj sağladığı gruplar arasında uyumluluk gözlenmektedir. DMF türlerine bakıldığında ise, poly-SIBTEST ve OLR yöntemlerinde ortak DMF içeren maddelerin TB DMF göstermesi bakımından tutarlılık görülmektedir.

Alt Problem 3e'ye İlişkin Bulgular ve Yorumlar

OLR, poly-SIBTEST ve GMH yöntemleriyle yapılan analizler sonucunda her üç yönteme göre DMF gösteren maddeler nelerdir?

Tablo 25'e bakıldığında her üç yönteme göre fen öğrenmeden zevk alma boyutunda 1., 2, 3. ve 5. maddeler ile araçsal güdülenme boyutunda 2., 3. ve 4. maddeler tüm yöntemler tarafından DMF tespit etmiştir. Bu ortak maddelerin 5 tanesi poly-SIBTEST tarafından önemli düzeyde (C) DMF belirlerken bir maddede ilk defa OLR yöntemi B düzeyinde DMF belirlemiştir. GMH ve poly-SIBTEST yöntemlerinin DMF belirlemediği 1 madde (ST 113 Q01) ortaktır.

Bölüm 5

Sonuç, Tartışma ve Öneriler

Araştırmanın bu bölümünde alt problem sırasına göre verilmiş araştırma sonuçlarına ve yapılacak araştırmalara yönelik önerilere yer verilmiştir.

Sonuçlar

Alt problem 1'e ilişkin sonuçlar. İngiltere-İrlanda verilerinde (aynı dil, benzer kültür) DMF analizleri sonucunda OLR yöntemi ve GMH yöntemi ile altı, poly-SIBTEST yöntemiyle yedi maddede DMF tespit edilmiştir. Yapılan analizde üç yönetime göre DMF içermeyen bir madde (araçsal güdülenme boyutunda 3. madde) varken, DMF gösteren beş madde ortaktır. Bunlar; fen öğrenmeden zevk alma boyutundaki 1.,3.,4. ve 5. maddeler ile araçsal güdülenme boyutundaki 1. maddededir. Ortak olan bu beş maddeden dördünde poly-SIBTEST yöntemine göre önemli düzeyde DMF belirlenmiştir (fen öğrenmeden zevk alma 1. ve 4. maddelerde C, 3 ve 5. maddelerde B düzeyde DMF). Maddelerin GMH ve poly-SIBTEST yöntemlerine göre avantaj sağladığı gruplara bakıldığında ise fen öğrenmeden zevk alma alt boyutunda 1. ve 3. maddelerin İngiltere; fen öğrenmeden zevk alma 2., 4. ve 5. ile araçsal güdülenme boyutunda 1. maddelerin ise İrlanda lehine çalıştığı gözlenmiştir. İngiltere lehine işleyen maddeler fene karşı ilgi duyma ve fen çalışırken mutlu olma iken fen hakkında yazıları okumaktan zevk alma, fen öğrenirken eğlenme ve fenin kariyerinde yardımcı olacağı düşüncesi İrlanda lehine çalışan maddelerdir. Bu sonuçlar bakımından iki yöntem arasında uyumluluk gözlenmiştir.

Alt problem 2'ye ilişkin sonuçlar. İngiltere-ABD verilerinde (aynı dil, farklı kültür) DMF analizleri sonucunda OLR ve GMH yöntemleri yedi, poly-SIBTEST yöntemiyle altı maddede DMF saptanmıştır. Analizde beş madde, üç yönetime göre ortak DMF göstermektedir. Bu maddeler fen öğrenmeden zevk alma boyutunda 3.,4. ve 5. maddeler ile araçsal güdülenme boyutunda 2. ve 3. maddelerdir. Bu ortak DMF'li maddelerden fen öğrenmeden zevk alma 3. ile araçsal güdülenme 2. ve 3. maddelerinde poly-SIBTEST tarafından önemli düzeyde DMF saptanmıştır. İngiltere-ABD veri analizinde her üç yönetime göre DMF içermeyen ortak maddelere rastlanmaz iken, GMH yönteminde fen öğrenmeden zevk alma boyutundaki 2. madde, OLR yönteminde ise araçsal güdülenme boyutundaki 4. madde diğer yöntemlerden farklı olarak DMF gözlenmiştir. DMF içeren maddelerin GMH ve poly-SIBTEST yöntemlerine

göre avantaj sağladığı gruplara bakıldığında ise fen öğrenmeden zevk alma alt boyutunda 3. madde ile araçsal güdülenme alt boyutundaki 2. maddenin İngiltere, fen öğrenmeden zevk alma 4. ve 5. maddeleri ile araçsal güdülenme 3. maddelerin ise ABD lehine çalıştığı gözlenmiştir. İngiltere lehine çalışan madde fen bilimlerinde yeni bilgiler edinmeden zevk alma iken; fen ile ilgili konulardan zevk alma, fen öğrenirken eğlenme ve kariyerine ilişkin önemini kavrama ABD lehine çalışan maddelerdir. Bu sonuçlar bakımından iki yöntem arasında uyumluluk gözlenmiştir.

Alt problem 3'e ilişkin sonuçlar. İngiltere- Türkiye verilerinde yapılan analize göre OLR yöntemi ile tüm maddelerde DMF belirlenirken, GMH yöntemiyle sekiz ve poly-SIBTEST yöntemiyle yedi maddede (beşi yüksek düzeyde) DMF belirlenmiştir. Her üç yönetime göre fen öğrenmeden zevk alma boyutunda 1., 2, 3. ve 5. maddeler ile araçsal güdülenme boyutunda 2., 3. ve 4. maddeler (toplam yedi madde) tüm yöntemler tarafından DMF tespit etmiştir. Bu ortak maddelerin beş tanesi poly-SIBTEST tarafından önemli düzeyde (C) DMF belirlenirken, çok hassas ölçümler yapması nedeniyle küçük değerler alan OLR yönteminde bir maddede (ST94 Q03) orta düzeyde (B) DMF'li madde belirlenmiştir. Farklı dil farklı kültüre göre yapılan bu analizde GMH ve poly-SIBTEST yöntemlerinin DMF belirlemediği bir madde (ST 113 Q01) ortaktır. Maddelerin GMH ve poly-SIBTEST yöntemlerine göre avantaj sağladığı gruplara bakıldığında ise fen öğrenmeden zevk alma 1.ve 3.; araçsal güdülenme 2. maddelerin İngiltere, fen öğrenmeden zevk alma 2. ve 5.; araçsal güdülenme 3. ve 4. maddelerin ise Türkiye lehine çalıştığı gözlenmektedir. İngiltere lehine çalışan maddeler fene karşı ilgi duyma, fen çalışırken mutlu olma, kariyerine ilişkin önemini kavrama gibi maddeler iken fen hakkında yeni bilgiler edinmekten zevk alma, öğrenirken eğlenme, kariyerinde katkı sağlayacağı ve bu amaçla çaba göstermeye değer bulma düşüncesi Türkiye lehine çalışan maddelerdir. Bu bakımdan GMH ve poly-SIBTEST yöntemlerine göre DMF içeren maddelerin avantaj sağladığı gruplar arasında uyumluluk gözlenmektedir.

Tartışma

Bu araştırmada PISA 2015 fene yönelik tutum maddelerinde dil ve kültür değişkenlerine göre DMF gösteren maddeler olup olmadığı incelenmiştir. Çok kategorili verilerde DMF analizleri için parametrik yöntemler arasından OLR, parametrik olmayan yöntemler arasından GMH ve poly-SIBTEST yöntemleri tercih

edilmiştir. Analiz sonuçları incelendiğinde DMF içeren maddelere ilişkin genel olarak yöntemler arasında tutarlılık görülmektedir. Bununla birlikte, DMF'li maddeler arasında en fazla uyum GMH ve OLR yöntemleri arasında gözlenmiştir.

Dil ve kültür değişkenleri bakımından seçilen İrlanda, İngiltere, ABD ve Türkiye veri setlerinde yapılan analizler sonucunda, İngiltere-İrlanda (aynı dil-benzer kültür) ve İngiltere-ABD (aynı dil-farklı kültür) karşılaştırmaları arasında her üç yöneme göre DMF gösteren maddelerde önemli bir farklılık görülmemiştir. İngiltere-Türkiye (farklı dil farklı kültür) karşılaştırması ise bu araştırmada en fazla DMF içeren grup olmuştur. Bu durum, Köse'nin (2015) belirttiği gibi maddelerin kültürel farklılıklardan ziyade dilsel farklılıklardan kaynaklı DMF gösterebileceği sonucuna varılabilir. Bu bulgu aynı zamanda, farklılaşmanın fazla olduğu grupların DMF'li madde sayısında da artışa neden olduğu benzer araştırmalarla da destekler niteliktedir (Asil ve Gelbal, 2012; Atalay, 2010; Huang, 2010; Köse, 2015; Mohorić ve Takšić, 2016). Diğer taraftan, ülke karşılaştırmalarının hangi grup lehine avantaj sağladığı incelendiğinde GMH ve poly-SIBTEST yöntemleri arasında tutarlılık gözlenmiştir.

DMF etki büyüklüğü incelendiğinde, çok kategorili verilerde DIFAS programı GMH'nin DMF türü ve etki büyüklüğüne ilişkin bilgi vermediğinden, bu yöntem için DMF düzeylerini karşılaştırma imkânı olmamıştır. OLR yönteminde ise hassas ölçüt aralıkları tercih edilmesine rağmen analiz sonuçlarının benzer çalışmalar gibi (Atar ve Kamata, 2011; Crane ve diğerleri, 2006; Choi, Gibbons ve Crane, 2011) çok küçük değerler almasından dolayı poly-SIBTEST yöntemi ile DMF düzeyleri arasında tutarlılık görülemediği. DMF düzeyleri arasındaki bu farklılığın nedeni olarak seçilen yöntemlerde DMF düzeyleri sınıflanmasının değer aralıklarındaki farklılıktan kaynaklanabileceği düşünülmektedir.

Öneriler

Bu çalışma sonucunda yapılacak öneriler iki başlık altında incelenebilir.

Araştırma sonucundan çıkan öneriler. Her üç yılda bir düzenli olarak uygulanan ve öğrenci başarısı, öğrenci durumları ve ülke karşılaştırmalarına olanak veren PISA uygulamasında DMF analizlerinin güncelliği adına çalışmalar düzenli olarak yapılmalıdır.

Elde edilen sonuçlar neticesinde, gruplar arasında farklılaşmaya neden olabilecek dil, kültür gibi değişkenlerden etkilenebilecek madde yapıları incelenerek olası DMF nedenleri belirlenmelidir.

Araştırmacılara yönelik öneriler. Bu çalışmada dilsel ve kültürel açıdan benzer ve farklı olacağı düşünülen ülkeler seçilmiştir. Yapılan araştırma sonucunda en fazla DMF içeren maddeler farklı dil ve farklı kültür karşılaştırmasında olmuştur. Bu bağlamda, PISA uygulamasının farklı dillerdeki çevirileri arasında da DMF çalışmaları yapılabilir ve bu maddelerin DMF nedenleri araştırılabilir.

Bu çalışma için kültür ve dil değişkenleri açısından farklı ve benzer olacağı düşünülen dört ülke tercih edilmiştir. Benzer dil kullanan üç ülke arasında yapılan analizlerde DMF'li madde sayısında farklılık gözlenmemiştir. Bu sebeple başka örneklemeler seçilerek yapılan DMF analizinde DMF içeren maddelerde dil ve kültür farkı olan ülkelere ait veriler seçilerek DMF açısından incelenebilir.

Bu araştırmada çok kategorili verilerde DMF analizleri için poly-SIBTEST, OLR ve GMH yöntemleri kullanılmıştır. OLR yöntemi DMF türü ve düzeyine ilişkin, poly-SIBTEST yöntemi DMF düzeyi ve hangi gruba avantaj sağladığına ilişkin ve GMH yöntemi ise hangi gruba avantaj sağladığına ilişkin bilgi vermiştir. Kullanılan üç yöntemde DMF varlığı dışında ortak karşılaştırma yapabilecek bir ölçüt bulunmamaktadır. Bu bakımdan, her yöntemin DMF belirlemedeki istatistiksel gücünün farklı olabileceği düşünüldüğünden çok kategorili verilerde bir simülasyon çalışması yapıldıktan sonra hangi yöntemin hangi şartlarda DMF belirlemede daha uygun olduğuna bakılabilir. Seçilen DMF belirleme yöntemleri ile elde edilen sonuçlar karşılaştırılabilir.

Çalışmada GMH yöntemi için kullanılan DIFAS programı iki kategorili verilerde DMF hakkında detaylı bilgi verirken, çok kategorili verilerde yalnız DMF yönü hakkında bilgi vermiştir. DIFAS programı analiz çıktıları incelendiğinde L-A LOR istatistiğinin mutlak değerleri yüksek değer alan verilerin poly-SIBTEST yöntemine göre C düzeyde ve düşük değer alan verilerin ise poly-SIBTEST yöntemine göre A düzeyde DMF içerdiği göze çarpmaktadır. Bu bakımdan, L-A LOR değerinin DMF düzeyi belirlemede kullanılıp kullanılmayacağına yönelik simülasyon ve uygulamalı araştırmalar yapılabilir. Bununla birlikte, çok kategorili verilerde DMF belirlenen maddeler için farklı kriter değerleri belirlenerek yöntemler arası uyum incelenebilir.

PISA 2015 uygulaması fene karşı tutum maddelerinin incelendiđi bu DMF alıřmasında, belli gruplara avantaj sađlayan maddelerin olası yanlılık nedenleri uzman grřleri ile arařtırılabilir.

Kaynaklar

- Abbott, M. L. (2007). A confirmatory approach to differential item functioning on an ESL reading assessment. *Language Testing*, 24(1), 7–36. Retrieved from <http://ltj.sagepub.com/content/24/1/7>
- Abedalaziz, N. (2010). Detecting gender related DIF using logistic regression and Mantel-Haenszel approaches. *Procedia - Social and Behavioral Sciences*, 7(2), 406–413. <https://doi.org/10.1016/j.sbspro.2010.10.055>
- Acock, A. C. (2005). Working with missing values. *Journal of Marriage and Family*, 65, 1012–1028.
- Agresti, A. (1984). Analysis of ordinal categorical data. *Biometrical Journal*, 29(1), 287.
- Akın Arıkan, Ç., Uğurlu, S., & Atar, B. (2016). MIMIC, SIBTEST, lojistik regresyon ve Mantel-Haenszel yöntemleriyle gerçekleştirilen DMF ve yanlılık çalışması. *Hacettepe Eğitim Dergisi*, 31(1), 34–52.
- Alavi, S. M., Rezaee, A. A., & Amirian, S. M. R. (2011). Academic discipline DIF in an English Language Proficiency Test. *Journal of English Language Teaching and Learning*, 5(7), 39–65.
- Ankenmann, R. D., Witt, E. A., & Dunbar, S. B. (1999). An investigation of the power of the likelihood ratio goodness of fit statistic in detecting differential item functioning. *Journal of Educational Measurement*, 36(4), 277–300.
- Arıkan, Ç. A. (2015). Değişen madde fonksiyonu belirlemede MTK-olabilirlik oranı, ordinal lojistik regresyon ve poly-sıbtest yöntemlerinin karşılaştırılmaları. *E-International Journal of Educational Research*, 6(1), 1–16. <https://doi.org/10.19160/e-ijer.24504>
- Asil, M., & 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 incelenmesi*. (Yüksek Lisans Tezi) Hacettepe Üniversitesi, Ankara.
- Atar, B., & Kamata, A. (2011). Comparison of IRT likelihood ratio test and logistic regression DIF detection procedures. *Hacettepe Üniversitesi Eğitim Bilimleri Dergisi*, 41, 36–47.
- Barnett, V., & Lewis, T. (1995). *Outliers in statical data* (3rd ed.). Chichester, England: Willey Series in Probabilty and Mathematical Statistics.

- Başokçu, T. O., & Öğretmen, T. (2013). Öğretmen öz-yeterlik ölçeğinde değişen madde fonksiyonlarının ağırlıklandırılmış cevap modeli ile belirlenmesi. *Ege Eğitim Dergisi*, 14(2), 63–78.
- Benito, G., & Navas-Ara, M. J. (2000). A comparison of Ki-kare, RFA and IRT based procedures in the detection of DIF. *Quality & Quantity*, 34, 17–31.
- Bolt, D. M. (2002). A Monte Carlo comparison of parametric and nonparametric polytomous DIF detection methods. *Applied Measurement in Education*, 15, 113–141.
- Butakor, P. (2015). A comparison of DIF detection and effect size measures among Mantel-Haenszel, SIBTEST and logistic regression using science test data. *African Journal of Science and Research*, 2(4), 12–15. Retrieved from <http://ajsr.rstpublishers.com/>
- Büyükköse, G. T. (2018). Farklı programlarda okutulan matematik dersi sınavlarına yönelik değişen madde fonksiyonu analizi. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi*, 4(1), 47–60.
- Büyükoztürk, Ş. (2002). Faktör Analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32(32), 470–483. Retrieved from <http://dergipark.gov.tr/kuey/issue/10365/126871>
- Büyükoztürk, Ş. (2013). *Sosyal bilimler için veri analizi el kitabı* (18. Baskı). Ankara: Pegem Akademi.
- Camilli, G., & Shepard, L. (1994). *Methods for identifying biased test items* (Vol. 4). California: Sage Publications.
- Chalmers, R. P. (2012). Package for the R environment. *Journal of Statistical Software*, 48(6), 1–29.
- Chang, H., Mazzeo, J., & Roussos, L. (1996). Detecting DIF for polytomously scored items: an adaptation of the SIBTEST procedure. *Journal of Educational Measurement*, 33, 333–353. <https://doi.org/10.1111/j.1745-3984.1996.tb00496.x>
- Childs, R. A. (1990). Gender bias and fairness. *Practical Assessment, Research & Evaluation*, 2(3), 1–5.
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif : An R package for detecting differential item functioning using iterative ordinal logistic regression/ item response theory and Monte Carlo simulations. *Journal of Statistical Software*, 39(8), 1–30.
- Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differential

- item functioning test items. *Issues and Practice*, 17(1), 31–44.
- Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and White Students in integrated colleges. *Journal of Educational Measurement*, 5(2), 115–124. <https://doi.org/10.1111/j.1745-3984.1968.tb00613.x>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Lawrence Erlbaum Associates.
- Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1), 15–18.
- Cox, D. R., & Snell, E. J. (1989). *Analysis of binary data* (2nd ed.). New York: Chapman and Hall.
- Crane, P. K., Gibbons, L., Ocepek-Weiklson, K., Cook, K., Cella, D., Narasimhalu, K., ... Teresi, J. A. (2007). A comparison of three sets of criteria for determining the presence of differential item functioning using ordinal logistic regression. *Quality of Life Research*, 16(1), 69–84.
- Crane, Paul K., Belle, G. Van, & Larson, E. B. (2004). Test bias in a cognitive test: Differential item functioning in the CASI. *Statistics in Medicine*, 23(2), 241–256. <https://doi.org/10.1002/sim.1713>
- Crane, Paul K., Gibbons, L. E., Jolley, L., & Belle, G. Van. (2006). Differential item functioning analysis with ordinal logistic regression techniques: DIF detect and difwithpar. *Medical Care*, 44, 115–123. <https://doi.org/10.1097/01.mlr.0000245183.28384.ed>
- Cromwell, S. (2002). *A primer on ways to explore Item bias*. Austin, Texas, USA: Educational Resources Information Center (ERIC).
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–298.
- Cuevas, M., & Cervantes, V. H. (2012). Differential item functioning detection with logistic regression. *Mathématiques et Sciences Humaines*, (199), 45–59. <https://doi.org/10.4000/msh.12274>
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (2. Baskı). Ankara: Pegem Akademi.
- Demir, E. (2013). Kayıp verilerin varlığında çoktan seçmeli testlerde madde ve test parametrelerinin kestirilmesi: SBS örneği. *Eğitim Bilimleri Araştırma Dergisi*, 3(2), 47–68.

- Demir, S. (2013). *PISA 2009 matematik okuryazarlığı alt testinde bulunan maddelerin Mantel-Haenszel, SIBTEST ve lojistik regresyon yöntemleri ile değişen madde fonksiyonunun incelenmesi*. (Yüksek lisans tezi), Abant İzzet Baysal Üniversitesi, Bolu.
- Demir, S., & Köse, İ. A. (2014). Mantel-Haenszel, SIBTEST ve lojistik regresyon yöntemleri ile değişen madde fonksiyonunu analizi. *Uluslararası İnsan Bilimleri Dergisi*, 11(1), 700–714. <https://doi.org/10.14687/ijhs.v11i1.2798>
- Demirus, K. B. (2015). *Ortak maddelerin değişen madde fonksiyonu gösterip göstermemesi durumunda test eşitlemeye etkisinin farklı yöntemlerle incelenmesi*. (Doktora tezi). Hacettepe Üniversitesi, Ankara.
- Dinçer, M. A., & Kolaşın, G. U. (2009). *Türkiye’de öğrenci başarısında eşitsizliğin belirleyicileri. Eğitim Reformu Girişimi*. (Türkiye’de eğitimde eşitliğin geliştirilmesi için verilere dayalı savunu projesi) Sabancı Üniversitesi, İstanbul.
- Dorans, N. J., & Schmitt, A. P. (1991). *Constructed response and differential item functioning: A pragmatic approach*. ETS Research Report Series. New Jersey: Educational Testing Service Princeton. <https://doi.org/10.1007/s00391-014-0816-5>
- Eells, K., Allison, D., Havighurst, R. J., Herrick, V. E., & Tyler, R. (1951). *Intelligence and cultural differences: A study of cultural learning and problem-solving*. Chicago: University of Chicago Press.
- Elashoff, J. D. (1972). A model for quadratic outliers in linear regression. *Journal of American Statistical Association*, 67(338), 478–485.
- Ellis, B. B., & Raju, N. S. (2003). Test and item bias : What they are, what they aren't, and how to detect them. Retrieved from <https://files.eric.ed.gov/fulltext/ED480042.pdf>
- Ercikan, K., Gierl, M. J., McCreith, T., Puhan, G., & Koh, K. (2004). Comparability of bilingual versions of assessments: Sources of incomparability of English and French versions of Canada’s National Achievement Tests. *Applied Measurement in Education*, 17(3), 301–321. https://doi.org/10.1207/s15324818ame1703_4
- Ercikan, K., & Koh, K. (2005). Examining the construct comparability of the English and French versions of TIMSS. *International Journal of Testing*, 5(1), 23–35. https://doi.org/10.1207/s15327574ijt0501_3
- Ersoy, Y. (2007). TIMSS-2007 : Uluslararası matematik fen araştırması-II: Başarıyı etkileyen örtük değişkenler ve genel eğilimler. <http://www.f2e2->

ogretmen.com/dagarcigimiz/f2e2-522.pdf adresinden erişilmiştir.

- Fidalgo, A. A., Pons, P. F., & Quintanilla, C. L. (2013). Análisis exploratorio del DIF y propiedades psicométricas del Test de Comprensión de las Emociones. In *Psicología-Reflexao eCritica*. Comunicación escrita presentada en el VI Congreso Internacional de Psicología Clínica, Santiago de Compostela.
- Fidalgo, Á. M. (2011). GMHDIF: A computer program for detecting dif in dichotomous and polytomous items using generalized Mantel-Haenszel statistics. *Applied Psychological Measurement*, 35(3), 247–249. <https://doi.org/10.1177/0146621610375691>
- Fidalgo, A. M., & Bartram, D. A. (2010). A comparison between some generalized Mantel-Haenszel statistics for detecting DIF in data simulated under the graded response model. *Applied Psychological Measurement*, 34(8), 600–606.
- Fidalgo, A. M., & Cobian, M. L. Q. (2018). Effect of the ability distribution shape on the generalized Mantel-Haenszel statistics used for detection. *Revista Brasileira De Biometria*, 36(2), 439–453.
- Fidalgo, A. M., & Scalón, J. D. (2012). Using Mantel-Haenszel methods for detecting differential item functioning. *Psicologia-Reflexao ECritica*, 25(1), 60–68.
- Field, A. (2000). Discovering statistics using SPSS for windows (pp. 1–32). New Delhi: Sage Publications.
- French, A. W., & Miller, T. R. (1996). Logistic regression and its use in detecting differential item functioning in polytomous items. *Journal of Educational Measurement*, 33(3), 315–332. <https://doi.org/10.1111/j.1745-3984.1996.tb00495.x>
- Gelin, M. N., & Zumbo, B. D. (2003). Differential item functioning results may change depending on how an item is scored: An illustration with the center for epidemiologic studies depression scale. *Educational and Psychological Measurement*, 63(1), 65–74. <https://doi.org/10.1177/0013164402239317>
- Gierl, M. J., & Jodoin, G. M. (2001). Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14, 329–349. https://doi.org/10.1207/S15324818AME1404_2
- Gierl, M. J., Khaliq, S. N., & Boughton, K. (1999). Gender differential item functioning in mathematics and science: Prevalence and policy implications. Canada: Paper Presented at the Symposium entitled “Improving Large-Scale Assessment in

- Education” at the Annual Meeting of the Canadian Society for the Study of Education.
- Gök, B., Atalay, K. K., & Kelecioğlu, H. (2014). PISA 2009 öğrenci anketi tutum maddelerinin kültüre göre değişen madde fonksiyonu açısından incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(1), 72–87.
- Gök, B., Kelecioğlu, H., & Doğan, N. (2010). Değişen madde fonksiyonunu belirlemede Mantel–Haenszel ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 35(156), 12–25.
- Groves, R. M. (2006). No nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 549–576.
- Hambleton, R. K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1(1), 1–13. <https://doi.org/10.1017/CBO9781107415324.004>
- Hambleton, R. K., & Rogers, H. J. (1989). Detecting potentially biased test items: Comparison of IRT area and Mantel-Haenszel methods. *Applied Measurement in Education*, 2(4), 313–334. https://doi.org/10.1207/s15324818ame0204_4
- Henderson, D. L. (2001). *Prevalence of gender DIF in mixed format High School Exit Examinations [microform]*. ERIC. Retrieved from <https://eric.ed.gov/?id=ED458284>
- Hidalgo, M. D., & López-Pina, J. A. (2004). Differential item functioning detection and effect size: A comparison between logistic regression and Mantel-Haenszel procedures. *Educational and Psychological Measurement*, 64(6), 903–915. <https://doi.org/10.1177/0013164403261769>
- Holland, P. W., & Thayer, D. T. (1986). *Program statistics research*. Princeton, NJ: Educational Testing Service. <https://doi.org/10.1002/j.2330-8516.1986.tb00186.x>
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Huang, X. (2010). *Different item functioning: The consequence of language, curriculum or culture?* (Doctoral dissertation). University of California, Berkeley.
- Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis*. Upper Saddle River, NJ.: Prentice Hall.
- Kalaycı, Ş. (2018). *SPSS uygulamalı çok değişkenli istatistik teknikler* (9.Baskı). Ankara: Asil Yayıncılık.

- Kamata, A., & Vaughn, B. K. (2004). An introduction to differential item functioning analysis. *Learning Disabilities: A Contemporary Journal*, 2(2), 49–69.
- Karami, H., & Nodoushan, M. A. S. (2011). Differential item functioning (DIF): Current problems and future directions. *International Journal of Language Studies*, 5(3), 133–142.
- Kaya, Y., Leite, W. L., & Miller, M. D. (2015). A comparison of logistic regression models for DIF detection in polytomous Items: The effect of small sample sizes and non-normality of ability distributions. *International Journal of Assessment Tools in Education*, 2(1), 22–39.
- Kelderman, H. (1989). Item bias detection using loglinear IRT. *Psychometrika*, 54(4), 681–697.
- Keleciođlu, H., Karabay, B., & Karabay, E. (2014). Seviye Belirleme Sınavı'nın madde yanlılıđı aısından incelenmesi. *İlköđretim Online*, 13(3), 934–953.
- Kim, S., Cohen, A. S., Alagoz, C., & Kim, S. (2007). DIF detection and effect size measures for polytomously scored items. *Journal of Education Measurement*, 44, 1–28. <https://doi.org/10.1111/j.1745-3984.2007.00029.x>
- Kılı, S. (2016). Cronbach'ın alfa gvenirlik katsayısı. *Journal of Mood Disorders (JMOOD)*, 6(1), 47–48. <https://doi.org/10.5455/jmood.20160307122823>
- Kline, R. B. (2011). Principles and practices of structural equation modelling. New York: The Guilford Press.
- Koak, D., okluk, ., & Kayri, M. (2016). Faktr sayısının belirlenmesinde MAP Testi, paralel analiz, K1 ve yama birikinti grafiđi yntemlerinin karřılařtırılması. *YY Eđitim Fakltesi Dergisi*, 8(1), 330–359.
- Kse, İ. A. (2015). PISA 2009 đrenci anketi alt leklerinde (Q32-Q33) bulunan maddelerin deđiřen madde fonksiyonu aısından incelenmesi. *Kastamonu Eđitim Dergisi*, 23(1), 227–240.
- Koyuncu, İ., Aksu, G., & Keleciođlu, H. (2018). Mantel-Haenszel, lojistik regresyon ve olabilirlik oranı deđiřen madde fonksiyonu inceleme yntemlerinin farklı yazılımlar kullanılarak karřılařtırılması. *İlköđretim Online*, 17(2), 909–925. <https://doi.org/10.17051/ilkonline.2018.419339>
- Kristjansson, E., Aylesworth, R., Mcdowell, I., & Zumbo, B. D. (2005). A comparison of four methods for detecting DIF in ordered response items. *Educational and Psychological Measurement*, 65, 935–953. <https://doi.org/10.1177/0013164405275668>

- Kurz, T. (1999). A review of scoring algorithms for multiple-choice tests (pp. 1–21). San Antonio, TX: Paper presented at the Annual Meeting of the Southwest Educational Research Association. <https://doi.org/ED419696>
- Lee, K. (2003). *Parametric and nonparametric IRT models for assessing differential item functioning*. (Unpublished Doctoral Dissertation). Wayne State University, USA.
- Li, H. H., & Stout, W. (1996). A new procedure for detection of crossing DIF. *Psychometrika*, *61*(4), 647–677.
- Lisova, T., & Kovalchuk, Y. (2016). Analysis of differential item functioning on some TIMSS 2011 data for Ukraine. In *Conference: ECER 2015, Education and Transition*. Nizhyn State Mykola Gogol University.
- Liu, I.-M., & Agresti, A. (1996). Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics*, *52*, 1223–1234.
- Mahalanobis, P. C. (1937). Normalisation of statistical variates and the use of rectangular co-ordinates in the theory of sampling distributions. *The Indian Journal of Statistics*, *3*(1), 1–40.
- Mantel, N. (1963). Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure. *Journal of American Statistical Association*, *58*(303), 690–700.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, *22*, 719–748.
- Maydeu-Olivares, A., Dragow, F., & Mead, A. D. (1996). Distinguishing among parametric item response models for polytomous ordered data. *Applied Psychological Measurement*, *18*(3), 245–256.
- Mellenbergh, G. J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, *7*(2), 105–118. <https://doi.org/10.3102/10769986007002105>
- Mellor, T. L. (1995). *A comparison of four differential item functioning methods for polytomous scored items*. (Unpublished doctor dissertation), The University of Texas, Austin.
- Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician*, *54*, 17–24.
- Miller, T. R., & Spray, J. (1994). Identifying nonuniform DIF in polytomously scored test

- items. *ACT Research Report Series*, 94(1), 1–16.
- Mohorić, T., & Takšić, V. (2016). DIF analiza upitnika emocionalne kompetentnosti Mantel-Haenszel metodom: kros-kulturalna usporedba (DIF analysis of ESCQ using Mantel-Haenszel procedure: Cross-cultural comparison). In *XX. Dani psihologije u Zadru*.
- Nagelkerke, N. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78, 691–692.
- OECD. (2016). *PISA 2015 Technical Report. Notes and Queries* (Vol. 146). <https://doi.org/10.1093/nq/CXLVI.mar01.162>
- Osborne, J. W. (2013). *Best practice in data cleaning: A complete guide to everything you need to do before and after collecting your data*. California: Sage Publications.
- Osterlind, S. J. (1983). *Test item bias*. Beverly Hills: Sage Publications. Retrieved from <file://catalog.hathitrust.org/Record/000476089>
- Osterlind, S. J., & Everson, H. (2009). *Differential item functioning* (2nd ed.). California: Sage Publications.
- Özberk, E. B. Ü., & Koç, N. (2017). WÇZÖ-IV Maddelerinin cinsiyet ve sosyo-ekonomik düzey açısından işlev farklılığının belirlenmesinde kullanılan yöntemlerin karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(1), 112–127. <https://doi.org/10.21031/epod.287577>
- Özçelik, D. A. (2010). *Ölçme ve değerlendirme*. Ankara: Pegem Yayıncılık.
- Özdamar, D. (2002). *Paket programlarda istatistiksel veri analizi-1* (4.Baskı). Eskişehir: Kaan Kitabevi.
- Pasca, L., Coello, M. T., Aragonés, J. I., & Frantz, C. M. P. (2018). The equivalence of measures on the connectedness to nature scale: A comparison between ordinal methods of DIF detection. *PloS One*, 13(11), e0207739. <https://doi.org/10.1371/journal.pone.0207739>
- Penfield, R. (2001). Assessing differential item functioning among multiple groups: A comparison of three Mantel-Haenszel procedures. *Applied Measurement in Education*, 14, 235–259.
- Penfield, R. (2013). DIFAS 5.0 Differential item functioning analysis system: User's manual. (Computer program).
- Penfield, R., & Algina, J. (2003). Applying the Liu-Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *Journal of Educational Measurement*, 40, 353–370.

- PISA 2015 Ulusal Raporu. (2016).
http://odsgm.meb.gov.tr/test/analizler/docs/PISA/PISA2015_Ulusal_Rapor.pdf
adresinden erişilmiştir.
- Polat, S. (2009). Türkiye’de eğitim politikalarının fırsat eşitsizliği üzerindeki etkileri. *DPT Uzmanlık Tezi*, 39(5), 561–563.
- Potenza, M. T., & Dorans, N. J. (1995). A framework for classification and evaluation. *Applied Psychological Measurement*, 19(1), 23–37.
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2009). *Measurement and assessment in teaching. Measurement and Assessment in Education*. Boston: Pearson.
<https://doi.org/10.1007/s13398-014-0173-7.2>
- Salı, G., & Sırgancı, G. (2018). Davranış ve duyguları değerlendirme ölçeği-2 (BERS-2)’nin kategoriler arasındaki psikometrik uzaklığının incelenmesi. *Kastamonu Eğitim Dergisi*, 26(4), 1217–1226.
- Sandilands, D. A. (2008). *Score scale comparability in international educational assessments*. (Master’s thesis), Universty British Columbia.
https://doi.org/10.4103/IJPC.IJPC_142_17
- Shealy, R., & Stout, W. (1991). *An item response theory model for test bias*. Hillsdale, NJ: Erlbaum.
- Stern, W. (1914). *The psychological methods of testing intelligence*. Baltimore: Warwick & York. Retrieved from file://catalog.hathitrust.org/Record/001280965
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361–370. <https://doi.org/10.1111/j.1745-3984.1990.tb00754.x>
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlik* (1st ed.). Ankara: Seçkin Yayıncılık.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (Fourth). Boston: Ally and Bacon.
- Tavşancıl, E. (2010). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel Yayın Dağıtım.
- Taylor, C., & Lee, Y. (2012). Gender DIF in reading and mathematics tests with mixed item formats. *Applied Measurement in Education*, 25(3), 246–280.
<https://doi.org/10.1080/08957347.2012.687650>
- Terzi, R., & Yakar, L. (2018). Seviye Belirleme Sınavında değişen madde ve değişen çeldirici fonksiyonu analizleri. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme*

- Dergisi*, 9(2), 137–150. <https://doi.org/10.21031/epod.368081>
- Tunç, E. B. (2017). *İki ve çok kategorili puanlanan maddelerde değişen madde fonksiyonlarının karşılaştırılması*. (Doktora tezi), Ankara Üniversitesi, Ankara.
- Turgut, M. F., & Baykul, Y. (2011). *Eğitimde ölçme ve değerlendirme* (3rd ed.). Ankara: Pegem Akademi.
- UNICEF. (2003). Eğitim toplumsal cinsiyet açısından incelenmesi, Türkiye 2003. https://www.unicef.org/turkey/pdf/_ge21.pdf adresinden erişilmiştir.
- Uyar, Ş., & Uyanık, G. K. (2016). PISA 2012 bilişsel maddelerin kültüre göre değişen madde fonksiyonu bakımından incelenmesi. *Eğitim Öğretim Araştırmaları Dergisi*, 5(3), 230–240.
- Uzun, N. B., & Gelbal, S. (2017). PISA fen başarı testinin madde yanlılığının kültür ve dil açısından incelenmesi. *Kastamonu Eğitim Dergisi*, 25(6), 2427–2446.
- Velicer, W. F., Eaton, C., & Fava, J. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. *Problems and Solutions in Human Assessment* (pp. 41–71). Boston, MA: Kluwer Academic Publishers.
- Wang, W. C., & Su, Y. H. (2004). Factors influencing the Mantel and generalized Mantel-Haenszel methods for the assessment of differential item functioning in polytomous items. *Applied Psychological Measurement*, 28(6), 450–480. <https://doi.org/10.1177/0146621604269792>
- Welch, C., & Hoover, H. D. (1993). Procedures for extending item bias techniques to polytomously scored items. *Applied Measurement in Education*, 6, 1–19.
- Wisnowski, J. W. (2001). Robust diagnostic regression analysis. *Journal of the American Statistical Association*, 33(3), 383–384. <https://doi.org/10.1198/jasa.2002.s465>
- Wongwiwatthananut, S., Popovich, N. G., & Bennett, D. E. (2000). Assessing pharmacy student knowledge on multiple-choice examinations using partial-credit scoring of combined-response multiple-choice items. *American Journal of Pharmaceutical Education*, 64(1), 1–10. <https://doi.org/10.1088/0266-5611/25/11/115016>
- Yıldırım, H. H., & Yıldırım, S. (2011). Correlates of communalities as matching variables in differential item functioning analyses. *Hacettepe Üniversitesi Eğitim Bilimleri Dergisi*, 40, 386–396.
- Yıldırım, K. (2012). PISA 2006 verilerine göre Türkiye’de eğitimin kalitesini belirleyen

- temel faktörler. *Türk Eğitim Bilimleri Dergisi*, 10(2), 229–255. Retrieved from <http://dergipark.gov.tr/download/article-file/256484>
- Yıldırım, S. (2008). Farklı işleyen maddelerin belirlenmesinde sınırlandırılmış faktör çözümlemesinin olabilirlik-oranı ve Mantel-Haenszel yöntemleriyle karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34, 297–307.
- Zumbo, B. D. (1999). Handbook on the theory and methods of differential item functioning: Logistic regression modeling as a unitary framework for binary and likert-type item scores. *Interpretation A Journal Of Bible And Theology* (pp. 1–57). Ottawa: Directorate of Human Resources Research and Evaluation.
- Zumbo, B. D. (2007). Three generations of DIF analyses: considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4(2), 223–233. <https://doi.org/10.1080/15434300701375832>
- Zumbo, B. D., & Thomas, D. R. (1996). *A measure of DIF effect size using logistic regression procedures*. National Board of Medical Examiners. Philadelphia, PA.
- Zwick, R. (1990). When do item response function and Mantel-Haenszel definition of differential item functioning coincide? *Journal of Educational Statistics*, 15(3), 185–197.
- Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance test. In *Presented at the annual meeting of the National Council of Measurement in Education* (pp. 4–18). San Francisco.

EK-A: PISA 2015 Öğrenci Anketinden Seçilen Tutum Maddeleri

ST094 (Enjoyment of Science)

How much do you disagree or agree with the statements about yourself below?

(Please select one response in each row.)

	Strongly disagree	Disagree	Agree	Strongly agree
I generally have fun when I am learning science topics.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I like reading about science.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am happy working on science topics.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I enjoy acquiring new knowledge in science.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am interested in learning about science.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

ST 113 (Enjoy of Learning Science)

How much do you agree with the statements below?

(Please select one response in each row.)

	Strongly disagree	Disagree	Agree	Strongly agree
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.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
What I learn in my science classes is important for me because I need this for what I want to do later on.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Studying science is worthwhile for me because what I learn will improve my career prospects.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Many things I learn in my science subject(s) will help me to get a job.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

EK-B: OLR ve poly SIBTEST Yöntemleri İçin R Programında Yazılan Kodlar

R programında OLR yöntemi analizine ilişkin örnek komutlar:

```
ENG_TURK$top_zevk <- rowSums(ENG_TURK[,3:7])
ENG_TURK$top_mot <- rowSums(ENG_TURK[,8:11])
ENG_USA_lordif_zevkalma <-lordif(ENG_USA[,3:7], ENG_USA[,1], criterion = c("Chisqr",
"R2", "Beta"), pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"), alpha = 0.01/5,
beta.change = 0.1, R2.change = 0.02, maxIter = 10, minCell = 5, control = list(), model =
"GRM", anchor = NULL, MonteCarlo = FALSE, nr = 100, weights = NULL, normwt = TRUE)
ENG_TUR_lordif_zevkalma
ENG_TUR_lordif_zevkalma$stats
plot(ENG_TUR_lordif_zevkalma,labels=c("eng","tur"))
write_sav( ENG_TUR_lordif_zevkalma$stats,"ENG_TUR_zevkalma_olr.sav")
```

R programında poly SIBTEST yöntemi analizine ilişkin örnek komutlar:

```
SIBTEST(ENG_TURK[,3:7],gr_eng_turk,1,focal=1)
SIBTEST(ENG_TURK[,3:7],gr_eng_turk,2,focal=1)
SIBTEST(ENG_TURK[,3:7],gr_eng_turk,3,focal=1)
SIBTEST(ENG_TURK[,3:7],gr_eng_turk,4,focal=1)
SIBTEST(ENG_TURK[,3:7],gr_eng_turk,5,focal=1),"ENG_TUR_zevkalma_sibtest.sav"
```

EK-C: OLR Yöntemine Göre Yapılan Analiz Çıktıları

OLR Yöntemine Göre İngiltere-İrlanda Çıktıları

```
> ENG_IRL_lordif_zevkalma <-lordif(ENG_IRL[,3:7], ENG_IRL[,1], criterion = c("Chisqr", "R2", "Beta"),
+                               pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"), alpha = 0.01, beta.change
= 0.1,
+                               R2.change = 0.02, maxlter = 10, minCell = 5,
+                               control = list(), model = "GRM", anchor = NULL, MonteCarlo = FALSE, nr = 100,
+                               weights = NULL, normwt = TRUE)
```

```
Iteration: 133, Log-Lik: -33231.949, Max-Change: 0.00010
```

```
(mirt) | Iteration: 1, 4 items flagged for DIF (1,3,4,5)
```

```
Iteration: 185, Log-Lik: -33040.084, Max-Change: 0.00010
```

```
(mirt) | Iteration: 2, 4 items flagged for DIF (1,3,4,5)
```

```
> ENG_IRL$top_zevk <- rowSums(ENG_IRL[,3:7])
```

```
> ENG_IRL$top_mot <- rowSums(ENG_IRL[,8:11])
```

```
> ENG_IRL_lordif_zevkalma
```

Call:

```
lordif(resp.data = ENG_IRL[, 3:7], group = ENG_IRL[, 1], criterion = c("Chisqr",
  "R2", "Beta"), pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"),
  alpha = 0.01, beta.change = 0.1, R2.change = 0.02, maxlter = 10,
  minCell = 5, control = list(), model = "GRM", anchor = NULL,
  MonteCarlo = FALSE, nr = 100, weights = NULL, normwt = TRUE)
```

Number of DIF groups: 2

Number of items flagged for DIF: 4 of 5

Items flagged: 1, 3, 4, 5

Number of iterations for purification: 2 of 10

Detection criterion: Chisqr

Threshold: alpha = 0.01

	item	ncat	chi12	chi13	chi23
1	1	4	0.0000	0.00	0.8142
2	2	4	0.2079	0.04	0.0276
3	3	4	0.0000	0.00	0.0125
4	4	4	0.0000	0.00	0.0752
5	5	4	0.0000	0.00	0.4381

```
> ENG_IRL_lordif_zevkalma$stats
```

	item	ncat	chi12	chi13	chi23	beta12	pseudo12.McFadden
1	1	4	0.0000	0.00	0.8142	0.0194	0.0067
2	2	4	0.2079	0.04	0.0276	0.0003	0.0001
3	3	4	0.0000	0.00	0.0125	0.0224	0.0041
4	4	4	0.0000	0.00	0.0752	0.0272	0.0056

```

5 5 4 0.0000 0.00 0.4381 0.0112 0.0029
pseudo13.McFadden pseudo23.McFadden pseudo12.Nagelkerke
1 0.0067 0e+00 0.0040
2 0.0003 2e-04 0.0000
3 0.0044 3e-04 0.0017
4 0.0057 1e-04 0.0021
5 0.0029 0e+00 0.0012

```

```

pseudo13.Nagelkerke pseudo23.Nagelkerke pseudo12.CoxSnell
1 0.0040 0e+00 0.0037
2 0.0002 1e-04 0.0000
3 0.0018 1e-04 0.0015
4 0.0022 1e-04 0.0019
5 0.0012 0e+00 0.0010

```

```

pseudo13.CoxSnell pseudo23.CoxSnell df12 df13 df23
1 0.0037 0e+00 1 2 1
2 0.0001 1e-04 1 2 1
3 0.0016 1e-04 1 2 1
4 0.0019 0e+00 1 2 1
5 0.0011 0e+00 1 2 1

```

```
> ENG_IRL_lordif_mot
```

Call:

```

lordif(resp.data = ENG_IRL[, 8:11], group = ENG_IRL[, 1], criterion = c("Chisqr",
  "R2", "Beta"), pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"),
  alpha = 0.01, beta.change = 0.1, R2.change = 0.02, maxIter = 10,
  minCell = 5, control = list(), model = "GRM", anchor = NULL,
  MonteCarlo = FALSE, nr = 100, weights = NULL, normwt = TRUE)

```

Number of DIF groups: 2

Number of items flagged for DIF: 1 of 4

Items flagged: 1

Number of iterations for purification: 2 of 10

Detection criterion: Chisqr

Threshold: alpha = 0.01

```
item ncat chi12 chi13 chi23
```

```

1 1 4 0.0000 0.0000 0.8904
2 2 4 0.5081 0.1378 0.0604
3 3 4 0.0399 0.1176 0.8074
4 4 4 0.4321 0.2258 0.1246
> ENG_IRL_lordif_mot$stats
  item ncat  chi12  chi13  chi23 beta12 pseudo12.McFadden
1  1  4 0.0000 0.0000 0.8904 0.0029      9e-04
2  2  4 0.5081 0.1378 0.0604 0.0002      0e+00
3  3  4 0.0399 0.1176 0.8074 0.0007      2e-04
4  4  4 0.4321 0.2258 0.1246 0.0000      0e+00
  pseudo13.McFadden pseudo23.McFadden pseudo12.Nagelkerke
1      9e-04      0e+00      5e-04
2      2e-04      1e-04      0e+00
3      2e-04      0e+00      1e-04
4      1e-04      1e-04      0e+00
  pseudo13.Nagelkerke pseudo23.Nagelkerke pseudo12.CoxSnell
1      5e-04      0e+00      4e-04
2      1e-04      1e-04      0e+00
3      1e-04      0e+00      1e-04
4      1e-04      1e-04      0e+00
  pseudo13.CoxSnell pseudo23.CoxSnell df12 df13 df23
1      4e-04      0e+00  1  2  1
2      1e-04      0e+00  1  2  1
3      1e-04      0e+00  1  2  1
4      1e-04      1e-04  1  2  1

```

OLR Yöntemine Göre İngiltere-ABD Çıktıları

```

> View(ENG_USA)
> ENG_USA_lordif_zevkalma <-lordif(ENG_USA[,3:7], ENG_USA[,1], criterion = c("Chisqr", "R2",
"Beta"),
+           pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"), alpha = 0.01, beta.change
= 0.1,
+           R2.change = 0.02, maxIter = 10, minCell = 5,
+           control = list(), model = "GRM", anchor = NULL, MonteCarlo = FALSE, nr = 100,
+           weights = NULL, normwt = TRUE)
Iteration: 126, Log-Lik: -32138.900, Max-Change: 0.00010
(mirt) | Iteration: 1, 3 items flagged for DIF (3,4,5)
Iteration: 197, Log-Lik: -32032.963, Max-Change: 0.00010
(mirt) | Iteration: 2, 3 items flagged for DIF (3,4,5)
> ENG_USA_lordif_zevkalma

```

Call:

```
lordif(resp.data = ENG_USA[, 3:7], group = ENG_USA[, 1], criterion = c("Chisqr",  
  "R2", "Beta"), pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"),  
  alpha = 0.01, beta.change = 0.1, R2.change = 0.02, maxlter = 10,  
  minCell = 5, control = list(), model = "GRM", anchor = NULL,  
  MonteCarlo = FALSE, nr = 100, weights = NULL, normwt = TRUE)
```

Number of DIF groups: 2

Number of items flagged for DIF: 3 of 5

Items flagged: 3, 4, 5

Number of iterations for purification: 2 of 10

Detection criterion: Chisqr

Threshold: alpha = 0.01

	item	ncat	chi12	chi13	chi23
1	1	4	0.1962	0.0204	0.0134
2	2	4	0.9377	0.1958	0.0712
3	3	4	0.0000	0.0000	0.0000
4	4	4	0.0003	0.0002	0.0517
5	5	4	0.0047	0.0174	0.7571

> ENG_USA_lordif_zevkalma\$stats

	item	ncat	chi12	chi13	chi23	beta12	pseudo12.McFadden	pseudo13.McFadden	pseudo23.McFadden	pseudo12.Nagelkerke	pseudo13.Nagelkerke	pseudo23.Nagelkerke	pseudo12.CoxSnell	pseudo13.CoxSnell	pseudo23.CoxSnell	df12	df13	df23	
1	1	4	0.1962	0.0204	0.0134	0.0002	0.0001	0.0004	0.0003	0.0000	0.0002	2e-04	0.0000	0.0001	1e-04	0.0000			
2	2	4	0.9377	0.1958	0.0712	0.0000	0.0000	0.0106	0.0013	0.0038	0.0001	0.0001	0.0000	0.0001	1e-04	0.0000			
3	3	4	0.0000	0.0000	0.0000	0.0549	0.0093	0.0008	0.0002	0.0002	0.0106	0.0013	0.0038	0.0044	5e-04	0.0034			
4	4	4	0.0003	0.0002	0.0517	0.0017	0.0006	0.0004	0.0000	0.0001	0.0004	0.0000	0.0001	0.0003	1e-04	0.0002			
5	5	4	0.0047	0.0174	0.7571	0.0007	0.0004	0.0004	0.0000	0.0001	0.0004	0.0000	0.0001	0.0001	0e+00	0.0001			

```

1      0.0002      1e-04  1  2  1
2      0.0001      1e-04  1  2  1
3      0.0039      5e-04  1  2  1
4      0.0003      1e-04  1  2  1
5      0.0001      0e+00  1  2  1

```

> ENG_USA_lordif_motivasyon

Call:

```

lordif(resp.data = ENG_USA[, 8:11], group = ENG_USA[, 1], criterion = c("Chisqr",
  "R2", "Beta"), pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"),
  alpha = 0.01/4, beta.change = 0.1, R2.change = 0.02, maxIter = 10,
  minCell = 5, control = list(), model = "GRM", anchor = NULL,
  MonteCarlo = FALSE, nr = 100, weights = NULL, normwt = TRUE)

```

Number of DIF groups: 2

Number of items flagged for DIF: 4 of 4

Items flagged: 1, 2, 3, 4

Number of iterations for purification: 2 of 10

Detection criterion: Chisqr

Threshold: alpha = 0.0025

```

item ncat  chi12 chi13 chi23
1  1  4 0.5301  0 0e+00
2  2  4 0.0000  0 0e+00
3  3  4 0.0000  0 7e-04
4  4  4 0.3123  0 0e+00

```

> ENG_USA_lordif_motivasyon\$stats

```

item ncat  chi12 chi13 chi23 beta12 pseudo12.McFadden
1  1  4 0.5301  0 0e+00 0.0002      0.0000
2  2  4 0.0000  0 0e+00 0.0188      0.0049
3  3  4 0.0000  0 7e-04 0.0078      0.0025
4  4  4 0.3123  0 0e+00 0.0001      0.0000

pseudo13.McFadden pseudo23.McFadden pseudo12.Nagelkerke
1      0.0010      0.0010      0.0000
2      0.0076      0.0027      0.0018
3      0.0030      0.0005      0.0012

```



```
4      0.0009      0.0008      0.0000
pseudo13.Nagelkerke pseudo23.Nagelkerke pseudo12.CoxSnell
```

```
1      0.0006      6e-04      0.0000
2      0.0027      1e-03      0.0016
3      0.0015      3e-04      0.0011
4      0.0006      6e-04      0.0000
```

```
pseudo13.CoxSnell pseudo23.CoxSnell df12 df13 df23
```

```
1      0.0005      5e-04  1  2  1
2      0.0025      9e-04  1  2  1
3      0.0013      2e-04  1  2  1
4      0.0005      5e-04  1  2  1
```

```
> plot(ENG_USA_lordif_motivasyon,labels=c("eng","usa"))
```

```
Error in plot.lordif(ENG_USA_lordif_motivasyon, labels = c("eng", "usa")) :
```

```
all items in ENG_USA_lordif_motivasyon have been flagged for DIF
```

OLR Yöntemine Göre İngiltere-Türkiye Çıktıları

```
> ENG_TURK_lordif_zevkalma <-lordif(ENG_TURK[,3:7], ENG_TURK[,1], criterion = c("Chisqr", "R2", "Beta"),
```

```
+ pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"), alpha = 0.01, beta.change = 0.1,
```

```
+ R2.change = 0.02, maxIter = 10, minCell = 5,
```

```
+ control = list(), model = "GRM", anchor = NULL, MonteCarlo = FALSE, nr = 100,
```

```
+ weights = NULL, normwt = TRUE)
```

```
Iteration: 111, Log-Lik: -34199.039, Max-Change: 0.00010
```

```
(mirt) | Iteration: 1, 4 items flagged for DIF (1,2,3,5)
```

```
Iteration: 121, Log-Lik: -33678.987, Max-Change: 0.00009
```

```
(mirt) | Iteration: 2, 5 items flagged for DIF (1,2,3,4,5)
```

```
Warning message:
```

```
In lordif(ENG_TURK[, 3:7], ENG_TURK[, 1], criterion = c("Chisqr", :
```

```
all items got flagged for DIF - stopping
```

```
> ENG_TUR_lordif_zevkalma$stats
```

```
item ncat chi12 chi13 chi23 beta12 pseudo12.McFadden
```

```
1 1 4 0.0000 0 0.0094 0.0076 0.0037
```

```
2 2 4 0.0000 0 0.0307 0.0515 0.0160
```

```
3 3 4 0.0000 0 0.0000 0.0673 0.0184
```

```
4 4 4 0.8826 0 0.0000 0.0000 0.0000
```

```
5 5 4 0.0000 0 0.0085 0.0054 0.0008
```

```
pseudo13.McFadden pseudo23.McFadden pseudo12.Nagelkerke
```

```
1 0.0040 0.0003 0.0022
```

2	0.0162	0.0002	0.0086
3	0.0209	0.0025	0.0078
4	0.0009	0.0009	0.0000
5	0.0011	0.0003	0.0003

pseudo13.Nagelkerke pseudo23.Nagelkerke pseudo12.CoxSnell

1	0.0024	2e-04	0.0020
2	0.0087	1e-04	0.0079
3	0.0089	1e-03	0.0072
4	0.0003	3e-04	0.0000
5	0.0005	1e-04	0.0003

pseudo13.CoxSnell pseudo23.CoxSnell df12 df13 df23

1	0.0022	2e-04	1	2	1
2	0.0080	1e-04	1	2	1
3	0.0081	9e-04	1	2	1
4	0.0003	3e-04	1	2	1
5	0.0004	1e-04	1	2	1

> ENG_TUR_ lordif_motivasyon

Call:

```
lordif(resp.data = ENG_TURK[, 8:11], group = ENG_TURK[, 1], criterion = c("Chisqr",
  "R2", "Beta"), pseudo.R2 = c("McFadden", "Nagelkerke", "CoxSnell"),
  alpha = 0.01/4, beta.change = 0.1, R2.change = 0.02, maxlter = 10,
  minCell = 5, control = list(), model = "GRM", anchor = NULL,
  MonteCarlo = FALSE, nr = 100, weights = NULL, normwt = TRUE)
```

Number of DIF groups: 2

Number of items flagged for DIF: 4 of 4

Items flagged: 1, 2, 3, 4

Number of iterations for purification: 1 of 10

Detection criterion: Chisqr

Threshold: alpha = 0.0025

item	ncat	chi12	chi13	chi23
1	1	4	0	0 0.9786
2	2	4	0	0 0.0000
3	3	4	0	0 0.0000

```

4 4 4 0 0 0.0000
> ENG_TUR_lordif_motivasyon$stats
  item ncat chi12 chi13 chi23 beta12 pseudo12.McFadden
1  1  4  0  0 0.9786 0.0053      0.0014
2  2  4  0  0 0.0000 0.0333      0.0114
3  3  4  0  0 0.0000 0.0213      0.0048
4  4  4  0  0 0.0000 0.0032      0.0007
  pseudo13.McFadden pseudo23.McFadden pseudo12.Nagelkerke
1      0.0014      0.0000      0.0008
2      0.0153      0.0039      0.0046
3      0.0061      0.0014      0.0026
4      0.0020      0.0013      0.0005
  pseudo13.Nagelkerke pseudo23.Nagelkerke pseudo12.CoxSnell
1      0.0008      0.0000      0.0007
2      0.0061      0.0015      0.0042
3      0.0034      0.0008      0.0024
4      0.0015      0.0010      0.0005
  pseudo13.CoxSnell pseudo23.CoxSnell df12 df13 df23
1      0.0007      0.0000  1  2  1
2      0.0056      0.0014  1  2  1
3      0.0031      0.0007  1  2  1
4      0.0014      0.0009  1  2  1

```

EK-Ç: Poly SIBTEST Yöntemine Göre Yapılan Analiz Çıktıları

Poly SIBTEST Yöntemine Göre İngiltere-İrlanda Analiz Çıktıları

> SIBTEST(ENG_IRL[,3:7],gr_eng_irl,1,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	0.129	0.011	150.559	1	0
CSIBTEST	1	4	1	0.129	NA	150.559	1	0

> SIBTEST(ENG_IRL[,3:7],gr_eng_irl,2,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	-0.028	0.011	5.903	1	0.015
CSIBTEST	1	4	1	0.028	NA	5.903	1	0.015

> SIBTEST(ENG_IRL[,3:7],gr_eng_irl,3,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	0.062	0.008	56.232	1	0
CSIBTEST	1	4	1	0.062	NA	56.232	1	0

> SIBTEST(ENG_IRL[,3:7],gr_eng_irl,4,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	-0.093	0.008	123.015	1	0
CSIBTEST	1	4	1	0.093	NA	123.015	1	0

> SIBTEST(ENG_IRL[,3:7],gr_eng_irl,5,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	-0.075	0.009	67.87	1	0
CSIBTEST	1	4	1	0.075	NA	67.87	1	0

> SIBTEST(ENG_IRL[,8:11],gr_eng_irl,1,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	3	1	-0.028	0.012	5.777	1	0.016
CSIBTEST	1	3	1	0.028	NA	5.777	1	0.016

> SIBTEST(ENG_IRL[,8:11],gr_eng_irl,2,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	3	1	1 0.037	0.012	9.91	1	0.002
CSIBTEST	1	3	1	1 0.037	NA	9.91	1	0.002

> SIBTEST(ENG_IRL[,8:11],gr_eng_irl,3,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	3	1	-0.007	0.012	0.315	1	0.575

```
CSIBTEST      1      3              1 0.007  NA  0.315 1 0.575
```

```
> SIBTEST(ENG_IRL[,8:11],gr_eng_irl,4,focal=1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2  df  p
SIBTEST      1      3      1      0.03  0.013  4.949  1  0.026
CSIBTEST      1      3      1      0.03   NA    4.949  1  0.026
```

Poly SIBTEST Yöntemine Göre İngiltere-ABD Analiz Çıktıları

```
> SIBTEST(ENG_USA[,3:7],gr_eng_usa,1,focal_name = 1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2  df  p
SIBTEST      1      4      1     -0.021  0.01   4.278  1  0.039
CSIBTEST      1      4      1     0.021   NA    4.278  1  0.039
```

```
> SIBTEST(ENG_USA[,3:7],gr_eng_usa,2,focal_name = 1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2  df  p
SIBTEST      1      4      1     -0.015  0.011  1.853  1  0.173
CSIBTEST      1      4      1     0.015   NA    1.853  1  0.173
```

```
> SIBTEST(ENG_USA[,3:7],gr_eng_usa,3,focal_name = 1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2      df  p
SIBTEST      1      4      1     0.095  0.008 131.044  1  0
CSIBTEST      1      4      1     0.095   NA   131.044  1  0
```

```
> SIBTEST(ENG_USA[,3:7],gr_eng_usa,4,focal_name = 1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2      df  p
SIBTEST      1      4      1     -0.039  0.008  21.645  1  0
CSIBTEST      1      4      1     0.039   NA   21.645  1  0
```

```
> SIBTEST(ENG_USA[,3:7],gr_eng_usa,5,focal_name = 1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2      df  p
SIBTEST      1      4      1     -0.03  0.009  11.163  1  0.001
CSIBTEST      1      4      1     0.03   NA   11.163  1  0.001
```

```
> SIBTEST(ENG_USA[,8:11],gr_eng_usa,1,focal_name = 1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2  df  p
SIBTEST      1      3      1     0.010  0.012  0.794  1  0.373
CSIBTEST      1      3      1     0.015   NA    1.560  2  0.458
```

```
> SIBTEST(ENG_USA[,8:11],gr_eng_usa,2,focal_name = 1)
```

```
      focal_group n_matched_set n_suspect_set beta  SE    X2      df  p
```

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	3	1	0.116	0.011	106.47	1	0
CSIBTEST	1	3	1	0.116	NA	106.47	1	0

> SIBTEST(ENG_USA[,8:11],gr_eng_usa,3,focal_name = 1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	3	1	-0.079	0.011	52.707	1	0
CSIBTEST	1	3	1	0.079	NA	52.707	1	0

> SIBTEST(ENG_USA[,8:11],gr_eng_usa,4,focal_name = 1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	3	1	-0.015	0.013	1.296	1	0.255
CSIBTEST	1	3	1	0.029	NA	5.068	2	0.079

Poly SIBTEST Yöntemine Göre İngiltere-Türkiye Analiz Çıktıları

> SIBTEST(ENG_TURK[,3:7],gr_eng_turk,1,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	0.108	0.011	94.623	1	0
CSIBTEST	1	4	1	0.108	NA	94.623	1	0

> SIBTEST(ENG_TURK[,3:7],gr_eng_turk,2,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	-0.227	0.011	447.197	1	0
CSIBTEST	1	4	1	0.227	NA	447.197	1	0

> SIBTEST(ENG_TURK[,3:7],gr_eng_turk,3,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2.	df	p
SIBTEST	1	4	1	0.176	0.009	356.389	1	0
CSIBTEST	1	4	1	0.176	NA	356.389	1	0

> SIBTEST(ENG_TURK[,3:7],gr_eng_turk,4,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	-0.015	0.009	2.652	1	0.103
CSIBTEST	1	4	1	0.027	NA	11.689	2	0.003

> SIBTEST(ENG_TURK[,3:7],gr_eng_turk,5,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	4	1	-0.051	0.01	28.529	1	0
CSIBTEST	1	4	1	0.051	NA	28.529	1	0

> SIBTEST(ENG_TURK[,8:11],gr_eng_turk,1,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2	df	p
SIBTEST	1	3	1	-0.019	0.012	2.513	1	0.113
CSIBTEST	1	3	1	0.007	NA.	2.520	2	0.284

> SIBTEST(ENG_TURK[,8:11],gr_eng_turk,2,focal=1)

	focal_group	n_matched_set	n_suspect_set	beta	SE	X2.	df	p

```

SIBTEST      1      3      1      0.214  0.011 356.208 1 0
CSIBTEST     1      3      1      0.214   NA   356.208 1 0
> SIBTEST(ENG_TURK[,8:11],gr_eng_turk,3,focal=1)
      focal_group n_matched_set n_suspect_set  beta  SE    X2  Df  p
SIBTEST      1      3      1      -0.092 0.012 59.704 1 0
CSIBTEST     1      3      1      0.092  NA   59.704 1 0
> SIBTEST(ENG_TURK[,8:11],gr_eng_turk,4,focal=1)
      focal_group n_matched_set n_suspect_set  beta  SE    X2  df  p
SIBTEST      1      3      1      -0.045 0.014 10.602 1 0.001
CSIBTEST     1      3      1      0.045  NA   10.602 1 0.001

```

EK-D: GMH Yöntemine Göre Yapılan Analiz Çıktıları

GMHDIF Programı İngiltere-İrlanda Analiz Çıktıları

GMHDIF Beta Version

GMH statistic = QMH1

Alpha level = 0,05

Results statistically significant at the 0,05 alpha level are marked with an asterisk.

VARIABLE 2

Stage 1: QMH = 154,8080 df = 3 p = 0,0000*

VARIABLE 3

Stage 1: QMH = 30,4224 df = 3 p = 0,0000*

VARIABLE 4

Stage 1: QMH = 63,4329 df = 3 p = 0,0000*

VARIABLE 5

Stage 1: QMH = 136,1824 df = 3 p = 0,0000*

VARIABLE 6

Stage 1: QMH = 70,5869 df = 3 p = 0,0000*

09:15 AM, Çarşamba 20-Mart-2019

Analyses time

Start = 09:15:29

End = 09:15:29

GMHDIF Beta Version

GMH statistic = QMH1

Alpha level = 0,05

Results statistically significant at the 0,05 alpha level are marked with an asterisk.

VARIABLE 7

Stage 1: QMH = 15,7465 df = 3 p = 0,0013*

VARIABLE 8

Stage 1: QMH = 9,2948 df = 3 p = 0,0256*

VARIABLE 9

Stage 1: QMH = 2,8590 df = 3 p = 0,4139

VARIABLE 10

Stage 1: QMH = 6,5562 df = 3 p = 0,0875

09:16 AM, Çarşamba 20-Mart-2019

Analyses time

Start = 09:16:31

End = 09:16:31

GMHDIF Programı İngiltere-ABD Analiz Çıktıları

GMHDIF Beta Version

GMH statistic = QMH1

Alpha level = 0,05

Results statistically significant at the 0,05 alpha level are marked with an asterisk.

VARIABLE 2

Stage 1: QMH = 15,3383 df = 3 p = 0,0015*

VARIABLE 3

Stage 1: QMH = 17,6073 df = 3 p = 0,0005*

VARIABLE 4

Stage 1: QMH = 167,7465 df = 3 p = 0,0000*

VARIABLE 5

Stage 1: QMH = 27,2388 df = 3 p = 0,0000*

VARIABLE 6

Stage 1: QMH = 16,8652 df = 3 p = 0,0008*

09:18 AM, Çarşamba 20-Mart-2019

Analyses time

Start = 09:18:26FOSE78

GMHDIF Beta Version

GMH statistic = QMH1

Alpha level = 0,05

Results statistically significant at the 0,05 alpha level are marked with an asterisk.

VARIABLE 7

Stage 1: QMH = 13,9241 df = 3 p = 0,0030*

VARIABLE 8

Stage 1: QMH = 106,8313 df = 3 p = 0,0000*

VARIABLE 9

Stage 1: QMH = 60,1230 df = 3 p = 0,0000*

VARIABLE 10

Stage 1: QMH = 6,5779 df = 3 p = 0,0866

09:19 AM, Çarşamba 20-Mart-2019

Analyses time

Start = 09:19:20

End = 09:19:20

GMHDIF Programı İngiltere-Türkiye Analiz Çıktıları

GMHDIF Beta Version

GMH statistic = QMH1

Alpha level = 0,05

Results statistically significant at the 0,05 alpha level are marked with an asterisk.

VARIABLE 2

Stage 1: QMH = 111,9338 df = 3 p = 0,0000*

VARIABLE 3

Stage 1: QMH = 434,7852 df = 3 p = 0,0000*

VARIABLE 4

Stage 1: QMH = 458,5244 df = 3 p = 0,0000*

VARIABLE 5

Stage 1: QMH = 16,5596 df = 3 p = 0,0009*

VARIABLE 6

Stage 1: QMH = 52,6104 df = 3 p = 0,0000*

09:09 AM, Çarşamba 20-Mart-2019

Analyses time

Start = 09:09:17

End = 09:09:17

GMHDIF Beta Version

GMH statistic = QMH1

Alpha level = 0,05

Results statistically significant at the 0,05 alpha level are marked with an asterisk.

VARIABLE 7

Stage 1: QMH = 7,6756 df = 3 p = 0,0532

VARIABLE 8

Stage 1: QMH = 394,4942 df = 3 p = 0,0000*

VARIABLE 9

Stage 1: QMH = 90,3939 df = 3 p = 0,0000*

VARIABLE 10

Stage 1: QMH = 68,8419 df = 3 p = 0,0000*

09:10 AM, Çarşamba 20-Mart-2019

Analyses time

Start = 09:10:21

End = 09:10:21

DIFAS Programı İngiltere-İrlanda Analiz Çıktıları

Opened the text file: C:\Users\kkata\YandexDisk\ENGIURL.txt

Number of Cases: 9574

Number of Variables: 10

DESCRIPTIVES

Name	Mean	SD	Min	Max	N
Var 1	1.5386	0.4985	1	2	9574
Var 2	2.7328	0.839	1	4	9574
Var 3	2.58	0.8668	1	4	9574
Var 4	2.7991	0.803	1	4	9574
Var 5	2.8747	0.8213	1	4	9574
Var 6	2.85	0.8528	1	4	9574
Var 7	1.9347	0.8231	1	4	9574
Var 8	2.1014	0.8976	1	4	9574
Var 9	1.9757	0.8355	1	4	9574

Var 10 2.0743 0.8638 1 4 9574

FREQUENCIES AND PERCENTAGES FOR: Var 1

Value	Freq	% Total	% Pres	Cum %
-------	------	---------	--------	-------

1	4417	46.14	46.14	46.14
2	5157	53.86	53.86	100

Present	9574	100
Missing	0	0
Total	9574	100

FREQUENCIES AND PERCENTAGES FOR: Var 2

Value	Freq	% Total	% Pres	Cum %
-------	------	---------	--------	-------

1	928	9.69	9.69	9.69
2	2206	23.04	23.04	32.73
3	4936	51.56	51.56	84.29
4	1504	15.71	15.71	100

Present	9574	100
Missing	0	0
Total	9574	100

FREQUENCIES AND PERCENTAGES FOR: Var 3

Value	Freq	% Total	% Pres	Cum %
-------	------	---------	--------	-------

1	1107	11.56	11.56	11.56
2	3130	32.69	32.69	44.26
3	4014	41.93	41.93	86.18
4	1323	13.82	13.82	100

Present	9574	100
Missing	0	0
Total	9574	100

FREQUENCIES AND PERCENTAGES FOR: Var 4

 Value Freq % Total % Pres Cum %

1	832	8.69	8.69	8.69
2	1745	18.23	18.23	26.92
3	5511	57.56	57.56	84.48
4	1486	15.52	15.52	100

Present	9574	100
Missing	0	0
Total	9574	100

 FREQUENCIES AND PERCENTAGES FOR: Var 5

 Value Freq % Total % Pres Cum %

1	800	8.36	8.36	8.36
2	1504	15.71	15.71	24.07
3	5366	56.05	56.05	80.11
4	1904	19.89	19.89	100

Present	9574	100
Missing	0	0
Total	9574	100

 FREQUENCIES AND PERCENTAGES FOR: Var 6

 Value Freq % Total % Pres Cum %

1	852	8.9	8.9	8.9
2	1751	18.29	18.29	27.19
3	4952	51.72	51.72	78.91
4	2019	21.09	21.09	100

Present	9574	100
Missing	0	0
Total	9574	100

 FREQUENCIES AND PERCENTAGES FOR: Var 7

Value	Freq	% Total	% Pres	Cum %
1	3094	32.32	32.32	32.32
2	4493	46.93	46.93	79.25
3	1505	15.72	15.72	94.97
4	482	5.03	5.03	100

Present	9574	100
Missing	0	0
Total	9574	100

 FREQUENCIES AND PERCENTAGES FOR: Var 8

Value	Freq	% Total	% Pres	Cum %
1	2766	28.89	28.89	28.89
2	3725	38.91	38.91	67.8
3	2429	25.37	25.37	93.17
4	654	6.83	6.83	100

Present	9574	100
Missing	0	0
Total	9574	100

 FREQUENCIES AND PERCENTAGES FOR: Var 9

Value	Freq	% Total	% Pres	Cum %
1	2939	30.7	30.7	30.7
2	4451	46.49	46.49	77.19
3	1662	17.36	17.36	94.55
4	522	5.45	5.45	100

Present	9574	100
Missing	0	0
Total	9574	100

FREQUENCIES AND PERCENTAGES FOR: Var 10

Value	Freq	% Total	% Pres	Cum %
1	2617	27.33	27.33	27.33
2	4254	44.43	44.43	71.77
3	2078	21.7	21.7	93.47
4	625	6.53	6.53	100

Present	9574	100
Missing	0	0
Total	9574	100

DIF analysis: Nonparametric tests for polytomous items

Stratifying variable: Sum of item responses

Stratum size: 1

Number of strata: 16

Number of reference group members: 4417

Number of focal group members: 5157

Grouping variable: Var1

Reference Value = 1, Focal Value = 2

DIF STATISTICS: POLYTOMOUS ITEMS

Name	Mantel	L-A LOR	LOR SE	LOR Z	COX'S B	COX SE	COX Z
Var 2	154.648	0.784	0.064	12.25	0.717	0.0577	12.426
Var 3	12.967	-0.232	0.064	-3.625	-0.221	0.0613	-3.605
Var 4	58.969	0.592	0.078	7.59	0.582	0.0758	7.678
Var 5	113.339	-0.825	0.08	-10.312	-0.814	0.0765	-10.641
Var 6	69.264	-0.604	0.073	-8.274	-0.583	0.07	-8.329

Reference Value = 1, Focal Value = 2

CONDITIONAL DIFFERENCES: Intervals of size 1.5

	5	6.5	8	9.5	11	12.5	14	15.5	17	18.5
Lower	5	6.5	8	9.5	11	12.5	14	15.5	17	18.5
Upper	6.5	8	9.5	11	12.5	14	15.5	17	18.5	20.1

Var 2	0.02	0.11	0.32	0.13	0.16	0.15	0.01	0.16	0.2	0.04
Var 3	0.01	0.05	-0.08	0.03	-0.01	-0.09	0	-0.05	-0.13	-0.03
Var 4	0.02	0.14	0.12	0.05	0.21	0.07	0	0.06	0.05	-0.01

Var 5 0 -0.04 -0.07 -0.17 -0.18 -0.07 -0.01 -0.08 -0.08 -0.01
 Var 6 -0.03 -0.16 -0.29 -0.06 -0.18 -0.05 -0.01 -0.11 -0.04 -0.01

 DIF analysis: Nonparametric tests for polytomous items

Stratifying variable: Sum of item responses

Stratum size: 1

Number of strata: 13

Number of reference group members: 4417

Number of focal group members: 5157

Grouping variable: Var1

Reference Value = 1, Focal Value = 2

DIF STATISTICS: POLYTOMOUS ITEMS

Name	Mantel	L-A LOR	LOR SE	LOR Z	COX'S B	COX SE	COX Z
Var 7	11.876	-0.22	0.064	-3.438	-0.206	0.0599	-3.439
Var 8	3.357	0.121	0.066	1.833	0.117	0.0638	1.834
Var 9	1.035	-0.066	0.065	-1.015	-0.062	0.0613	-1.011
Var 10	5.653	0.139	0.058	2.397	0.124	0.0523	2.371

Reference Value = 1, Focal Value = 2

CONDITIONAL DIFFERENCES: Intervals of size 1.2

Lower	4	5.2	6.4	7.6	8.8	10	11.2	12.4	13.6	14.8
Upper	5.2	6.4	7.6	8.8	10	11.2	12.4	13.6	14.8	16.1

Var 7	-0.02	-0.07	-0.1	-0.01	-0.02	-0.01	0.02	-0.26	-0.04	-0.06
Var 8	-0.01	-0.06	0.06	0	0.06	0	0.02	0.22	-0.01	0.01
Var 9	-0.01	-0.02	0.02	0	0	-0.05	-0.03	0.15	0.05	-0.02
Var 10	-0.01	0.15	0.01	0.02	0.01	0.06	-0.02	-0.11	-0.01	-0.01

DIFAS Programı İngiltere-ABD Analiz Çıktıları

Opened the text file: C:\Users\kkata\YandexDisk\ENGUSA.txt

Number of Cases: 9417

Number of Variables: 10

DESCRIPTIVES

Name	Mean	SD	Min	Max	N
------	------	----	-----	-----	---

Var 1	1.531	0.4991	1	2	9417
Var 2	2.8069	0.8176	1	4	9417
Var 3	2.593	0.8411	1	4	9417
Var 4	2.8045	0.7918	1	4	9417
Var 5	2.871	0.804	1	4	9417
Var 6	2.8503	0.8366	1	4	9417
Var 7	1.9367	0.7979	1	4	9417
Var 8	2.0904	0.862	1	4	9417
Var 9	2.0123	0.8202	1	4	9417
Var 10	2.0995	0.8467	1	4	9417

 FREQUENCIES AND PERCENTAGES FOR: Var 1

Value	Freq	% Total	% Pres	Cum %
1	4417	46.9	46.9	46.9
2	5000	53.1	53.1	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 2

Value	Freq	% Total	% Pres	Cum %
1	762	8.09	8.09	8.09
2	1946	20.66	20.66	28.76
3	5057	53.7	53.7	82.46
4	1652	17.54	17.54	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 3

Value	Freq	% Total	% Pres	Cum %
1	956	10.15	10.15	10.15

2	3159	33.55	33.55	43.7
3	4064	43.16	43.16	86.85
4	1238	13.15	13.15	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 4

Value	Freq	% Total	% Pres	Cum %
1	718	7.62	7.62	7.62
2	1898	20.16	20.16	27.78
3	5308	56.37	56.37	84.15
4	1493	15.85	15.85	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 5

Value	Freq	% Total	% Pres	Cum %
1	689	7.32	7.32	7.32
2	1662	17.65	17.65	24.97
3	5241	55.65	55.65	80.62
4	1825	19.38	19.38	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 6

Value	Freq	% Total	% Pres	Cum %
1	750	7.96	7.96	7.96

2	1856	19.71	19.71	27.67
3	4865	51.66	51.66	79.34
4	1946	20.66	20.66	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 7

Value	Freq	% Total	% Pres	Cum %
1	2872	30.5	30.5	30.5
2	4711	50.03	50.03	80.52
3	1392	14.78	14.78	95.31
4	442	4.69	4.69	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 8

Value	Freq	% Total	% Pres	Cum %
1	2523	26.79	26.79	26.79
2	4108	43.62	43.62	70.42
3	2198	23.34	23.34	93.76
4	588	6.24	6.24	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 9

Value	Freq	% Total	% Pres	Cum %
1	2613	27.75	27.75	27.75
2	4572	48.55	48.55	76.3

3	1735	18.42	18.42	94.72
4	497	5.28	5.28	100

Present	9417	100
Missing	0	0
Total	9417	100

 FREQUENCIES AND PERCENTAGES FOR: Var 10

Value	Freq	% Total	% Pres	Cum %
1	2360	25.06	25.06	25.06
2	4353	46.22	46.22	71.29
3	2111	22.42	22.42	93.7
4	593	6.3	6.3	100

Present	9417	100
Missing	0	0
Total	9417	100

 DIF analysis: Nonparametric tests for polytomous items

Stratifying variable: Sum of item responses

Stratum size: 1

Number of strata: 16

Number of reference group members: 4417

Number of focal group members: 5000

Grouping variable: Var1

Reference Value = 1, Focal Value = 2

 DIF STATISTICS: POLYTOMOUS ITEMS

Name	Mantel	L-A LOR	LOR SE	LOR Z	COX'S B	COX SE	COX Z
Var 2	5.147	-0.148	0.065	-2.277	-0.14	0.0619	-2.262
Var 3	0.126	-0.024	0.069	-0.348	-0.023	0.0643	-0.358
Var 4	134.756	0.894	0.079	11.316	0.863	0.0744	11.599
Var 5	20.698	-0.355	0.079	-4.494	-0.354	0.0778	-4.55
Var 6	15.616	-0.288	0.073	-3.945	-0.282	0.0713	-3.955

 Reference Value = 1, Focal Value = 2

CONDITIONAL DIFFERENCES: Intervals of size 1.5

Lower	5	6.5	8	9.5	11	12.5	14	15.5	17	18.5
Upper	6.5	8	9.5	11	12.5	14	15.5	17	18.5	20.1

Var 2	0.01	0.22	0.21	0	-0.03	-0.05	-0.01	0.01	-0.08	-0.02
Var 3	0	-0.04	-0.06	-0.01	0.02	-0.04	0	0.01	0.16	0.02
Var 4	0.03	0.15	0.1	0.09	0.31	0.16	0.01	0.03	-0.01	0.01
Var 5	0	-0.15	0.02	-0.05	-0.16	-0.03	0	-0.01	-0.06	0
Var 6	-0.01	-0.22	-0.27	0	-0.14	0.01	0	-0.06	-0.02	-0.01

DIF analysis: Nonparametric tests for polytomous items

Stratifying variable: Sum of item responses

Stratum size: 1

Number of strata: 13

Number of reference group members: 4417

Number of focal group members: 5000

Grouping variable: Var1

Reference Value = 1, Focal Value = 2

DIF STATISTICS: POLYTOMOUS ITEMS

Name	Mantel	L-A LOR	LOR SE	LOR Z	COX'S B	COX SE	COX Z
Var 7	1.835	0.089	0.066	1.348	0.084	0.0619	1.357
Var 8	69.274	0.578	0.07	8.257	0.563	0.0676	8.328
Var 9	56.121	-0.515	0.07	-7.357	-0.49	0.0654	-7.492
Var 10	2.833	-0.103	0.062	-1.661	-0.093	0.0552	-1.685

Reference Value = 1, Focal Value = 2

CONDITIONAL DIFFERENCES: Intervals of size 1.2

Lower	4	5.2	6.4	7.6	8.8	10	11.2	12.4	13.6	14.8
Upper	5.2	6.4	7.6	8.8	10	11.2	12.4	13.6	14.8	16.1

Var 7	-0.01	0.01	0.01	-0.02	0.02	0.2	-0.02	0.06	0.27	-0.06
Var 8	0	-0.03	0.08	0.02	0.21	0.03	0.09	0.17	0	0.01
Var 9	0	-0.09	-0.07	-0.02	-0.11	-0.13	-0.06	-0.12	-0.13	-0.03
Var 10	0.01	0.11	-0.02	0.01	-0.09	-0.1	0	-0.11	-0.14	-0.01

DIFAS Programı İngiltere-Türkiye Analiz Çıktıları

Opened the text file: C:\Users\kkata\YandexDisk\ENGTURK.txt

Number of Cases: 9506

Number of Variables: 10

DESCRIPTIVES

Name	Mean	SD	Min	Max	N
Var 1	1.5353	0.4988	1	2	9506
Var 2	2.7063	0.8797	1	4	9506
Var 3	2.6224	0.8809	1	4	9506
Var 4	2.7216	0.8588	1	4	9506
Var 5	2.8023	0.8667	1	4	9506
Var 6	2.7969	0.881	1	4	9506
Var 7	1.9315	0.8096	1	4	9506
Var 8	2.0398	0.8449	1	4	9506
Var 9	1.9982	0.8364	1	4	9506
Var 10	2.0933	0.8684	1	4	9506

FREQUENCIES AND PERCENTAGES FOR: Var 1

Value	Freq	% Total	% Pres	Cum %
1	4417	46.47	46.47	46.47
2	5089	53.53	53.53	100

Present	9506	100
Missing	0	0
Total	9506	100

FREQUENCIES AND PERCENTAGES FOR: Var 2

Value	Freq	% Total	% Pres	Cum %
1	1123	11.81	11.81	11.81
2	2115	22.25	22.25	34.06
3	4699	49.43	49.43	83.49
4	1569	16.51	16.51	100

Present	9506	100
Missing	0	0
Total	9506	100

 FREQUENCIES AND PERCENTAGES FOR: Var 3

Value	Freq	% Total	% Pres	Cum %
1	1122	11.8	11.8	11.8
2	2794	29.39	29.39	41.2
3	4141	43.56	43.56	84.76
4	1449	15.24	15.24	100

Present	9506	100
Missing	0	0
Total	9506	100

 FREQUENCIES AND PERCENTAGES FOR: Var 4

Value	Freq	% Total	% Pres	Cum %
1	1029	10.82	10.82	10.82
2	2109	22.19	22.19	33.01
3	4847	50.99	50.99	84
4	1521	16	16	100

Present	9506	100
Missing	0	0
Total	9506	100

 FREQUENCIES AND PERCENTAGES FOR: Var 5

Value	Freq	% Total	% Pres	Cum %
1	975	10.26	10.26	10.26
2	1770	18.62	18.62	28.88
3	4920	51.76	51.76	80.63
4	1841	19.37	19.37	100

Present	9506	100
---------	------	-----

Missing	0	0
Total	9506	100

 FREQUENCIES AND PERCENTAGES FOR: Var 6

Value	Freq	% Total	% Pres	Cum %
1	1009	10.61	10.61	10.61
2	1823	19.18	19.18	29.79
3	4764	50.12	50.12	79.91
4	1910	20.09	20.09	100

Present	9506	100
Missing	0	0
Total	9506	100

 FREQUENCIES AND PERCENTAGES FOR: Var 7

Value	Freq	% Total	% Pres	Cum %
1	2981	31.36	31.36	31.36
2	4677	49.2	49.2	80.56
3	1366	14.37	14.37	94.93
4	482	5.07	5.07	100

Present	9506	100
Missing	0	0
Total	9506	100

 FREQUENCIES AND PERCENTAGES FOR: Var 8

Value	Freq	% Total	% Pres	Cum %
1	2649	27.87	27.87	27.87
2	4392	46.2	46.2	74.07
3	1903	20.02	20.02	94.09
4	562	5.91	5.91	100

Present	9506	100
Missing	0	0

Total 9506 100

FREQUENCIES AND PERCENTAGES FOR: Var 9

Value	Freq	% Total	% Pres	Cum %
1	2777	29.21	29.21	29.21
2	4525	47.6	47.6	76.81
3	1648	17.34	17.34	94.15
4	556	5.85	5.85	100

Present	9506	100
Missing	0	0
Total	9506	100

FREQUENCIES AND PERCENTAGES FOR: Var 10

Value	Freq	% Total	% Pres	Cum %
1	2484	26.13	26.13	26.13
2	4349	45.75	45.75	71.88
3	1975	20.78	20.78	92.66
4	698	7.34	7.34	100

Present	9506	100
Missing	0	0
Total	9506	100

DIF analysis: Nonparametric tests for polytomous items

Stratifying variable: Sum of item responses

Stratum size: 1

Number of strata: 16

Number of reference group members: 4417

Number of focal group members: 5089

Grouping variable: Var1

Reference Value = 1, Focal Value = 2

DIF STATISTICS: POLYTOMOUS ITEMS

Name Mantel L-A LOR LOR SE LOR Z COX'S B COX SE COX Z

Var 2	96.998	0.634	0.065	9.754	0.559	0.0568	9.842
Var 3	410.485	-1.361	0.069	-19.725	-1.2	0.0592	-20.27
Var 4	347.205	1.304	0.077	16.935	1.237	0.0664	18.63
Var 5	2.784	-0.123	0.074	-1.662	-0.121	0.0726	-1.667
Var 6	36.788	-0.429	0.07	-6.129	-0.413	0.068	-6.074

Reference Value = 1, Focal Value = 2

CONDITIONAL DIFFERENCES: Intervals of size 1.5

Lower	5	6.5	8	9.5	11	12.5	14	15.5	17	18.5
Upper	6.5	8	9.5	11	12.5	14	15.5	17	18.5	20.1

Var 2	0.04	0.28	0.32	0.05	0.19	0.17	0.02	0.14	0.22	-0.02
Var 3	-0.01	-0.42	-0.4	-0.07	-0.28	-0.42	-0.01	-0.18	-0.2	-0.11
Var 4	0.03	0.2	0.25	0.15	0.44	0.37	0.01	-0.01	-0.01	0.02
Var 5	0	0.2	0.07	-0.03	-0.1	0.01	0	-0.04	-0.04	0.01
Var 6	-0.02	-0.04	-0.25	-0.06	-0.25	-0.05	-0.01	0.05	0.04	0

DIF analysis: Nonparametric tests for polytomous items

Stratifying variable: Sum of item responses

Stratum size: 1

Number of strata: 13

Number of reference group members: 4417

Number of focal group members: 5089

Grouping variable: Var1

Reference Value = 1, Focal Value = 2

DIF STATISTICS: POLYTOMOUS ITEMS

Name	Mantel	L-A LOR	LOR SE	LOR Z	COX'S B	COX SE	COX Z
Var 7	6.268	-0.16	0.064	-2.5	-0.146	0.0583	-2.504
Var 8	275.484	1.086	0.068	15.971	1.034	0.0623	16.597
Var 9	67.881	-0.535	0.066	-8.106	-0.494	0.06	-8.233
Var 10	19.354	-0.259	0.06	-4.317	-0.225	0.051	-4.412



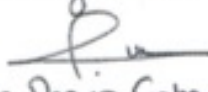
Reference Value = 1, Focal Value = 2

CONDITIONAL DIFFERENCES: Intervals of size 1.2

Lower	4	5.2	6.4	7.6	8.8	10	11.2	12.4	13.6	14.8
Upper	5.2	6.4	7.6	8.8	10	11.2	12.4	13.6	14.8	16.1

Var 7	-0.02	-0.04	-0.12	-0.01	-0.02	0.08	0.01	0.04	-0.16	-0.09
Var 8	0	-0.08	0.17	0.03	0.4	0.25	0.14	0.53	0.4	0.04
Var 9	-0.01	0.05	0.02	-0.03	-0.19	-0.15	-0.1	-0.31	-0.09	-0.02
Var 10	0.03	0.08	-0.07	0	-0.18	-0.18	-0.05	-0.26	-0.15	-0.02

EK-E: Etik Kurul İzin Muafiyet Formu

	HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ TEZ ÇALIŞMASI ETİK KURUL İZİN MUAFİYETİ FORMU
HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ EĞİTİM BİLİMLERİ ANABİLİM DALI BAŞKANLIĞI'NA	
Tarih: 05/07/2016	
Tez Başlığı / Konusu: <u>2012 Yılında Uygulanan PISA Matematik Testindeki Maddelerin Türkiye'de Bölgelere Göre Yetkinlik İncelenmesi</u>	
Yukarıda başlığı/konusu gösterilen tez çalışmam:	
<ol style="list-style-type: none">1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır,2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmasını gerektirmemektedir.3. Beden bütünlüğüne müdahale içermemektedir.4. Gözlemsel ve betimsel araştırma (anket, ölçek/skala çalışmaları, dosya taramaları, veri kaynakları taraması, sistem-model geliştirme çalışmaları) niteliğinde değildir.	
Hacettepe Üniversitesi Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.	
Gereğini saygılarımla arz ederim.	
Adı Soyadı: <u>Emre Gül ÇÖR</u>	Tarih ve İmza
Öğrenci No: <u>N12220608</u>	
Anabilim Dalı: <u>Eğitim Bilimleri</u>	
Programı: <u>Eğitimde Ölçme ve Değerlendirme</u>	
Statüsü: <input checked="" type="checkbox"/> Y.Lisans <input type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr.	
DANISMAN GÖRÜŞÜ VE ONAYI	
<u>Etik kuruldan izin alınmasına gerek yoktur.</u>	
 Yrd. Doç. Dr. Derin Çobanoğlu Aktaş (Unvan, Ad Soyad, İmza)	
Telefon: 0-312-2976860	Detaylı Bilgi: http://www.sosyalbilimler.hacettepe.edu.tr
	Faks: 0-3122992147
	E-posta: sosyalbilimler@hacettepe.edu.tr

EK-F: Etik Beyanı

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

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

beyan ederim.

30.10.2019

Esengül GÜR



EK-G: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu

26/06/2019

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

Tez Başlığı : PISA 2015 Uygulamasındaki Maddelerin Kültüre Göre Değişen Madde Fonksiyonu Açısından İncelenmesi.

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

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
26/06/2019	135	182129	30/05/2019	%7	1147170961

Uygulanan filtreler:

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

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

Ad Soyad: Esengül GÜR

Öğrenci No.: N12220408

Ana Bilim Dalı: Eğitim Bilimleri

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü: Y.Lisans Doktora Bütünleşik Dr.

İmza

DANIŞMAN ONAYI

UYGUNDUR.

Dr. Öğretim Üyesi Kübra ATALAY KABASAKAL

EK-H: Thesis/Dissertation Originality Report

26/06/2019

HACETTEPE UNIVERSITY
Graduate School of Educational Sciences
To The Department of Educational Sciences

Thesis Title: An Investigation Of The PISA 2015In Terms Of Differantial Item Functioning Based On Culturing

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

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
26/06/2019	135	182129	30/05/2019	%7	1147170961

Filtering options applied:

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

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

I respectfully submit this for approval.

Name Lastname: Esengül GÜR
Student No.: N12220408
Department: Educational Sciences
Program: Measurement and Evaluation in Education
Status: Masters Ph.D. Integrated Ph.D.


Signature

ADVISOR APPROVAL


APPROVED
Assist. Prof. Dr. Kubra ATALAY KABASAKAL

EK-I: Yayımlama ve Fikri Mülkiyet Hakları Beyanı

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

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

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

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

...30.05.2019
Esengül GÜR

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

- (1) Madde 6.1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metallerin kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3 şahıslara veya kurumlara haksız kazanç, imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7.1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

* Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

