



HACETTEPE ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı

Eğitimde Ölçme ve Değerlendirme Programı

ORTAK DEĞİŞKENLİ ÇOK BOYUTLU ÇOK DÜZEYLİ KARMA MODEL İLE DMF İNCELEMESİ

Ömer DOĞAN

Doktora Tezi

Ankara, 2024

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

ORTAK DEĞİŞKENLİ ÇOK BOYUTLU ÇOK DÜZEYLİ KARMA MODEL İLE DMF
İNCELEMESİ

DIF ANALYSIS WITH A MULTIDIMENSIONAL MULTILEVEL MIXTURE MODEL WITH
COVARIATES

Ömer DOĞAN

Doktora Tezi

Ankara, 2024

Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

mer DOĐAN'ın hazırladıđı "Ortak Deđiřkenli ok Boyutlu ok D¼zeyli Karma Model ile DMF İncelemesi" bařlıklı bu alıřma j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde lme ve Deđerlendirme Bilim Dalında Doktora Tezi** olarak kabul edilmiřtir.

J¼ri Bařkanı	Prof. Dr. H¼lya KELECİĐLU	İmza
J¼ri Üyesi (Danıřman)	Prof. Dr. Burcu ATAR	İmza
J¼ri Üyesi	Prof. Dr. Dilara BAKAN KALAYCIOĐLU	İmza
J¼ri Üyesi	Do. Dr. Derya OBANOĐLU AKTAN	İmza
J¼ri Üyesi	Do. Dr. Melek G¼řah řAHİN	İmza

Bu tez Hacettepe niversitesi Lisans¼st¼ Eđitim, đretim ve Sınav Ynetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 27 / 06 / 2024 tarihinde uygun gr¼lm¼ř ve Enstit¼ Ynetim Kurulunca / / tarihi itibarıyla kabul edilmiřtir.

Prof. Dr. İsmail Hakkı MİRİCİ
Eđitim Bilimleri Enstitüsü M¼d¼r¼

Öz

Bu araştırmanın amacı test sürecinde geçerliğe ait kanıt toplamak için yararlanılan araçlardan biri olan değişen madde fonksiyonunu (DMF) kullanarak gözlenen grupları ve karma modellerle oluşturulan örtük sınıfları DMF'li madde sayısı ve etki büyüklüğü düzeyleri açısından karşılaştırmak ve DMF'nin doğasını araştırmaktır. Bu doğrultuda 22 ülkenin yer aldığı eTIMSS 2019 1. kitapçığında yer alan 10 matematik ve 10 fen bilimleri olmak üzere 20 maddelik veri seti oluşturulmuştur. Yapısı incelenen veri setinin çok boyutlu ve çok düzeyli bir yapı sergilediği tespit edilmiştir. Analizler öncesi düzey, boyut ve ortak değişken durumuna göre oluşturulan 54 model için bilgi ölçütü kriterleri incelenmiş ve uygun modellere ait betimsel istatistikler oluşturulmuştur. Gözlenen grup değişkenleri olarak alınan cinsiyet ve evdeki kitap sayısı değişkenleri, karma modellerden ise veriye en iyi uyumu gösteren modeller için DMF analizleri yapılmıştır. Karma modellere ortak değişken olarak öğrenci düzeyinde toplam puan ile en yüksek korelasyona sahip evdeki kitap sayısı değişkeni, ülke düzeyinde ise her ülkeye ait katılımcıların evdeki kitap sayısı ortalaması seçilmiştir. Araştırma sonuçlarına göre her iki gözlenen grup değişkeni için sadece A etki düzeyinde ihmal edilebilir DMF'li maddeler mevcutken karma modellere göre B ve C etki düzeyinde birçok madde tespit edilmiştir. Her model için oluşturulan örtük ülke sınıflarında yüksek ve düşük puanlı sınıflara ait ülke dağılımları incelenmiştir. Model karşılaştırmaları sonucunda 2 PL modelin veriye daha iyi uyum gösterdiği bulunmuştur. Ayrıca en iyi uyumu sağlayan CB2-C2 adlı model üzerinde yapılan inceleme ile veri seti bağlamında matematik ve fen bilimleri maddelerinin konu ve bilişsel alanı ile örtük sınıfların yapısıyla DMF kaynakları daha net biçimde ortaya çıkmıştır.

Anahtar sözcükler: değişen madde fonksiyonu, karma model, çok boyutluluk, çok düzeylilik, ortak değişken, örtük sınıf

Abstract

The aim of this study is to compare the manifest groups using the differential item functioning (DIF) and the latent classes created with mixture models in terms of the number of items with DIF and effect size levels, and to investigate the nature of the DIF. In this direction, a data set of 20 items (10 mathematics and 10 science items) from eTIMSS 2019 booklet 1, which includes 22 countries, was created. The structure of the data set was analyzed and it was determined that it displayed a multidimensional and multilevel structure. DIF analyses were conducted for mixture models with the variables of gender and the number of books at home, which were taken as manifest group variables. The number of books at home was selected as a covariate for the mixture models. According to the results of the study, for the manifest group variables, there were negligible DIF items only at effect level A, while many items were found at effect levels B and C according to the mixture models. In the latent country classes created for each model, the country distributions of the high and low scoring classes were examined. As a result of the model comparisons, it was found that the 2 PL model fit the data better. In addition, in the context of the data set, the subject and cognitive domain of mathematics and science items and the structure of latent classes and DIF sources emerged more clearly.

Keywords: differential item functioning, mixture model, multidimensionality, multilevelness, covariate, latent class

Teşekkür

Danışman seçiminden tezin teslimine kadar geçen zamanın tamamında ilgisi, desteği, sabrı, nezaketi, güveni ve akademik bilgisi ile muazzam katkı sağlayan, tez konusu seçiminden tez sürecinin son aşamasına kadar açık fikirli, yol gösterici ve cesaretlendirici kişiliğiyle bu çalışmanın ortaya çıkmasını sağlayan tez danışmanın Prof. Dr. Burcu ATAR'a en derin teşekkürlerimi sunarım.

Tez İzleme Komisyonu'nda yer alıp tezin süreç içindeki değişimlerinde farklı bakış açılarını ortaya koyup tezin çok daha iyi hale gelmesini sağlayan, dikkatli gözlerle, içtenlikle uyarılarını yapan Prof. Dr. Dilara BAKAN KALAYCIOĞLU ve Doç. Dr. Derya ÇOBANOĞLU AKTAN'a çok teşekkür ederim. Tez savunma jürimde yer alıp tezin son şeklinin verilmesinde önemli katkıları olan Prof. Dr. Hülya KELECİOĞLU ve Doç. Dr. Melek Gülşah ŞAHİN'e çok teşekkür ederim. Ders sürecinden itibaren yaklaşımları, sıcaklıkları, örneklikleri ve samimiyetleriyle katkı sunan hocalarım Prof. Dr. Selahattin GELBAL'a, Prof. Dr. Nuri DOĞAN'a ve Doç. Dr. Kübra ATALAY KABASAKAL'a çok teşekkür ederim. Üretkenliği ve çalışkanlığıyla örnek olan ölçme değerlendirme alanındaki yüksek lisans danışmanın Prof. Dr. İsmail KARAKAYA'ya çok teşekkür ederim. İlk yüksek lisans tez sürecimden itibaren dostluklarını her zaman hissettiğim değerli arkadaşlarım Dr. Ali TOSUN'a ve Dr. Yusuf GİDİŞ'e çok teşekkür ederim.

Uzun soluklu lisansüstü eğitim sürecinde her daim destekçim olan, eksiklerimi kapatan, önüme çıkan engelleri beraber kaldırdığımız babam Himmet DOĞAN'a, annem Ayşe DOĞAN'a, kardeşim Hüseyin DOĞAN'a, eşi Ayşe DOĞAN'a, oğulları Yusuf DOĞAN'a çok teşekkür ederim. Son olarak da en büyük teşekkürü hak eden, en zoru kolaylaştırıp gücüme güç katan, vakitlerini çaldığım çekirdek ailemin üyeleri eşim Cennet Kevser DOĞAN'a, oğullarım Furkan ve Muhammed Tarık DOĞAN'a çok çok teşekkür ederim.

İçindekiler

Kabul ve Onay.....	ii
Öz.....	iii
Abstract.....	iv
Teşekkür.....	v
Tablolar Dizini.....	viii
Şekiller Dizini.....	x
Simgeler ve Kısaltmalar Dizini.....	xi
Bölüm 1 Giriş.....	1
Problem Durumu.....	1
Araştırmanın Amacı ve Önemi.....	12
Araştırma Problemi.....	14
Sayıtlılar.....	15
Sınırlılıklar.....	15
Bölüm 2.....	16
Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	16
DMF, DMF Metotları ve DMF Kaynağı.....	16
İlgili Araştırmalar.....	30
Bölüm 3 Yöntem.....	38
Araştırmanın Türü.....	38
Araştırmanın Örnekleme.....	38
Veri Toplama Süreci.....	39
Veri Toplama Aracı.....	39
Verilerin Analizi.....	42
Bölüm 4 Bulgular, Yorumlar ve Tartışma.....	66
Birinci Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma.....	66
İkinci Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma.....	77

Üçüncü Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma	81
Dördüncü Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma.....	85
Bölüm 5 Sonuç ve Öneriler.....	93
Sonuçlar	93
Öneriler	96
Kaynaklar	99
EK A: KMTK Modeline Ait Mplus Kodlarından Biri.....	115
EK B: ÇDKMTK Modeline Ait Mplus Kodlarından Biri	117
EK C: ÇDÇBKMTK Modeline Ait Mplus Kodlarından Biri	120
EK D: odÇDÇBKMTK Modeline Ait Mplus Kodlarından Biri	124
EK E: Araştırma Etik Komisyon İzin Muafiyeti Formu	128
EK-F: Etik Beyanı	129
EK-G: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu	130
EK-H: Thesis/Dissertation Originality Report.....	131
EK-I: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı	132

Tablolar Dizini

Tablo 1 <i>Kitapçık 1 için Madde Türü, Konu Alanı ve Bilişsel Alanlara Ait Bilgiler ..</i>	41
Tablo 2 <i>20 Maddelik Veri Setinin Tek ve İki Boyutlu Yapı Sergilemesi Durumunda DFA Sonuçları</i>	49
Tablo 3 <i>Matematik ile Fen Bilimleri Verilerine Ait 20 Madde İçin Konu Alanları ve Bilişsel Alanlar</i>	51
Tablo 4 <i>eTIMSS 2019 Katılımcısı 22 Ülkeye Ait Betimsel İstatistikler.....</i>	52
Tablo 5 <i>Veri Setine Ait Betimsel İstatistikler.....</i>	53
Tablo 6 <i>Model İsimlerine Ait Örnekler ve Açıklamalar</i>	58
Tablo 7 <i>Bilgi Ölçütü İndeksi En İyi Olan Modeller</i>	59
Tablo 8 <i>İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 KMTK Modelinde Öğrenci ve Ülke Düzeyi Örtük Sınıfları için Nihai Sınıf Sayıları ve Oranları</i>	60
Tablo 9 <i>İki Düzeye Ortak Değişkenin eklendiği 2 PL CB2-C2 KMTK Modeli Öğrenci Düzeyi için Nihai Modelin Madde Güçlük İndeksleri</i>	61
Tablo 10 <i>İki Ortak Değişkenli 2 PL CB2-C2 KMTK Modeli Ülke Düzeyi için Nihai Modelin Madde Güçlük İndeksleri</i>	62
Tablo 11 <i>İki Ortak Değişkenli 2 PL CB2-C2 ÇBÇDKMTK Modeli Örtük Sınıflara Ait Puan Ortalamaları ve Standart Sapmaları.....</i>	63
Tablo 12 <i>İki Ortak Değişkenli 2 PL CB2-C2 ÇBÇDKMTK Modeli Ülke Düzeyi Örtük Sınıflarda Yer Alan Ülkeler</i>	64
Tablo 13 <i>Gözlenen Değişken Olan Cinsiyete Ait Sayılar, Oranlar, Puan Ortalamaları ve Standart Sapmaları.....</i>	65
Tablo 14 <i>Cinsiyet Değişkenine Göre Yapılan DMF Analizi Sonuçları</i>	66
Tablo 15 <i>Evdeki Kitap Sayısı Değişkenine Göre Yapılan DMF Analizi Sonuçları</i>	67
Tablo 16 <i>İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 Öğrenci Düzeyinde KMTK Modelinde ÖÜ1'e Ait DMF Analizi Sonuçları</i>	69
Tablo 17 <i>İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 Öğrenci Düzeyinde KMTK Modelinde ÖÜ2'ye Ait DMF Analizi Sonuçları</i>	70
Tablo 18 <i>İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 Ülke Düzeyinde KMTK Modeline Ait DMF Analizi Sonuçları</i>	71
Tablo 19 <i>Veri Setine En Uygun Modeller Kullanılarak Yapılan Öğrenci Düzeyindeki DMF Analizi Sonuçları</i>	72

Tablo 20 <i>Veri Setine En Uygun Modeller Kullanılarak Yapılan Ülke Düzeyindeki DMF Analizi Sonuçları</i>	75
Tablo 21 <i>Veri Setine En Uygun Modeller için Öğrenci Düzeyinde DMF'li Madde Sayısı, DMF'li Maddeler ve Etki Büyüklüğü Düzeyleri</i>	77
Tablo 22 <i>Veri Setine En Uygun İki Düzeyli Modeller için Ülke Düzeyinde DMF'li Madde Sayısı, DMF'li Maddeler ve Etki Büyüklüğü Düzeyleri</i>	79
Tablo 23 <i>Veri Setine En Uygun Modellere Ait Bilgi Ölçütü İndeksleri</i>	81
Tablo 24 <i>Seçilen İki Düzeyli Modellerden Oluşturulan Ülkelere Ait Sınıflar</i>	83
Tablo 25 <i>Tek Boyutlu CB2-C2 2 PL Modele ve İki Düzeye Ortak Değişkenin Eklendiği Modele Ait Madde Güçlük İndeksleri</i>	84
Tablo 26 <i>Örtük Sınıflara Ait Birey Sayıları, Ortalamalar ve Korelasyon Değerleri ile EKS Ortalamaları</i>	85
Tablo 27 <i>Ülke Düzeyi Örtük Sınıflardaki Öğrencilerin Örtük Sınıflardaki Sayıları</i> 86	
Tablo 28 <i>Öğrenci düzeyi B ve C Düzeyi DMF'li Maddelere Ait Test Bölümleri Konu Alanları ve Bilişsel Alanlar</i>	88
Tablo 29 <i>Ülke düzeyi B ve C Düzeyi DMF'li Maddelere Ait Test Bölümleri Konu Alanları ve Bilişsel Alanlar</i>	89

Şekiller Dizini

Şekil 1 Çok Düzeyli Yapının Analiz Kapsamına Alındığı Karma Model Şeması ..	23
Şekil 2 ÇDKMTK Modelinde Grup İçi ve Gruplar arası Yapıyı Gösteren Örnek Yapı	24
Şekil 3 Veri Setini Oluşturan 20 Madde İçin Yamaç Birikinti Grafiği.....	43
Şekil 4 20 Maddelik Veri Setinde Boyutluluk Analizi Kapsamında Paralel Analiz Sonuç Görseli.....	44
Şekil 5 10 Maddelik Matematik Veri Seti İçin Yamaç Birikinti Grafiği	45
Şekil 6 10 Maddelik Matematik Veri Setinde Boyutluluk Analizi Kapsamındaki Paralel Analiz Sonuç Görseli.....	46
Şekil 7 10 Maddelik Fen Bilimleri Veri Seti İçin Yamaç Birikinti Grafiği.....	47
Şekil 8 10 Maddelik Fen Bilimleri Veri Setinde Boyutluluk Analizi Kapsamında Paralel Analiz Sonuç Görseli.....	48
Şekil 9 20 Maddelik Veri Setine Ait Kayıp Veri Bilgileri	55

Simgeler ve Kısaltmalar Dizini

AERA: American Educational Research Association

AFA: Açımlayıcı Faktör Analizi

AIC: Akaike Bilgi Kriteri

APA: American Psychological Association

BIC: Bayesian Bilgi Kriteri

ÇBÇDKMTK: Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramı

ÇDKMTK: Çok Düzeyli Karma Madde Tepki Kuramı

DFA: Doğrulayıcı Faktör Analizi

DMF: Değişen Madde Fonksiyonu

eTIMSS: elektronik Uluslararası Matematik ve Fen Eğilimleri Araştırması

ILSA: Uluslararası Geniş Ölçekli Değerlendirmeler

KİK: Küme İçi Korelasyon

KMO: Kaiser-Meyer-Olkin

KMTK: Karma Madde Tepki Kuramı

KRM: Karma Rasch model

KTK: Klasik Test Kuramı

LR: Lojistik Regresyon

MCMC: Markov Zinciri Monte Carlo

MH: Mantel-Haenszel

MLE: En Yüksek Olabilirlik Kestirimi

MTK: Madde Tepki Kuramı

NCME: National Council on Measurement in Education

odÇBÇDKMTK: Ortak Değişkenli Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramı

ÖSA: Örtük Sınıf Analizi

RMSE: Ortalama Kareler Hata Karekökü

SABIC: Örneklem Büyüklüğü Ayarlı Bayesian Bilgi Kriteri

SE: Standart Hata

TIMSS: Uluslararası Matematik ve Fen Eğilimleri Araştırması

1PL: 1-Parametrelili Lojistik

2PL: 2-Parametrelili Lojistik

3PL: 3-Parametrelili Lojistik

Bölüm 1

Giriş

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

Problem Durumu

Bireylerin diploma ve terfi alma ile işe ve okula seçimi gibi çeşitli kararları almak üzere düzeyi bilinmek istenen yetenek, ilgi, tutum, beceri vb. bilişsel yapıların gerçeğe olabildiğince yakın ölçülmesinde en kullanışlı araçlardan biri testlerdir. Standartlar'da (American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME), 2014) test, belirli bir alandaki incelemeye tabi tutulmuş davranış örnekleminde elde edilen ve ardından standartlaştırılmış bir süreç kullanılarak değerlendirilen ve puanlanan bir araç veya yöntem olarak tanımlanır. Günümüzde testler sınıf içi ölçmelerden ülkeler arası eğitim araştırmalarına kadar geniş bir yelpazede kullanılmaktadır. Test sonuçlarına dayanarak birçok yargıya varılmakta ve bireyler, kurumlar ya da ülkeler arası karşılaştırmalar yapılmaktadır. Testten elde edilen puanlar ile verilen kararların isabetli, tutarlı, hatasız ve adil olması gerekir. Bu sebeple testin öncelikli olarak güvenilir olması ve test puanlarına ait geçerlik kanıtlarının elde edilmesi gerekir.

Tanımlı ve kapsamı tarihsel süreçte sürekli değişen geçerlik son 40 yıllık dönemde Standartlar (AERA ve diğerleri, 1985, 1999, 2014) ve Kane (2006) tarafından ortaya atılan kanıta dayalı yaklaşım ile ele alınmıştır. Buna göre geçerlik testin bir özelliği değildir ve test puanlarından yapılan çıkarımlarla ilgilidir. Geçerlik için Standartlar (AERA ve diğerleri, 2014) testin önerilen kullanımı için test puanlarının yorumlarının kanıt ve teorilerle destekleme derecesi olduğunu ifade eder. Dolayısıyla geçerlik test değerlendirme ve test geliştirmenin en önemli temelidir. Geçerlik süreci önerilen puan yorumlarında sağlam bir bilimsel temel

için gerekli kanıtları toplamayı içerir. Geçerlik testin kendisi için değil, değerlendirmede önerilen kullanımlar için test puanlarının yorumlarıdır.

Geçerlik, test puanlarının amaçlanan yorumu ve bunların önerilen kullanımla ilgisi lehine ve aleyhine argümanlar oluşturma ve değerlendirme süreci olarak ele alındığında beş çeşit kanıt kaynağından söz edilmektedir. Bunlar test kapsamına dayalı kanıt, yanıtlama süreçlerine dayalı kanıt, iç yapıya dayalı kanıt, diğer değişkenlerle ilişkilere dayalı kanıt ve testin sonuçlarına dayalı kanıttır (AERA ve diğerleri.,2014). Bu kanıt türlerinden araştırma kapsamında ilgilenilen iç yapıya dayalı kanıtta testin iç yapısının analizleri ile test maddesi ve test bileşenleri arasındaki ilişkilerin önerilen test puanlarının dayandığı yapıya ne ölçüde uyduğunu gösterilebilir. Bu kanıtlar için, uygulama öncesi duyarlılık analizleri (alan, ölçme ve dil uzmanlarının yapacağı analizler), boyutluluk analizleri, değişen madde fonksiyonu (DMF) analizleri, ölçme, yapı ve parametre değişmezliği analizleri, iç tutarlığa dayalı güvenilirlik analizleri yapılabilir. Ayrıca madde tepki kuramı (MTK) kapsamında model-veri uyumu analizleri, MTK artık analizleri ve duruma göre çok boyutlu ölçekleme, kümeleme analizi, örtük sınıf analizi, bilişsel tanı modelleri analizi yapılabilir. Bunun yanında test bilgisi ve ölçmenin standart hatasından faydalanılabilir.

DMF analizleri geçerliğin testin iç yapısına dayalı kanıt toplanmasında kullanılır ve testin adilliğinin sağlanmasında katkı sağlar. Fukuhara ve Kamata (2011), testin adilliğinin, test puanlarının geçerliğini belirlerken standartlaştırılmış testlerde önemli bir husus olduğunu ve bir testin hedeflenen bir popülasyonun alt grupları için adil olup olmadığını araştırmak için, yaygın olarak DMF analizi yapıldığını ifade etmişlerdir. Bir madde, ilgilenilen örtük bir özellik (örneğin yetenek) üzerindeki performans düzeylerindeki farklılıklar kontrol edildikten sonra bir referans gruba kıyasla bir odak grupta farklı işlev gördüğünde DMF sergilediği söylenir (Clauser & Mazor, 1998; Holland & Wainer, 1993; Potenza & Dorans, 1995; Scheuneman, 1979). Bir testte farklı işleyen bir maddenin varlığı, o maddenin testin geri kalanıyla ilgisi olmayan bazı rahatsız edici (nuisance) boyut ya da boyutları ölçtüğünü gösterir (Ackerman, 1992). DMF belirlemeye yönelik yaklaşımların çoğu, cinsiyet ve etnik

köken gibi özelliklerle tanımlanan sınava giren gruplar arasındaki farklılıklara odaklanarak testleri daha adil hale getirmek üzere tasarlanmıştır. İncelenen gruplar ne olursa olsun, DMF test geçerliği için ciddi bir tehdit olarak kabul edilir çünkü bir grubun bir madde üzerinde diğer bir gruba kıyasla haksız bir avantaja sahip olduğu anlamına gelir. DMF'nin test adaleti ile ilgisi göz önüne alındığında, ölçme araştırmacıları DMF'yi araştırmak için çok sayıda yöntem geliştirmiştir (De Mars & Lau, 2011).

DMF, test kalitesinin önemli bir göstergesidir ve DMF belirleme yöntemleri parametrik olabilir veya olmayabilir. Mantel-Haenszel (MH) (Mantel & Haenszel, 1959; Holland & Thayer, 1988) ve SIBTEST (Shealy & Stout, 1993), klasik test kuramı (KTK) çerçevesinde kullanılan parametrik olmayan DMF belirleme süreçlerinin örnekleridir. Lord'un (1980) ki-kare yöntemi, Raju'nun (1988, 1990) alan ölçümleri ve olasılık oranı testi (Thissen ve diğerleri, 1988) dahil olmak üzere çeşitli MTK tabanlı yöntemler de mevcuttur. Bu yöntemler, test uygulamasında DMF'yi belirlemede çok faydalıdır, ancak Roussos ve Stout'a göre (1996) DMF'nin nedenlerini anlamada çok az ilerleme kaydedilmiştir (aktaran Yao & Li, 2015). Gözlenen gruplar DMF analizinde, DMF maddeleri istatistiksel anlamlılıkları ve etki büyüklükleri için istatistiksel analizler kullanılarak tespit edilir. DMF belirleme yöntemleri genel olarak benzer sonuçlar verse de yöntemlerde farklı algoritma ve eşleme ölçütleri kullanıldığı için hepsinin tam bir uyum içerisinde olduğu söylenemez. Ayrıca, DMF'li maddeleri belirlemek için kullanılan kesme noktaları da farklıdır (Bakan Kalaycıoğlu & Berberoğlu, 2011; Doğan & Öğretmen, 2008; Gök ve diğerleri, 2010; Uyar, 2015). Bu sebeple, araştırmacıların ve test geliştiricilerin DMF analizi için birden fazla yöntem kullanmaları önerilmektedir (Hambleton, 2006).

Gözlenen gruplara ait DMF analizi için kabul edilen paradigma nispeten güçlü bir varsayıma dayanmaktadır. Bu varsayım bir gözlenen grubun tüm üyeleri aynı stratejileri veya problem çözme tekniklerini kullanır ve bir maddenin içeriğiyle ilgili olarak benzer geçmişlere sahip olmasıdır. Bu nedenle, bir maddeye doğru yanıt verme olasılığı sınav katılımcısının yeteneğine bağlı olsa da bir gözlenen grubun tüm üyeleri o maddeyi aynı

şekilde görür (Samuelsen, 2008). Gözlenen grup değişkenlerini tespit etmek için kullanılan DMF belirleme yöntemlerinin uzun bir geçmişi vardır. Ancak, DMF'ye yol açan deneyimler veya bilişsel süreçler gözlenen gruplarla mükemmel bir korelasyon göstermiyorsa, DMF'nin altında yatan örtük grupları ortaya çıkarmak daha bilgilendirici olacaktır (De Mars & Lau, 2011). Cohen ve Bolt (2005), DMF'ye yönelik gözlenen gruplar yaklaşımının sorunlu olduğuna dair ampirik kanıtlar sunmuştur. DMF için bir olabilirlik oranı testi kullanarak madde parametrelerini tahmin etmişler ve 32 test maddesinden beşinde cinsiyete göre DMF olduğunu bulmuşlardır. Elde edilen madde parametreleri daha sonra üç parametrelili lojistik (3-PL) karma modelde kullanılmış ve katılımcıların cinsiyetlerinden bağımsız olarak örtük gruplarda yer almaları sağlanmıştır. Daha sonra karma oranları, örtük sınıf yetenek dağılımları ve bireysel sınav katılımcıları için sınıf üyelikleri tahmin edilmiştir. Sonuçlar, popülasyondaki erkek ve kadın yüzdelerine yakın karma oranlar verse de daha yakından incelendiğinde "erkek ve kadınların cinsiyetlerinin karşısındaki örtük sınıfa sınıflandırılmasında önemli ölçüde geçiş olduğu" görülmüştür. Yani bir gözlenen gruba ait DMF'li maddeler o grubun tüm üyelerine karşı farklı işlev görmemiştir.

Hu ve Dorans (1989), marjinal dağılımlara odaklanan DMF analizlerinde bir kusur olduğunu göstermiş ve DMF analizlerinde etkileşimlerin de dikkate alınması gerektiğini ortaya koymuştur. Kadınların lehine olan bir maddenin çıkarılmasının, kadınlar için biraz daha düşük ve erkekler için biraz daha yüksek puanlarla sonuçlandığını bulmuşlardır. Bu beklenen durumun yanında araştırmacılar hem Hispaniklerin hem de Asya kökenli Amerikalıların puanlarının erkeklerin puanlarından daha fazla yükselmiş olduğunu; yani bu gruplardaki kadınların maddenin kaldırılmasıyla aslında bir avantaj elde etmiş olduklarını belirlemişlerdir. Sıklıkla incelenen gözlenen grupların gerçekten ilgilenilen gruplar olmaması da mümkündür. Sosyoekonomik durum, okul türü (ev, özel, devlet) ve okul büyüklüğü gibi diğer gözlenen gruplama değişkenleri DMF açısından potansiyel olarak ilginç ve bilgilendiricidir. Bunlar aynı zamanda daha homojen alt popülasyonları da temsil edebilir (Samuelsen, 2008).

Bireylerin niteliksel farklılıkları açısından alt gruplara ayrılmasında kullanılabilecek bazı istatistiksel yöntemler bulunmaktadır ve bu yöntemlerden birisi de örtük sınıf analizidir (ÖSA). ÖSA örtük alt grupları tanımlayan belli tepki kombinasyonlarının belirlenmesine imkân sağlamaktadır ve böylece birbirine benzeyen tepki örüntülerine sahip olan bireyler aynı sınıflarda gruplanarak benzer özelliklere sahip alt sınıflar elde edilebilmektedir (Hagenaars & McCutcheon, 2002). ÖSA'nın temel amacı, gözlenen değişkenleri kullanarak bireyleri sınıflamak ve sınıfları en iyi şekilde ayıran gözlenen değişkenleri veya maddeleri belirlemektir (Akbaş & Kahraman, 2019).

Örtük sınıfların maddelere veya madde türlerine nasıl tepki verdikleri önemlidir. Örtük sınıfların demografik bilgiler gibi kişi özellikleriyle ne ölçüde ilişkilendirilebileceği araştırılarak karakterize edilebilir. Bu nedenle, kişi özellikleriyle örtük bir DMF analizi yapmak, örtük sınıflar arasındaki DMF maddelerini belirlemenin yanı sıra DMF'yi açıklamak için de yararlıdır (Cho ve diğerleri, 2015). DMF gözlenen gruplara dayalı olarak belirlendiğinde, ilgilenilen grup üyeliği DMF analizi için cinsiyet ve etnik köken grupları gibi bilinen bir grup olarak önsel (a priori) olarak belirlenir. İlgilenilen grup üyeliğini önsel olarak belirlemek yerine, örtük sınıflarla bilinmeyen homojen bir alt grup, karma modelleme kullanılarak verilerden bulunabilir (McLachlan & Peel, 2000). Örtük DMF analizinde, bir örtük grubun homojen olduğu varsayılır çünkü karma madde tepki modellerinden aynı örtük gruba sınıflandırılan kişiler benzer madde tepki davranışları gösterir. Gözlenen grup üyeliği ile örtük grup üyeliği (verilerden tespit edilen) arasında büyük bir örtüşme olduğunda gözlenen DMF ve örtük DMF analizlerinin sonuçları benzer olacaktır (Cho ve diğerleri, 2015).

Örtük sınıf yaklaşımının kullanılması, farklılığın neden ortaya çıktığını iki şekilde açıklamaya yardımcı olabilir. İlki, farklı işleyen tüm maddeler tanımlanacağından, DMF'nin büyüklüğüne ilişkin daha doğru bir gösterge ile, araştırmacı için daha fazla bilgi mevcut olmasıdır. Mevcut stratejiler, farklı işleyen maddelerin yalnızca bir alt kümesini belirleyebilir ve DMF'in büyüklüğünü olduğundan az göstererek farklı işleyişin nedenini izole etmeyi daha

zor hale getirebilir. İkincisi, örtük bir strateji, araştırmacıların gözlenen grup üyeliği yerine örtük sınıfın yordayıcıları olarak ortak değişkenleri dahil etmelerine olanak tanıyabilir. Örneğin Bandalos vd. (2006), Florida Kapsamlı Değerlendirme Testinin 2003 yılı dokuzuncu sınıf Matematik Testinden elde edilen verilerin incelenmesinde, örtük sınıflar arasındaki önemli farklılıkları anlamının bir yolu olarak hem kategorik hem de sürekli ortak değişkenlerin dahil edilmesini göstermiştir (Samuelsen, 2008).

Örtük DMF analizinde, gözlenen DMF analizine kıyasla üç adım daha atılması gerekmektedir. Birinci adım, çeşitli aday modeller arasından en uygun ölçme modelinin (yani bir karma madde tepki modeli) bulunması gereken bir model seçim aşamasıdır. Aday modeller, sınıf içi modeli sabit tutarak, örtük sınıf sayısı (yani kategorik örtük değişken sayısı) açısından farklılık gösterir. İkinci adım, seçilen ölçüm modeli göz önüne alındığında, her bir kişi için bir örtük sınıf üyeliği elde etmektir. Üçüncü adım, örtük sınıfları karakterize edebilecek veya yorumlayabilecek DMF maddelerini belirlemektir. Araştırmacılar, (örneğin, Cho ve diğerleri, 2013; Cho ve diğerleri, 2010; Finch & Finch, 2013; Leite & Cooper, 2010; Maij-de Meij ve diğerleri, 2008, 2010) örtük sınıf üyeliğini belirledikten sonra örtük DMF'yi belirlemek için gözlenen DMF belirlemesine ait analizleri kullanmışlardır (Cho ve diğerleri, 2015).

Örtük değişken yönteminin sosyal bilimlerde kullanılan başka bir üyesi ise MTK modelleridir. Ölçülen bir özelliğin ya da yeteneğin test maddesine verilen tepki ile olan ilişkisinin matematiksel bir model ile tanımlandığı modern psikometrik teoridir. MTK modelinde doğru yanıt olasılıkları tahmin edilir, madde parametreleri deneklerin yetenek parametreleri ile birlikte örtük süreklilik içine yerleştirilir ve niceliklidir. ÖSA modellerinde ise örtük değişken kategorik ve niteliklidir (Dai, 2009; De Ayala & Santiago, 2017; Li, 2014). Standart bir MTK modelinde madde parametrelerinin her sınav katılımcısı için aynı olduğu varsayılır. Bir kişinin örtük yetenek parametresi, genel olarak, testi alan kişinin gözlemlenemeyen bir özelliği olarak kavramsallaştırılır. Bir kişinin bu varsayılan özelliği MTK modelleri aracılığıyla ölçülür. ÖSA'da gözlemlenmeyen örtük sınıf üyeliği koşullu

bağımsızlık ile tanımlanır. Yani, bireylerin örtük sınıf üyeliğine koşullandıktan sonra, bireylerin gözlenen yanıtları birbirinden bağımsızdır. Gözlenen verilere dayanarak, ÖSA'nın amacı denekleri farklı örtük sınıflara sınıflandırmaktır. Bu, MTK modelinin hem maddeler hem de kişiler için nicel parametreleri tahmin etmeyi amaçlaması bakımından MTK'dan farklıdır. Karma MTK (KMTK) modelleri ise MTK modelleri ve ÖSA modellerinin birleşimidir ve iki yaklaşımının kombinasyonu olarak, KMTK modelleri katılımcıların hem sürekli örtük özelliklerini hem de örtük sınıf üyeliklerini aynı anda tahmin eder (Dai, 2009; Li, 2014).

Son yıllarda, KMTK modelleme yaklaşımı, örtük DMF'nin belirlenmesi (Cho, 2007; Choi ve diğerleri, 2015; Cohen & Bolt, 2005; Dai, 2009; De Ayala ve diğerleri, 2002; De Boeck ve diğerleri, 2011; De Jong & Steenkamp, 2010; Samuelsen, 2005), test hızılığının tespiti (Bolt ve diğerleri, 2002; Boughton & Yamamoto, 2007; Meyer, 2010) ve alternatif bilişsel özelliklere veya stratejilere (Jeon, 2019; Lee ve diğerleri, 2017; Mislevy, & Verhelst 1990; Rijmen & De Boeck, 2005) sahip test katılımcılarının sınıflandırılması dahil olmak üzere test geliştirmede çeşitli önemli psikometrik sorunların üstesinden gelmek için uygulanmıştır. Ayrıca, bu yaklaşım birçok pratik duruma da uygulanmıştır. Örneğin, psikopatolojik testlerde araştırmacılar ve klinisyenler için denekleri en olası davranış bozukluğu türlerine atamak büyük bir öneme sahiptir ve bu tür senaryolarda, KMTK modelleri bir müdahale programı uygulamak üzere tanılama amacıyla kullanılabilir (Li, 2014).

Rost (1990)'a göre karma modeller, temel modellerin varsayımlarını kaybetmeden niteliksel bireysel farklılıkları dikkate almayı vaat eden bir yoldur. KMTK modeli madde tepki örüntüleri açısından homojen olan sınav katılımcılarının örtük sınıflarını tanımlamak için kullanılır. KMTK'da her bir örtük sınıfın üyeleri yetenek bakımından farklılık gösterir ve cevap stratejileri sınıflar arasında farklılık gösterecek şekilde oluşturulur. KMTK modeli kullanılarak aynı anda hem kişiler arasındaki fark ortaya çıkar hem de aynı madde için onların yetenekleri hesaplanabilir (Cho & Cohen, 2010). KMTK modeli, Karma Rasch modele (KRM), 2-PL ve 3-PL modele göre geliştirilebilmektedir (Samuelsen, 2008).

KMTK modelleri, test geliştirme ve DMF belirleme de dahil olmak üzere birçok önemli psikometrik sorunun üstesinden gelmek için kıymetli bir metodoloji olarak hizmet eder (Li, 2014). Ancak bu modele ait önemli bir sınırlılık vardır, o da eğitim testi verilerinin çoğunda öğrenci düzeyinin ötesinde mevcut olan temel çok düzeyli yapıyı esasen göz ardı etmesidir (Cho & Cohen, 2010). Davranışsal ve sosyal araştırmalarda hiyerarşik yapıların yaygınlığına rağmen, geçmiş çalışmalar genellikle veri analizinde bunları yeterince ele almamıştır. Bu ihmal, büyük ölçüde, iç içe geçmiş yapılara sahip doğrusal modellerin kestirimine yönelik geleneksel istatistiksel tekniklerdeki sınırlamaları yansıtmaktadır. Sosyal bilimlerdeki araştırmalarda bu sınırlamalar, kümeleme yanlılığı, yanlış tahmin edilen kesinlik ve "analiz birimi" sorununa ilişkin endişeler yaratmıştır (Raudenbush & Bryk, 2002). Bu sebeple KMTK modeline düzeylilik unsuru da dahil edilerek çok düzeyli karma madde tepki kuramı (ÇDKMTK) modeli elde edilmiştir.

Çok düzeyli veri yapısı göz ardı edildiğinde ve küme içi korelasyon (KİK, Raudenbush & Bryk, 2002) değerleri sıfır olmadığında ortaya çıkan analizler muhtemelen madde parametrelerinin ve bunlarla ilişkili standart hataların yanlış kestirimlerini verecek ve bu da hatalı DMF belirleme sonuçlarına yol açacaktır (Finch & Finch, 2013). Bunun yanında çok düzeyli verilere tek düzeyli modellerin uygulanması hem istatistiksel hem de kavramsal problemlere yol açmaktadır (De Leeuw & Kreft, 1986, Hox, 2002 aktaran Atar, 2010). Cho ve Cohen (2010) bir simülasyon çalışmasına ve gerçek bir veri setinin analizine dayanarak ÇDKMTK modelinin hem DMF'nin varlığını belirlemede hem de sınava giren ve okul düzeyindeki ortak değişkenlerin modele dahil edilmesi yoluyla DMF'nin nedenlerini araştırmada etkili bir araç olduğunu göstermiştir. ÇDKMTK modelinin karma kısmı, öğrenci ve okul düzeyinde olmak üzere iki düzeyde örtük sınıfları tanımlamak için kullanılır. Öğrenci düzeyinde, örtük sınıflar belirli tepki stratejilerini kullanmada homojen olan kişilerden oluşur (Mislevy & Verhelst, 1990; Rost, 1990). Okul düzeyinde, örtük sınıflar benzer özellikleri paylaşan okullardan oluşur (Vermunt & Magidson, 2005). Çok düzeyli karma modeller birçok eğitim verisinin özelliği olan doğal yuvalanmayı açıklar ve aynı zamanda, tepki

modellerine göre homojen grubun (yani, örtük sınıfın) elde edilmesiyle örtük sınıflar arasındaki farklılıkların neden ortaya çıkmış olabileceğini açıklamanın bir yolunu sağlar. Bu tür bilgiler (ÇDKMTK modelinde verilen bilgiler), standart DMF testinden daha büyük bir açıklayıcı güce sahiptir. Bu tür bilgilerin sağlanması, maddelerin farklı şekilde çalışmasının nedenlerini kavramsallaştırmaya yardımcı olabilir (Cho, 2007).

ÇDKMTK modellerinin önemli bir sınırlaması ise, grup üyeliği hakkında modele dahil edilen gözlenen yordayıcılar tarafından verilenin ötesinde bilgi sağlamamasıdır (Cho & Cohen, 2010). Bunun yanında gerçek dünyadaki birçok değerlendirme durumu, sınava girenlerin aynı anda birden fazla yapı üzerinde değerlendirilmesini içermektedir. Bu tür senaryolar, örneğin bir matematik sınavında hesaplama, problem çözme, geometri gibi matematik becerisinin birden fazla yönünün değerlendirilmesinde olduğu gibi, birkaç alt becerinin aynı anda değerlendirildiği konu alanı testlerinin uygulanmasını içerebilir (Finch & Finch, 2013). Çoğu zaman psikometrik analizler doğası gereği yalnızca niceldir ve asıl test içeriğini ve sınava girenlerin testteki maddeleri cevaplamak için ihtiyaç duyduğu bilişsel süreçleri göz ardı ederek öncelikle sınava girenlerin doğru yanıtlarına odaklanır. Birçok eğitimsel ve psikolojik test doğası gereği çok boyutludur, yani bu testler iki veya daha fazla boyutu veya yapıyı ölçer ve sınava girenlerin aynı anda birden fazla yapı üzerinde değerlendirilmesini içerir. Bu tür senaryolar birkaç alt becerinin aynı anda değerlendirildiği konu alanı testlerinin yönetimini içerebilir. Başka bir bağlamda, sınava girenlere aynı anda dil ve matematik gibi iki ayrı yapının testleri verilebilir. Bu gibi durumlarda, örtük özelliğin tek boyutlu olduğuna dair standart MTK varsayımı muhtemelen geçerli olmayacak ve bu da hem madde hem de kişi parametrelerinin tahmininde potansiyel zorluklara yol açacaktır (Ackerman, 1994; Ackerman ve diğerleri, 2003; Reckase, 1985).

Çok boyutluluk eğitimsel ve psikolojik testlerde yaygındır. Bu testlerle ölçülen tüm yetenekler kasıtlıdır ve açımlayıcı faktör analizi kullanan araştırmalar genellikle verilerde altta yatan çok boyutlu bir yapı bildirir. Çok boyutluluk verileri iki türe ayrılabilir: maddeler arası (between) ve madde içi (within) çok boyutluluk. Farklı örtük yetenekleri ölçen ve test

maddelerinin tek bir yapıyı yansıttığı birkaç alt testi içeren bir test, basit bir test yapısını temsil eder ve maddeler arası çok boyutlu verileri verir. Terra Nova Çoklu Değerlendirme Testi ve Minnesota Çok Yönlü Kişilik Envanteri, çok boyutlu testler için iki temsili örnektir (Adams ve diğerleri, 1997; Mazor ve diğerleri, 1995; Wang ve diğerleri, 1997 aktaran Chen & Jin, 2018). Bu gibi durumlarda, örtük özelliğin tek boyutluluğuna ilişkin standart MTK varsayımı muhtemelen geçerli olmayacak ve hem madde hem de birey parametrelerinin tahmininde potansiyel zorluklara yol açacaktır (Ackerman, 1994; Reckase, 1985). Bu duruma ek olarak, veriler çok boyutlu olduğunda, standart DMF belirleme yöntemleri de sınava giren kişinin yeteneklerini tek bir örtük özelliğe göre koşullandırdığı göz önüne alındığında etkili olmayabilir (örneğin, MH testi, lojistik regresyon (LR), SIBTEST, olabilirlik oran testi). Bu gibi durumlarda hem doğru parametre tahmini sağladığı hem de ayrı yapılar hakkında ek bilgi sağlayabileceği için çok boyutlu madde tepki kuramı modeli önerilir.

Yukarıda belirtilen avantajlı ve dezavantajlı durumlar göz önüne alındığında DMF değerlendirmesi için hem düzey hem de boyutluluğu kapsayan karma bir model olan çok boyutlu çok düzeyli karma madde tepki kuramı (ÇBÇDKMTK) modelinin faydalı olacağını göstermektedir. ÇBÇDKMTK modeli ÇDKMTK modeli ile hemen hemen aynı şekilde kullanılmak üzere tasarlanmıştır ancak birden fazla örtük özelliğin aynı anda değerlendirildiği duruma uygun olarak değiştirilmiştir. Ayrıca çok boyutlu analizler ile madde güçlükleri tek boyutluluk varsayımı altında değil, doğru çok boyutlu model altında tahmin edilir. ÇBÇDKMTK modeli örtük sınıf üyeliğinin yalnızca tek bir örtük özelliği değerlendiren maddeler yerine birden fazla örtük özelliği ölçen bir dizi madde tarafından belirlenmesidir. Bu nedenle, sınıflar daha ayrıntılı bir dizi özelliği yansıtacak ve böylece potansiyel olarak, her seferinde yalnızca bir örtük yapının kullanılması durumunda ortaya çıkacak olandan daha farklı madde tepki örüntülerinin daha gerçekçi bir resmini ortaya çıkaracaktır. Dolayısıyla, örneğin, yalnızca tek bir yapıya dayalı olarak popülasyonun alt gruplarını anlamak yerine, ÇBÇDKMTK modeli aynı anda iki veya daha fazla örtük özelliğe göre sınava giren alt gruplara ilişkin bilgi sağlama potansiyeline sahiptir (Finch & Finch, 2013).

Son olarak, KMTK, ÇDKMTK ve ÇBÇDKMTK modellerinde verilerin her seviyesindeki ortak değişkenler (kovaryantlar) DMF'nin doğasını araştırmak için kullanılabilir. Modeller cinsiyet, etnik köken, sosyoekonomik durum ya da derse yönelik tutum gibi kişi düzeyindeki ortak değişkenleri dahil etmek için kişi düzeyinde genişletilebilirken varsa ikinci düzeyde model, her düzeydeki örtük sınıfların bileşimini tanımlamak için sosyoekonomik durum ve kentsel/kırsal konum gibi ikinci düzeydeki ortak değişkenleri içerecek şekilde genişletilebilir (Cho, 2007). Bir KMTK modeline bir ortak değişkenin dahil edilmesinin, örtük sınıfların tespiti ve ayrıca örtük sınıflar arasındaki farklılıkların karakterize edilmesi için yararlı olduğu gösterilmiştir (örneğin, Bilir, 2009; Choi ve diğerleri, 2015; Dai, 2013; Smit ve diğerleri, 1999). Ortak değişkenlerin sınıf üyeliği ve madde parametreleri üzerindeki etkileri sınıfa özgüdür. Ortak değişkenlerin madde parametreleri üzerindeki önemli etkisi, madde parametre tahminleri için daha küçük standart hatalar vermesidir. Önemli bir etkileşim terimi de örtük sınıf üyeliğini tahmin etme ve açıklama katsayıları üzerinde etkili olmuştur (Karadavut ve diğerleri, 2019). Belirli maddeler için DMF bulunması halinde, bu ortak değişkenlerin dahil edilmesi olası kaynaklarının açıklanmasına yardımcı olabilir (Finch & Finch, 2013). Bu sebeple ÇBÇDKMTK modeline ortak değişkenler ekleyerek DMF'nin doğasının ve olası sebeplerinin anlaşılmasının kolaylaşacağı düşünülmektedir. Bu sebeple araştırma kapsamında ortak değişkenler içeren çok boyutlu çok düzeyli karma madde tepki kuramı (odÇBÇDKMTK) modeli yer almıştır.

DMF belirlenmesine yönelik analizlerde gözlenen gruplara göre yapılan analizlerin sınıfların homojen dağılmaması sebebiyle doğru sonuçlar vermediği çeşitli çalışmalarda (Cho ve diğerleri, 2015; Cohen & Bolt 2005; De Mars & Lau, 2011; McLachlan & Peel, 2000) ifade edilmiştir. Karma modeller kullanılarak yapılan DMF belirleme çalışmalarında testin boyutluluğu ve düzeyliliği göz ardı edildiği için Tip I hata oranının şişmesi ve hatalı DMF çıkarımları yapılması durumlarının görüldüğü belirtilmiştir (Finch & Finch, 2013). Çok düzeyli karma model ile düzeylilik göz önüne alınmıştır ancak testin birden fazla boyut içermesi durumunda yine tam olarak doğru sonuçlar elde edilemeyebilir. Araştırma

kapsamında literatürde belirtilen gözlenen gruplara, karma modele, çok düzeyli karma modele ve çok boyutlu çok düzeyli karma modele göre DMF analizlerinin olumlu ve olumsuz yanlarının gerçek veri üzerinde incelemesi yapılmıştır. Elde edilen bulgular ile modellere göre en uygun örtük sınıf sayılarının, DMF'li maddelerin, DMF'li madde sayısındaki değişimin ve DMF'nin kaynaklarının bulunmasına dair çeşitli çıkarımlar yapılmıştır. Tek bir analize birden fazla yapıyı dahil ederek, örnekleme mevcut olan sınava giren birey tipleri hakkında daha fazla bilgi edinilebilir ve bu da eğitimcilere, farklı madde tepki kalıplarına sahip bireylerle nasıl daha iyi çalışabilecekleri konusunda fikir verebilir. Ayrıca ÇBÇDKMTK modelinin çok boyutlu doğasının, araştırmacılara ve eğitimcilere DMF'nin çok yönlü doğası hakkında, ölçeklerin ayrı ayrı analiz edilmesinden daha fazla bilgi sağlayacağı düşünülmektedir.

Araştırmanın Amacı ve Önemi

Standartlar'da (AERA ve diğerleri, 2014) DMF yöntemlerinin DMF belirlemede etkili ancak DMF kaynaklarının belirlenmesinde yetersiz olduğu vurgulanmıştır. Zumbo (2007) DMF'nin üç neslinden bahseder ve ilk nesilde yaygın kullanılan terimin madde yanlılığı olduğunu ve testlerde madde yanlılığını tespit etmek için genellikle iki grubu karşılaştıran yöntemlerin uygulandığını belirtir. İkinci nesilde ise madde yanlılığından DMF terimine geçiş yaşanmıştır ve DMF'nin kaynağına dair açıklamalar sağlamak yerine, öncelikle DMF içeren maddeleri belirlemek veya "işaretlemek" için gelişmiş istatistiksel yöntemler geliştirmeye odaklanılmıştır. Üçüncü nesilde ise DMF'nin kaynağının belirlenmek istenmesi, DMF çalışmasını teorileştirmenin erken bir işareti olarak görülür. Yani, DMF'nin üçüncü nesli en açık şekilde, DMF'nin test maddesinin ve/veya test durumunun, ilgilenilen temel yetenekle (ve dolayısıyla testin amacıyla) ilgili olmayan bazı özellikleri nedeniyle ortaya çıktığını düşünmek olarak tanımlanmaktadır. DMF'nin ilk iki nesline (çok boyutlu model de dahil olmak üzere) hakim olan olası DMF nedenlerine "test durumu" da eklenerek, DMF pratiği büyük ölçüde genişletilmekte ve test yapısının (ve dolayısıyla çok boyutluluğun) ötesindeki konulara odaklanılmakta, dolayısıyla çok boyutlu DMF modelinin ötesine geçilmektedir.

Pektaş (2018) bir teste ait geçerlik ve güvenirlik gibi psikometrik özellikleri olumlu yönde artırabilmek için testin orta ve yüksek düzeyde DMF gösteren maddelerden arındırılması gerektiğini belirtip araştırma sonuçlarına bağlı olarak özellikle karar almak için uygulanan testlerin geçerlik ve güvenirlik belirleme sürecinde DMF analizlerinin yapılmasını önermiştir. Buna göre üçüncü nesil DMF ile birlikte sadece basit bir analiz yardımıyla testin adil ve geçerli yapılacağını düşünmek yerine testin maddelerinden ya da test durumlarından kaynaklanmış olabilecek farklılaşmayı doğru tespit etmek ve farklılaşmanın sebebini bulabilmek önemli hale gelmiştir.

DMF belirlemede maddelerin farklı işlediği gözlenen gruplar yanlılığın gerçek kaynağı ile örtüşmeyebilir. Örtük bir DMF değişkeni içeren bir model altında DMF belirlenmesinin bu yanlılık kaynağına karşı daha hassas olması beklenmektedir. Simülatif bir çalışmada, örtük bir gruplama değişkeni içeren bir KMTK modelinin DMF'li maddeleri belirlemede yalnızca gözlenen değişkenleri kullanan DMF belirleme yöntemlerine göre daha iyi performans gösterdiği bulunmuştur (Maij-de Meij ve diğerleri, 2010). O halde DMF'nin sadece gözlenen bir kategorideki üyelikle ilişkili olduğunu varsaymak yerine önce verilerde madde tepki profilleri açısından farklılık gösteren grupları belirlenebilir ve ardından örneğin bir veya daha fazla madde için tek tip DMF'nin mevcut olup olmadığı incelenebilir. Ayrıca, bu örtük sınıflar, cinsiyet veya evde bulunan kitap sayısı gibi sınav katılımcılarıyla ilişkili gözlemlenebilir özellikler açısından karakterize edilebilir. Bu şekilde, Cohen ve Bolt'un (2005) yanı sıra diğer araştırmacılar da (örneğin, Cohen ve diğerleri, 2005; Samuelson, 2005) KMTK yaklaşımının DMF'nin öncüllerine ilişkin olarak, yalnızca gözlenen gruplarla ilişkili DMF'yi belirmeye dayanan yöntemlere kıyasla daha fazla içgörü sağlayabileceğini göstermiştir. Buna ek olarak, DMF değerlendirmesine yönelik bu yaklaşım, kendisi de bazı açılardan önyargılı olabilecek, önceden belirlenmiş sınav katılımcıları sıralamasına dayanmayan daha kapsamlı analizler sağlama potansiyeline sahiptir (Finch & Finch, 2013)

Araştırmanın amacı gözlenen değişkenlere göre oluşturulan gruplar ile KMTK, ÇDKMTK, ÇBÇDKMTK ve odÇBÇDKMTK modelleri ile elde edilen örtük sınıflardaki DMF'li

maddeleri belirleme ve DMF'nin olası kaynaklarını açıklayabilme performanslarını karşılaştırmaktır. Bu doğrultuda elektronik Uluslararası Matematik ve Fen Eğilimleri Araştırması (eTIMSS) 2019 çalışmasının birinci kitapçığında yer alan matematik ve fen bilimlerine ait 20 maddenin bulunduğu teste ait maddeler gözlenen değişkenlere ait gruplar için MH yöntemiyle, KMTK modellerine ait düzey, boyut ve ortak değişkenin dahil edilmesiyle elde edilen örtük gruplar için MH ve genelleştirilmiş Mantel-Haenszel (GMH) yöntemleriyle DMF analizi yapılmıştır. Araştırma neticesinde literatüre olabilecek kuramsal faydaların yanında toplumsal ve psikometrik açıdan da özellikle testlerin adilliğini sağlama ve geçerlik kanıtlarını artırma bakımından katkıda bulunacağı düşünülmektedir. Araştırmacılar açısından ise standart DMF uygulamaları yerine testin yapısına uygun DMF belirleme yönteminin seçiminin çalışmaların kalitesini artıracığının anlaşılacağı düşünülmektedir. Ayrıca DMF sürecinin sadece madde belirlemeden ibaret olmadığı ve DMF'nin kaynağını bulmanın önemli olduğu düşüncesinin özellikle ortak değişkenlerin katkısıyla öne çıkacağı beklenmektedir. Birden fazla yapıyı tek bir analize dahil ederek, örnekleme mevcut olan birey tipolojileri hakkında daha fazla bilgi edinilebileceğini ve bunun da eğitimcilere farklı madde tepki örüntülerine sahip bireylerle nasıl daha iyi çalışabileceklerine dair iç görü sağlayacağına inanılmaktadır.

Araştırma Problemi

Araştırmada gözlenen gruplara, Karma Madde Tepki Kuramına, Çok Düzeyli Karma Madde Tepki Kuramına, Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına ve ortak değişkenin dahil olduğu Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına göre yapılan DMF analizlerinde örtük sınıf sayıları, DMF'li maddeler ve sayıları ile madde kestirimleri farklılaşmakta mıdır? sorusuna cevap aranacaktır.

Alt Problemler

- 1) Cinsiyet ve evdeki kitap sayısı değişkenleri için Mantel Haenszel ve Genelleştirilmiş Mantel Haenszel yöntemlerine, Karma Madde Tepki Kuramına,

Çok Düzeyli Karma Madde Tepki Kuramına, Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına ve ortak değişkenin dahil olduğu Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına göre yapılan analizlerde DMF'li madde sayıları ve etki büyüklüğü düzeyleri farklılaşmakta mıdır?

- 2) Karma Madde Tepki Kuramına, Çok Düzeyli Karma Madde Tepki Kuramına ve Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına göre yapılan analizlerde 2PL ve 3PL modellere göre en uygun olan örtük sınıf modeli, DMF'li madde sayıları ve etki büyüklükleri farklılaşmakta mıdır?
- 3) Veriye uyarlanan en uygun modellere göre ortaya çıkan örtük sınıflarda cinsiyet, ülke ve madde parametrelerine ilişkin özelliklerin dağılımı ve bulgular farklılaşmakta mıdır?
- 4) Ortak değişkenlerin dahil olduğu Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına göre yapılan DMF analizlerinde DMF'nin kaynağını belirlemeye yönelik zengin içerikli çıkarımlar yapılabilmekte midir?

Sayıtlılar

Çalışma kapsamında kullanılan eTIMSS 2019 uygulamasına ait veriler ölçme ve değerlendirme kriterlerine uygun olarak toplanmıştır.

Sınırlılıklar

Çalışmada sadece ikili puanlanan test maddeleri kullanılmıştır.

Çalışmada sadece eTIMMS 2019 uygulamasına katılan (benchmark haricindeki) ülkelerdeki bireylerin birinci kitapçığa ait verileri kullanılmıştır.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Araştırmanın kuramsal temeli; DMF, DMF metotları ve DMF kaynağı, karma madde tepki kuramı modeli, çok düzeyli madde tepki kuramı modeli, çok boyutlu çok düzeyli karma madde tepki kuramı modeli ve ortak değişkenlerin eklendiği çok boyutlu çok düzeyli karma madde tepki kuramı modeli olmak üzere beş konu başlığı çerçevesinde oluşturulmuştur.

DMF, DMF Metotları ve DMF Kaynağı

DMF'nin, bir testle ölçülmesi amaçlanan aynı yetenek düzeyi göz önüne alındığında, bir odak grup ile bir referans grup arasında bir maddeyi doğru yanıtlama olasılığının farklı olması şeklinde tanımlanır (Fukuhara & Kamata, 2011). DMF'nin varlığı geleneksel olarak puan geçerliğine yönelik bir tehdit olarak görülmekte ve bu nedenle çok istenmeyen bir durum olarak değerlendirilmekte olup çeşitli katılımcı grupları için puan karşılaştırılabilirliğini büyük ölçüde etkileyebilmektedir. Testteki DMF'li maddelerin sayısı arttıkça DMF'nin etkisi de artar. Test geçerliği ve adillğini sağlamak için DMF değerlendirmesi, test yapımı sırasında rutin bir madde analizi prosedürü haline gelmiştir. DMF maddelerini değerlendirmek için istatistiksel yöntemler geliştirmek önemli bir görevdir ve bunların ilk kullanımı ikinci nesil DMF analizinin başlangıcı olarak görülmüştür (Finch & French, 2018; Zumbo, 2007). DMF belirlemek için kullanılan metotlar ise farklı sınıflar altında ele alınmıştır. Gomez-Benito ve Navas-Ara (2000) DMF belirleme metotlarını KTK, MTK, ki-kare ve faktör analizine dayalı olarak sınıflamışlardır. Penfield ve Camilli (2007) DMF belirleme metotlarını genel odds oranı, LR, oran farkları ve ortalama farklara göre metotlar şeklinde sınıflandırmışlardır. Gelin (2005) ise DMF belirleme metotlarını gözlenen gruplara ve örtük değişkene dayalı olmak üzere iki temel grup altında sınıflandırmıştır. Cinsiyet, etnik köken, yaşanılan bölge ya da ülke vb. kriterler ile bireyler gruplara ayrıldığında bu gruplar gözlenen gruplar olarak adlandırılmaktadır. Gözlenen gruplar arasında farklılaşmayı incelemek için de KTK bağlamında MH, dönüştürülmüş madde güçlüğü (Abedlazez, 2010)

ve LR metotları öne çıkarken MTK bağlamında ise Lord'un ki-karesi ve Raju'nun alan ölçümü en bilinenleridir. Örtük DMF analizinde ise gözlenen gruplar DMF analizine kıyasla üç adım daha atılması gerekmektedir. İlk adım, çeşitli aday modeller arasından en uygun ölçme modelinin (yani bir KMTK modeli) bulunması gereken bir model seçim aşamasıdır. Aday modeller, sınıf içi modeli sabit tutarak, örtük sınıf sayısı açısından farklılık gösterir. İkinci adım, seçilen ölçme modeli göz önüne alındığında, her bir kişi için bir örtük sınıf üyeliği elde etmektir. Son adım ise, örtük sınıfları referans alarak DMF'li maddeleri belirlemektir. Araştırmacılar (örneğin, Cho ve diğerleri, 2010; Cho ve diğerleri, 2013; Finch & Finch, 2013; Leite & Cooper, 2010; Maij-de Meij ve diğerleri, 2008, 2010), örtük sınıf üyeliğini belirledikten sonra örtük DMF'yi belirlemek için gözlenen gruplar için DMF belirleme metotlarını kullanmışlardır (Cho ve diğerleri, 2015). Angoff'a göre (1993) DMF belirleme metotları yıllar içinde ilerlemiş ve sonuçları daha doğru hale gelmiş olsa da test geliştiricileri ve madde yazarları genellikle bir çıkmazla karşı karşıya kalmaktadır: *"hiçbir tartışma, gayet mantıklı bazı maddelerin neden büyük DMF değerine sahip olduğunu açıklamaya yardımcı olmuyor gibi görünmektedir"*.

Son zamanlarda araştırmacılar DMF kaynaklarını hem nitel hem de nicel yollarla belirlemeye yönelik yöntemler formüle etmeye çalışmışlardır ve bu birleşik yaklaşım DMF analizinin üçüncü nesline işaret etmektedir. DMF kaynakları bilindiğinde, bu maddelerin revizyonuna daha fazla rehberlik edilebilir; buna karşılık, maddelerin kalitesi dikkatle izlenebilir ve test geçerliği ve adilliği korunabilir. Nitel yöntemler için, genellikle bir uzmanlar panelinden değerlendirilen ve DMF'li olarak kabul edilen maddelerin olası kaynaklarını belirlemeleri istenir (Ercikan, 2002; Zumbo, 2007).

Ancak uzman panelinin DMF belirleme konusunda mükemmel iş çıkardığı söylenemez. Bunun nedeni, uzmanların incelemesinin mesleki deneyimlerine dayanması olabilirken, bazı DMF maddeleri için farklı işleyiş, uzmanların deneyimleri dışındaki nedenlerden kaynaklanması olabilir. Ayrıca, uzmanlar DMF için öznel bir inceleme yapmış olabilir ve izlenimleri istatistiksel DMF analizlerinin sonuçlarıyla tutarsız olabilir (Cheng ve

diğerleri, 2020; Sireci ve diğerleri, 1998). Örneğin, Allalouf vd. (1999) İsrail Psikometri Giriş Testinin hem İbranice hem de Rusça versiyonlarındaki madde cevap verilerini analiz ederken maddelerin %34'ünün DMF sergilediğini bulmuştur. Bunun üzerine sekiz üyeli bir komiteden 42 DMF maddesi için olası kaynakları belirlemeleri istenmiştir. Komitenin kaynak konusunda bir anlaşmaya varamadığı yedi madde olmuştur. Bunun yanında uzman panelinden çıkan görüşlerle gözlenen gruplara göre yapılan DMF analizlerinin sonuçlarının tutarlı olmadığı görülmektedir. Örneğin Ceylan'ın (2022) yaptığı araştırmada madde yanlılığı süreci ters yönde işletilmiş ve test maddelerinin tamamı öncelikle uzman görüşüne gönderilmiştir. Çalışmanın sonucunda uzmanların yanlı olduğunu belirttiği dört maddede, LR yönteminin maddelerin birini tespit edebildiği, MH yönteminin maddelerden iki tanesini, SIBTEST yönteminin ise maddelerden üç tanesini tespit edebildiği görülmüştür. Ayrıca uzmanlarca yanlı olarak görülmeyen maddelerde de MH ve SIBTEST yöntemlerinin B ve C düzeyinde DMF belirlediği ortaya çıkmıştır. Bu nedenle, DMF'nin kaynakları hakkında nitel bilgi içermeyen yalnızca DMF değerlendirmesinin sonuçlarını sunmak, madde yazarları için madde revizyon sürecine rehberlik etmekte başarısız olur. Bunun aksine, DMF'nin kaynakları bilindiğinde, maddelerin çeşitli sınav gruplarında neden farklı işlev gördüğünün açıklanması daha olasıdır (Cheng ve diğerleri, 2020). DMF kaynağını bulmada nicel analizlerin sonuçlarını olumsuz etkileyen bir durum da grupların homojen olmamasıdır. Eğer bir araştırmacı gözlenen gruplara göre DMF belirlemeyi cinsiyete dayalı olarak test ederse, DMF'nin cinsiyete dayalı olduğunu varsayar ve böylece diğer potansiyel kaynakları engellemiş olur. Dolayısıyla, DMF'ye neden olan mekanizma doğrudan cinsiyetle ilişkili değilse veya sadece kısmen ilişkiliyse, bu gözlemlenen grup yöntemleri bunu tespit etmek için etkili araçlar olmayabilir (Finch & Finch, 2013). DMF için rakip bir paradigma, sınav katılımcılarının test maddelerine yanıt verme biçimlerinde niteliksel farklılıklar olduğunu ve bu farklılıkların gözlenen grup üyeliğinden ziyade örtük grup üyeliğinden kaynaklandığını varsayar (Samuelson, 2008). Bu muammaya tepki olarak Cohen ve Bolt (2005), tek biçimli DMF'yi belirlemek için daha önceki yöntemlerden bazı temel açılardan farklı olan bir karma model yaklaşımı tanımlamıştır.

Karma Madde Tepki Kuramı Modeli

Geleneksel MTK modelleri çeşitli varsayımlarla tanımlanır: tek boyutluluk, değişmezlik (invariance), yerel bağımsızlık, monotonluk ve bir test maddesine doğru cevap verme olasılığı ile sınava giren kişinin örtük özelliği ve test maddesinin özellikleri arasındaki ilişkiyi tanımlayan sürekli bir fonksiyonun varlığı. MTK modellerinin faydalı olduğu görülmüş olmasına rağmen bu varsayımların ihlaline karşı her zaman dayanıklı değildir. İhlaller meydana geldiğinde, "sınav katılımcıları ve test maddeleri arasındaki etkileşimlerin karmaşıklığını daha doğru yansıtmak" için daha karmaşık modellere ihtiyaç duyulabilir (Reckase, 2009). İlk varsayım (tek boyutluluk) ihlal edilirse, çok boyutlu bir MTK modeli uygulanabilir. Değişmezlik varsayımı ihlal edilirse, KMTK modelleri uygulanabilir (von Davier ve diğerleri, 2007). KMTK modelleri, MTK modellerinde bulunan madde parametrelerinin ilgili gruplar arasında değişmemesi varsayımına ihtiyaç duymamakta ve örtük sınıflar arasında madde parametrelerinin değişmesine izin vermektedir (De Ayala & Santiago, 2017; von Davier & Yamamoto, 2007). ÖSA ve MTK modellerinin karşılanması gereken güçlü varsayımları vardır. MTK ve ÖSA modellerini birleşimi olan KMTK modelleri, psikometrik araştırmalarda iki modelin de temel varsayımlarını ihlal edebilecek verilerin analizi için kullanılabilir (Rost, 1990).

KMTK modelleri niteliksel olarak farklı olan ancak sürekli bir örtük değişkene dayalı bir ölçme modelinin geçerli olduğu örtük alt grupların bir karışımını içeren durumların ele alınmasına olanak tanır. Bu modelleme çerçevesinde, öğrenciler hem sürekli bir örtük değişken üzerindeki konumlarına hem de örtük sınıf üyeliklerine göre karakterize edilebilir (De Ayala & Santiago, 2017).

Bir KMTK modeli için iki amaç vardır. Birincisi madde parametrelerini ve kişi parametrelerini (yani, deneklerin örtük özelliklerini) hesaplamak, diğeri ise deneklerin örtük sınıf üyeliklerini tahmin etmektir. KMTK modellerinde, her örtük sınıf için bir dizi madde parametresinin tahmin edildiği unutulmamalıdır. Kişi parametresi için, bir denek için yalnızca bir sınıfa ait olduğu varsayılsa da sınıf üyeliği kesin olarak bilinmemektedir, bu nedenle

tahmin sırasında her sınıfa üyelik koşuluna bağlı olarak bir kişi parametresi tahmin edilir. Sınıf üyeliği açısından, her deneğin tam olarak bir örtük sınıfa ait olduğu varsayılır, ancak sınıf üyeliği gözlemlenmediği için, tahmin edilenler her bir sınıfta olma olasılıklarıdır (Dai, 2009).

KMTK modellerinin faydası, doğrudan test verilerinden kaynaklanan boyutsallıkla oluşan grupları tespit etmek için bir araç sağlamasıdır. Bu gruplar anlamlı olduğu ölçüde, deneklerin nasıl ve neden bu şekilde yanıt verdiklerini anlamak için potansiyel olarak önemli bir araç sağlarlar (Cho & Cohen, 2010). KMTK modelleri aynı zamanda DMF'nin modellenmesi için de bir mekanizma sağlar, ancak bu mekanizma gözlenen grupların aksine örtük gruplarla ilgilenir (Cohen & Bolt, 2005).

En basit KMTK modeli, literatürde yaygın olarak kullanılan KRM modelidir. KRM modelinin geliştirilmesine katkıda bulunan ilk çalışmalar Kelderman ve Macready (1990), Mislevy ve Verhelst (1990) ve Rost (1990) tarafından gerçekleştirilmiştir. Bu çalışmalarda KRM, örtük bir özelliğin örtük bir kategorik değişkenle birleştirilmesi fikrine dayalı olarak türetilmiştir. Bu modelde tahmin edilecek parametreler Rasch güçlük ve sınıfa özgü yetenek parametreleridir:

(1)

$$P(y_{ij} = 1 | \theta_i) = \sum_{g=1}^G \pi_g \frac{\exp[(\theta_{ig} - b_{ig})]}{1 + \exp[(\theta_{ig} - b_{ig})]}$$

Burada π_g g örtük sınıfına ait olan bireylerin oranıdır, θ_{ig} i kişinin g örtük grubundaki yetenek düzeyidir; b_{jg} j maddesinin g örtük grubu için güçlüğüdür. Literatürde θ_{ig} için g örtük grup üyeliğinin alt simgesinin genellikle atlandığı belirtilmektedir, çünkü bireyler örtük grup üyelikleri göz önüne alındığında yalnızca bir yetenek parametresine sahiptir.

Modele bir parametre daha eklenerek, 2-parametrelilik (2PL) KMTK modeli ayırt edicilik varsayımını gevşetir ve KRM modelinin bir uzantısı olarak görülebilir. Örtük grupların her biri için bir 2PL modelinin geçerli olduğu varsayılırken madde güçlüğü ve ayırt edicilik

parametrelerinin farklı olmasına izin verilir. Mevcut arařtırmalar, farklı KMTK modellerini karřılařtırarak dođru sayıda örtük sınıf belirlemenin isabetliliđini arařtırmıřtır (Li ve diđerleri, 2009; řen, 2014). řen (2014) arařtırmasında KRM modellerinin tahmin edilmesinin, dađılımlar iki modlu ve tekdüze olduđunda verilerde hatalı örtük sınıf sorunlarına yol ađtıđını, Karma 2PL ve Karma 3PL MTK modellerinin örtük normal dıřılıđa (non-normality) karřı daha dayanıklı olduđu bulmuřtur. Model seđiminde Akaike bilgi kriteri (AIC) ve Bayesian bilgi kriteri (BIC) kullanılan arařtırmada çođu kořul için, dođru modelin seđiminde BIC indeksinin performansı AIC indeksinden daha iyi olmuřtur.

Örtük üyeliđin iyileřtirilmesi, test uzunluđundan bađımsız olarak KMTK 2PL modeli için en iyi olarak bulunmuřtur. Özellikle, daha karmařık bir model olan KMTK 2PL modeli, örtük normalliđin olmaması nedeniyle daha az hatalı örtük sınıf çözümleri üretme eđiliminde olmuřtur.

(2)

$$P(y_{ij} = 1 | \theta_i) = \sum_{g=1}^G \pi_g \frac{\exp[a_{jg} (\theta_{ig} - b_{ig})]}{1 + \exp[a_{jg} (\theta_{ig} - b_{ig})]}$$

Burada a_{jg} g örtük grubu için j maddesinin ayırt ediciliđidir.

Benzer řekilde, 3 parametrelili lojistik (3PL) KMTK modeli, madde güçlüđü, ayırt edicilik ve tahmin parametrelerinin örtük gruplar arasında farklı olmasına izin verilen örtük grupların her biri için 3PL modelinin geđerli olduđunu varsayar. Karma 3PL modeli sınırı sayıda çalıřmada uygulanmıřtır (Choi ve diđerleri, 2015; Cohen & Bolt, 2005; Jensuttiwetchakul ve diđerleri, 2016; Li ve diđerleri, 2009; řen, 2014). Karma 3PL modelinde dođru yanıt olasılıđı řu řekilde gösterilir:

(3)

$$P(y_{ij} = 1 | \theta_i) = \sum_{g=1}^G \pi_g [c_{jg} + (1 - c_{jg}) \frac{\exp[a_{jg} (\theta_{ig} - b_{ig})]}{1 + \exp[a_{jg} (\theta_{ig} - b_{ig})]}$$

Burada c_{jg} g örtük grubu için j maddesinin düřük asimptot parametresidir.

KMTK modellerinin önemli bir özelliği, örtük sınıflar arasındaki ve her bir örtük sınıf içindeki tepki verilerinin heterojenliğinin iyi bir şekilde temsil edilebilmesidir. Sınıf üyeliğinin örtük gruplar arasındaki tepki örüntülerinin niteliksel farklılıklarını yansıttığı düşünülürken, yeteneğin bir örtük grup içindeki bireyler arasındaki niceliksel farklılıkları yansıttığı düşünülmektedir. Karma modeller daha fazla esneklik sağlar ancak bazı sınırlamalar nedeniyle eleştirilmektedir (Zhang,2017). Bu sınırlamalardan birisi yuvalanmış veri yapısı var olduğunda ortaya çıkan düzeyliliğin sonuçları etkileme durumudur. MTK modelleri tarafından göz ardı edilen bu unsur, madde yanıtlarının gözlenemeyen heterojenliğinin bir kısmının, bireylerin okullardaki öğrenciler gibi gruplar halinde iç içe geçtiği çok düzeyli veri yapılarından kaynaklanmasındır. Bu durumda, aynı okul bağlamını paylaşan öğrencilerin örtük özellik düzeyleri açısından farklı okullara mensup arkadaşlarına kıyasla daha benzer olduklarını varsaymak mantıklıdır ve sonuç olarak, ilgili madde yanıtlarının bağımsız olduğu varsayılmaz (Gnaldi ve diğerleri, 2015). KMTK bağlamında bu tür modeller her hiyerarşik seviyede farklı örtük özelliklerin belirlenmesine dayanmaktadır. KMTK modelleri, verilerin kümelenmiş yapısını hesaba katmak için çok düzeyli modellere genişletilip ÇDKMTK modeli olarak adlandırıldı.

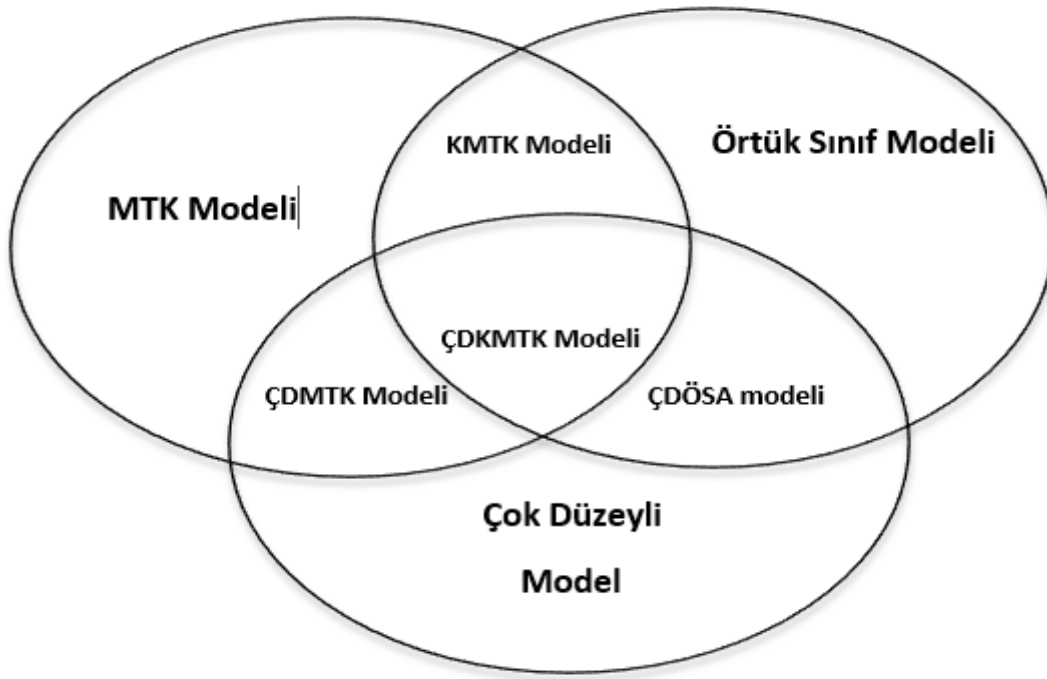
Çok Düzeyli Karma Madde Tepki Kuramı Modeli

Hiyerarşik verilerin ne kadar önemli olduğunu fark ettikten sonraki adım, istatistiksel tekniklerin bu hiyerarşik yapıyı dikkate almasının yollarını düşündürmektir (Raudenbush & Bryk, 2002). ÇDKMTK modelleri, birimlerin (örneğin öğrenciler) daha üst düzey birimler (örneğin sınıflar, okullar veya ülkeler) içinde kümelenmesini hesaba katarak, iç içe geçmiş birimlerdeki ortaklıklar nedeniyle ortaya çıkabilecek cevaplardaki benzerlikler nedeniyle MTK'nın temel varsayımlarından olan yerel bağımsızlığın ihlal edilmemesini sağlar (Dras, 2023). Kümelenmeden kaynaklanan bağımlılıkları hesaba katmak için ÇDKMTK modelleri Cho'nun (2007) araştırmasından uyarlanan Şekil 1'de gösterildiği gibi üçüncü bir bileşeni entegre eder. Karma modellemeye ek olarak, DMF literatürüne oldukça yeni bir katkı da bu tür değerlendirmelerde yaygın olan çok düzeyli veri yapısıyla başa çıkma yöntemlerinin

ortaya çıkması olmuştur (French & Finch, 2010). ÇDKMTK modeline geçiş, örtük sınıf tabanlı DMF analizi araştırmalarının hiyerarşik veri setleri üzerinde yürütülmesine olanak tanır.

Şekil 1

Çok Düzeyli Yapının Analiz Kapsamına Alındığı Karma Model Şeması



Halen birçok DMF değerlendirmesinde bu tür bir örnekleme yapısı analizlerde genellikle göz ardı edilmekte, bu da potansiyel olarak parametre tahmini ve hipotez testi sonuçlarını tehlikeye atmaktadır. Özellikle, çok düzeyli yapı göz ardı edildiğinde standart hataların yanlış tahmini ortaya çıkmakta (Raudenbush & Bryk, 2002), bu da yanlış istatistiksel testlere yol açabilmektedir ve DMF analizleri için, yanlış DMF belirlenmesi (Finch & Finch, 2013; French & Finch, 2010), şişirilmiş Tip I hata oranları (Samuelsen, 2005) şeklinde sonuçlanmaktadır.

ÇDKMTK modelleri çok düzeyli veri setlerinde heterojenlikle başa çıkmak için (Cho & Cohen, 2010; Tay ve diğerleri, 2011) önerilmiştir. ÇDKMTK modelinde doğru bir yanıtın koşullu olasılığı şu şekilde verilebilir:

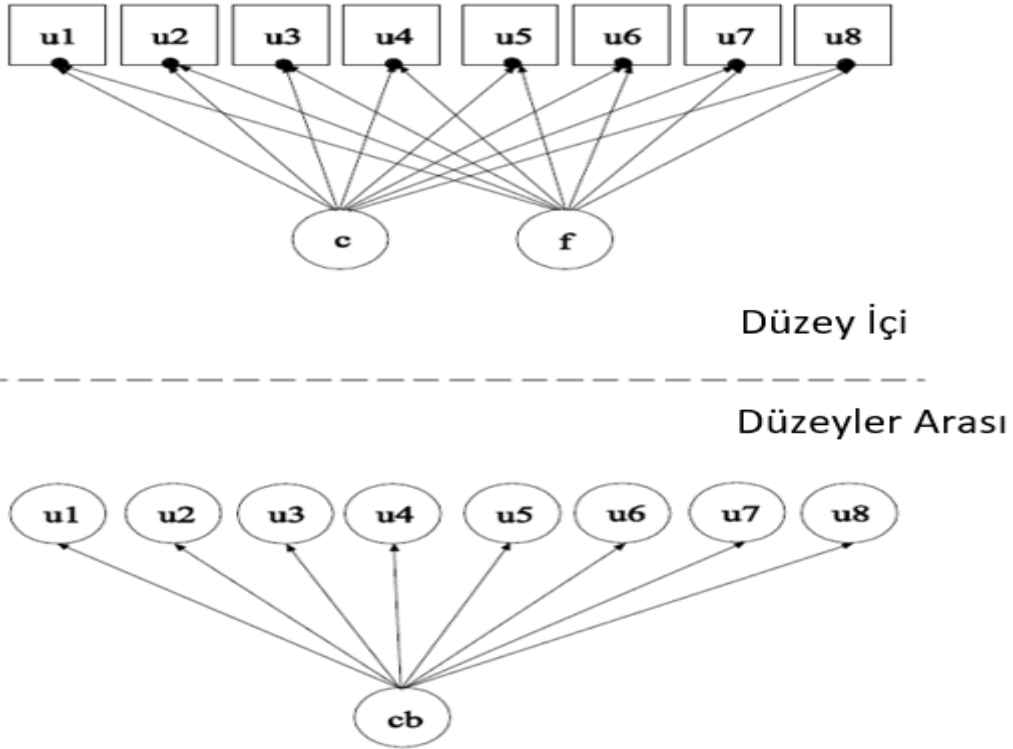
$$P(Y_{ijt} = 1 | g, k, \theta_{jtgk}) = \frac{1}{1 + \exp[-(\theta_{jtgk} - \beta_{igk})]}$$

(4)

burada y_{ijt} , t okulundaki j sınav katılımcısının i maddesine verdiği yanıttır, burada $g = 1, \dots, G$ öğrenci düzeyindeki örtük sınıflar için bir indistir, $k = 1, \dots, K$ okul düzeyindeki örtük sınıflar için bir indis, $j = 1, \dots, J$ sınava girenler için bir indis, $t = 1, \dots, T$ okullar için bir indistir, $i = 1, \dots, I$ maddeden oluşan bir test için, β_{igk} , g ve k örtük sınıfları için i maddesinin güçlüğüdür ve θ_{jtgk} , t okulunda ve g ve k örtük sınıflarında j sınav katılımcısının yeteneğidir. θ_{jtgk} yetenekleri normal dağılımla ortalama μ_{gk} ve varyans θ_{gk}^2 sahiptir. Şekil 2'de ÇDKMTK modeline ait bir görsel örneği verilmiştir.

Şekil 2

ÇDKMTK Modelinde Grup İçi ve Gruplar arası Yapıyı Gösteren Örnek Yapı



Kaynak: Muthen ve Muthen (1998-2017).

Modelde hem birey düzeyinde sınıflar hem de düzeyler arası sınıflar bulunmaktadır. Birey düzeyindeki sınıflar, örneğin öğrenciler gibi bireylerden oluşur. Düzeyler arası sınıflar, örneğin okullar gibi kümelerden oluşur. Modelin düzey içi kısmında, u1, u2, u3, u4, u5, u6, u7 ve u8 faktör göstergelerinin bireysel düzeydeki kategorik örtük değişken c'nin sınıfları arasında değişen eşiklere (thresholds) sahip olduğu bir KMTK modeli belirtilmiştir. Faktör göstergelerini işaret eden okların sonundaki içi dolu daireler, faktör göstergelerinin eşiklerinin rastgele olduğunu göstermektedir. Bunlar, düzeyler arası kısımda u1, u2, u3, u4, u5, u6, u7 ve u8 olarak adlandırılmaktadır. Rastgele eşikler u1, u2, u3, u4, u5, u6, u7 ve u8, kümeler (düzeyler arası birimler) arasında değişen sürekli örtük değişkenler oldukları için modelin düzeyler arası kısmında daireler içinde gösterilmiştir. Rastgele eşiklerin sınıf içi varyansı yoktur ve sınıflar arası kategorik örtük değişken cb'nin sınıfları arasında değişiklik gösterirler.

Kavramsal olarak, ÇDKMTK modelinin ayırt edici bir özelliği, gözlemlenen hiyerarşik birimlerle sınırlandırılmamış, ilgilenilen yapılara (örneğin kişilik, tutumlar ve iklim) ilişkin ölçüm gruplamalarının organik olarak türetilmesidir. ÇDKMTK modeli, sadece teorik meseleleri ele almak için yeni metodolojik olanaklar sunarak örgütsel araştırmalara katkıda bulunmakla kalmaz, aynı zamanda yeni kavramsal ve metodolojik gelişmeleri de teşvik edebilir (Tay ve diğerleri, 2011). Modelin öğrenci düzeyi, soruları yanıtlama stratejilerinde farklılık gösteren örtük sınıfların var olup olmadığını belirleme fırsatı sunmaktadır. Okul düzeyindeki bilgiler, örtük sınıflar arasındaki müfredat veya pedagojik farklılıklardan kaynaklanabilecek olası farklılıkları ortaya çıkarmak için kullanılabilir (Cho & Cohen, 2010).

Lee vd., (2018) ÇDKMTK modelinin gerekliliğini incelemek için, çok düzeyli verilerin uygun olmayan şekilde modellemenin etkisini araştıran ek araştırmalara ihtiyaç olduğunu belirtir. Bu alanda yapılan simülasyon çalışmaları, daha yüksek seviyeli bir iç içe geçme yapısını göz ardı eden yanlış belirlenmiş modellerin, bir karma büyüme modelinde (Chen ve diğerleri, 2010) ve bir örtük sınıf modelinde (Kaplan & Keller, 2011; Park & Yu, 2016)

daha az doğru tahminlere ve daha düşük sınıflandırma doğruluğuna neden olduğunu göstermiştir. Chen vd. (2010), uygun birçok düzeyli karma büyüme modeliyle karşılaştırıldığında, karma büyüme modeli kullanan kişilerin sınıflandırılmasının daha az doğru olduğunu ve sınıflandırma doğruluğunun küme içi korelasyon (KİK), karma oranlarından ve rastgele etkilerin sınıf içi varyanslarından ve kovaryanslarından etkilendiğini bulmuştur. Bunun yanında modelin yetersiz kalabileceği bazı durumlar da mevcuttur. Çünkü çoğu zaman, büyük ölçekli test programları, sınava girenlere birden fazla yapıyı ölçen sınavların uygulanmasını içerir. Tek boyutlu modeller, DMF'nin yorumlanmasına yardımcı olacak faydalı bir mekanizma sağlamamaktadır. Buna karşılık, çok boyutlu bir model, ikincil yetenekleri izole etmenin ve yorumlamanın ihtiyatlı olacağını öne sürer (Camilli, 1992). Bu sebeple çok boyutlu çok düzeyli karma madde tepki kuramı modeli ortaya çıkmıştır ve model hem birçok test programının çok düzeyli doğasını hem de birçok test programının çok boyutlu doğasını içermektedir.

Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramı Modeli

Klasik MTK modellerini karakterize eden tek boyutluluk varsayımına göre, bir dizi maddeye verilen yanıtlar yalnızca tek bir örtük özelliğe bağlıdır ve bu da eğitim ortamında öğrencinin yeteneği olarak yorumlanabilir. Eğer tek boyutluluk sağlanamazsa, test maddeleri gerçekten de birden fazla yeteneği ölçtüğü için öğrencilerin yeteneklerini tek bir puanla özetlemek yanıltıcı olabilir ve bir öğrencinin yeteneğinin doğru bir şekilde değerlendirilmesi çok boyutlu yöntemlerin geliştirilmesini gerektirir (Gnaldi ve diğerleri, 2015; Jiao ve diğerleri, 2012). Yao ve Li (2015), çok boyutlu verilere tek boyutlu MTK'nın uygulanmasının, iki grup farklı yeteneklere sahip olduğunda maddenin DMF'li olarak algılandığı şişirilmiş bir yanlış pozitif oran verdiğini bulmuş ve çok boyutlu MTK'nın kullanılmasını önermiştir.

Eğitsel ve psikolojik testler genellikle birbiriyle ilişkili birden fazla yapıyı ölçen birden fazla alt testten oluşur. Buna ek olarak, tanınmış, büyük ölçekli bir uluslararası eğitim değerlendirmesi olan TIMSS matematik testleri, matematik beceri setlerinin üç bilişsel

alanının (bilme, uygulama ve akıl yürütme) yanı sıra üç konu alanına (sayılar, geometrik şekiller ve ölçüler ve veri gösterimi) sahip çoklu boyutlara dayanmaktadır. Bu nedenle, tartışılan doğrulayıcı KMTK modelinin uygulamalarının kapsamını genişletmek amacıyla, modelin çok boyutlu bir uzantısını dikkate almak faydalı olacaktır (Jeon, 2019). ÇBÇDKMTK modeli, yukarıda tartışılan ÇDKMTK modeli ile aynı şekilde, ancak birden fazla örtük özelliğin aynı anda değerlendirildiği durumlarda kullanılmak üzere tasarlanmıştır. KMTK modelinin sağladığı potansiyel avantajlar ve ÇDKMTK modelinin çok düzeyli verilerle DMF değerlendirmesi için güçlü yönleri göz önüne alındığında, ÇBÇDKMTK modeli hem çok boyutlu hem de çok düzeyli verileri barındıran modelin bir uzantısını tanımlamaktadır. ÇBÇDKMTK modeli şu şekilde ifade edilebilir:

$$P(Y_{ijt} = 1 | g, k, \theta_{jtgkm}) = \frac{1}{1 + \exp[\Sigma - (\theta_{jtgkm} - \beta_{igk})]}$$

(5)

Denklem 5'teki terimler Denklem 4'te tanımlandığı gibidir ve θ_{jtgkm} , g ve k örtük sınıflarından t okulundaki j katılımcısının m boyutu için örtük özelliğidir.

ÇDKMTK modelinde olduğu gibi, ÇBÇDKMTK modelinde de hem düzey 1 hem de 2'de örtük sınıf üyeliği tahminlerinin yanı sıra sınıfa özgü madde güçlüğü tahminleri sağlanır. Ancak bu durumda, madde güçlükleri tek boyutluluk varsayımı altında değil, doğru çok boyutlu model altında tahmin edilir; birden fazla örtük yapı aynı anda değerlendirildiğinde durum böyle olmayacaktır ve ÇDKMTK modelinde olduğu gibi, örtük özellik ortalamaları ve varyansları, örtük özellik için bir ölçek oluşturmak, yani modeli tanımlamak amacıyla ortalaması ve varyansı 0 ve 1 olarak ayarlanmış olan hariç, hem düzey 1 hem de 2'deki her örtük sınıf için tahmin edilir (Finch & Finch, 2013). Kanovský vd.'nin (2020) çalışmalarının sonucu katılımcıların puanlarının gözlenemeyen (örtük sınıf) birey ve ülke üyeliğine, ölçme aracının çok boyutlu yapısına bağlı olduğunu göstermiş ve bireylerin ve ülkelerin karşılaştırmalı psikolojik değerlendirmesinde ÇBÇDKMTK modelinin kullanımını haklı çıkarmıştır. Bununla birlikte Embretson'un (2007) da belirttiği gibi, KMTK modeli kullanmanın bir dezavantajı, modeller tarafından tanımlanan örtük grupların nitel anlamını

yorumlamada zorluk yaşanmasıdır. Buna ek olarak, gözlenen tepki verileri, özellikle karmaşık istatistiksel modeller için parametrelerin kesin tahminlerini üretmek için tek başına yeterli olmayabilir. Bu sebeple eğitim ve davranış bilimlerine ait alanlardaki çeşitli çalışmalarda araştırmacıların birincil ilgi alanına giren değişkenlere ek olarak ikincil bilgiler de elde edilmektedir. Alanyazında ortak değişkenler (kovaryant değişkenler) olarak da adlandırılan bu gibi ikincil bilgiler genel olarak etnik köken, cinsiyet ve eğitim süresi gibi arka plan özelliklerini içerir. Bu tür "dışsal değişkenler" bazı zamanlar istatistiksel modellemede görmezden gelinse de, modellenen birincil ilgi değişkenleriyle potansiyel olarak önemli ilişkilere sahip olabileceklerine inanılmaktadır (Li ve diğerleri, 2016). Bu sebeple analize ortak değişkenleri eklemek DMF kaynağını bulmada önemli katkı sağlayacaktır.

Ortak Değişkenler ile Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramı Modeli

KMTK modellerini kullanmanın en zorlu görevlerinden biri popülasyonun heterojenliğine neyin sebep olduğunu belirlemektir; bu aynı zamanda araştırmacıların örtük sınıf modellemesi kullanırken sıklıkla sordukları bir sorudur. Bireyle ilgili özellikler gibi ikincil bilgilerin modellenen değişkenlerin tahminini karıştırabileceği yaygın olarak gösterilmiştir. Bu potansiyel olarak etkili ortak değişkenlerin dahil edilmesi, örtük sınıfı tanımlamanın zorluklarını hafifletmeye ve model parametrelerinin tahminini iyileştirmeye yardımcı olabilir. Hem karma modelleme hem de geleneksel MTK literatüründe ortak değişkenlerin dahil edilmesinin faydalarına ilişkin kanıtlar bulunmuştur (Zhang, 2017).

Smit vd.'ye (1999) göre ortak değişkenler, verilerdeki örtük sınıfların belirlenmesine yardımcı olmak için özellikle yararlı olabilir. Ayrıca mevcut çalışma ile Karadavut vd. (2019), Dai, (2009, 2013), Li, (2014) ve Li vd.'nin (2016) çalışmalarında örtük sınıf değişkeniyle ilişkili ortak değişkenlerin modele dahil edilmesi ile model parametre kestirimlerinin nasıl iyileştiği, standart hataların nasıl azaltılabileceği ve sınıflandırma doğruluğunun nasıl geliştirilebileceği gösterilmiştir. Benzer şekilde daha doğru model parametre tahminleri elde etmek, örtük sınıf ataması ve örtük sınıfların numaralandırılması gibi farklı amaçlara ulaşmak için diğer karma modelleme çerçevelerinde olduğu gibi KMTK'da da ortak

değişkenlerin dahil edilmesi (örneğin, Lubke & Muthen, 2005; Smit ve diğerleri, 1999) önerilmiştir (Li ve diğerleri, 2016). Cho vd. (2006), ortak değişken ve örtük gruplar arasında %80 veya daha fazla örtüşme olduğunda ve tahmin simülasyonu yapılmadığında, ortak değişken kullanımının hem madde güçlüklerinin hem de grup üyeliğinin tespitini iyileştirdiğini bulmuştur (Cho, 2007).

Ortak değişkenler Karadavut vd. (2019) tarafından gözlenen ve örtük olmak üzere ikiye ayrılmıştır. Gözlenen ortak değişkenlere örnek demografik bilgiler ve evdeki kitap sayısı iken örtük ortak değişkenlere örnek tutumlar ve inançlardır. Gözlenen ortak değişkenler her zaman yeterince bilgilendirici değildir, ancak örtük sınıfların oluşmasına neden olan değişkenle yalnızca orta derecede ilişkili olma eğilimindedirler. Örneğin DMF bağlamında, gözlenen gruplama değişkenleri gruplar arası farklılıkların nedenlerini açıklamada çok yardımcı olmamıştır (Cohen & Bolt, 2005). Ayrıca, bu gözlenen değişkenler tarafından açıklanan varyans oranları, anlamlı olsalar bile genellikle küçüktür. Çalışma gözlenen ve örtük ortak değişkenlerin çıkarılan temel örtük sınıfların sayısı üzerinde bir etkiye sahip olmadığını, ancak örtük sınıfların özelliklerini açıklamaya yardımcı olduklarını göstermiştir. Ortak değişkenler örtük sınıf üyelik oranlarını değiştirmiş, ancak sınıf yeteneği ortalamaları üzerinde güçlü bir etki göstermemiştir. Örtük ortak değişkenler, gözlenen ortak değişkenlere kıyasla örtük sınıf üyeliğinin özelliklerini açıklamak için daha kullanışlıdır (Karadavut ve diğerleri, 2019).

Örtük sınıf sayısı arttıkça tahmin edilen parametre sayısı da hızla artmaktadır. Bunun nedeni, hem her bir sınav katılımcısı için yetenek ve üyelik parametrelerinin hem de her bir sınıf için madde parametrelerinin tahmin edilmesi gerekmesidir (Choi, 2010). Bu durum, ikiden fazla örtük sınıf tanımlandığında modelin yorumlanmasını zorlaştıracaktır. Bu durumda, tahmin edilecek parametre sayısındaki artışla başa çıkabilmek için büyük bir örneklem büyüklüğüne ihtiyaç duyulmaktadır (Li ve diğerleri, 2015). Buna ek olarak, KMTK modelleri tek başlarına katılımcıların örtük kategorizasyonunu açıklamamaktadır. Bu nedenle, örtük grupların nasıl ve neden farklı olduğunun anlaşılmasını iyileştirmek için

ikincil/ek bilgilerin dahil edilmesi önerilmektedir (Smit ve diğeri, 1999). Ortak deęişken seçimi adımımda ise, ilişkinin gücünü belirlemek için aday ortak deęişkenler ile ilgilenilen örtük deęişken arasındaki korelasyon katsayıları incelenmelidir (Li, 2014; Li ve diğeri, 2016; Karadavut ve diğeri, 2019; Uysal Saraç, 2022).

İlgili Araştırmalar

Cho ve Cohen'in (2010) simülatif ve gerçek veri ile ÇDKMTK modelini kullanarak DMF belirlemesi yaptıkları çalışmada hem öğrenci düzeyinde hem de okul düzeyinde çok düzeyli madde tepki verilerinin modellenmesi sağlanmıştır. Bunun yanında hem simülatif hem de gerçek veri analizinde ortak deęişkenlerin analizlere dahil edilip edilmediği koşullar da çalışmada yer almıştır. DMF analizleri için elde edilen örtük sınıflara ait madde güçlükleri Highest Posterior Density (HPD) metodu ile incelenmiş ve DMF'li maddeler bu şekilde belirlenmiştir. 40 maddelik test için DMF etki büyüklükleri, okul düzeyinde DMF'li madde yüzdeleri, öğrenci ve okul örneklem büyüklüğü kombinasyonları iki örtük sınıf olasılığı modeli olmak üzere koşullar simülatif olarak incelenmiştir. Gerçek veri ise 2006 yılında geniş ölçekli bir matematik sınavına aittir ve %20 rastgele örnekleme ile 206 okul ve 8362 öğrenciye ait veri araştırma kapsamında incelenmiştir. Simülatif veri için label switching (etiket deęişimi) incelenmiş, yakınsama (convergence) gözlenmiş, model seçimi yapılmış ve generating (üretme) parametrelerinin iyileştirilmesi yapılmıştır. Deneysel örnekte, Bayesian Information Criteria (BIC) değerlerine göre, ortak deęişkenli ve ortak deęişkensiz modele dayalı olarak öğrenci düzeyinde dört ve okul düzeyinde iki örtük sınıf seçilmiştir. Her öğrenci ve okul örtük sınıfı için sınıfa özgü madde güçlükleri elde etmenin ve bunları aynı ölçekte ifade etmenin mümkün olduğu gösterilmiştir. Araştırmada belirlenen DMF'li maddelerinin sayısının, gözlenen gruplar kullanılarak yapılan olağan DMF analizinden beklenenden daha fazla olduğu görülmüştür.

Finch ve Finch (2013) çalışmalarında üçüncü sınıf öğrencileri için hazırlanan bir başarı testinde engelli olan ve olmayan öğrencilerden elde edilen verilerdeki DMF'nin

belirlenmesi için ÇDÇBKMTK modeli analizini kullanmışlardır. Çok boyutlu yapının matematik ve dil dersleri ile oluştuğu belirtilen çalışmada çok düzeyli yapının birinci düzeyinde öğrenciler, ikinci düzeyinde ise okullar yer almaktadır. Bunun yanında DMF kaynağına ulaşmak için birinci düzeyde bireylerin engel durumu, ikinci düzeyde okullarda bulunan engelli öğrencilerin okuldaki toplam öğrenci sayısına oranı ortak değişken olarak analize dahil edilmiştir. DMF'nin varlığını belirlemek için ÇDÇBKMTK modeli yaklaşımın kullanılmasını, gözlenen grupların kullanılmasına kıyasla DMF'nin nedeni hakkında bilgi edinmek için potansiyel olarak üstün şekilde lanse etmişlerdir. Çalışmaya ait analiz sonuçları, çok boyutlu modelin DMF'nin doğası hakkında tek boyutlu ayrı modellere göre daha eksiksiz bilgi sağladığını göstermiştir. Birden fazla boyutun dahil edilmesinin, öğrencilerin öğrendikleri ve değerlendirildikleri gerçek dünya bağlamlarını yansıtacak şekilde, birden fazla yapıdaki göreceli yeterliliklerini eş zamanlı olarak kullanılarak öğrenci ve okul düzeylerinin daha eksiksiz bir karakterizasyonunu sağladığı belirtilmiştir.

Choi vd.'nin (2015) çalışmaları TIMSS 2007 uygulamasında matematik performansındaki farklılıklara nelerin katkıda bulunabileceğini araştırmayı amaçlamaktadır. 11 çoktan seçmeli ve 15 tane iki kategorili hale getirilmiş kısa cevaplı yapılandırılmış madde olmak üzere 26 maddeden oluşan 4. sınıf düzeyine ait test çalışma kapsamında analiz edilmiştir. Grup üyeliğinin belirlenmesine yardımcı olmak için üç parametrelili KMTK modeline bir ortak değişken eklenen çalışmada açıklayıcı bir amaç için kullanılan ortak değişken, örneklemden katılımcılara sunulan internet erişimiyle ilgili bir soruya verilen yanıtlardan oluşmuştur. Çalışma sonucunda KMTK modelleme yaklaşımının avantajı olarak örnekleme tespit edilen örtük sınıfların üyelerinin özelliklerinin ve belirli örtük sınıflar için farklı şekilde daha zor veya daha kolay olan madde özelliklerinin, DMF'nin olası nedenlerinin daha doğrudan yorumlanmasına yardımcı olabilmesi olduğu ifade edilmiştir. Ayrıca çalışmanın sonuçlarının, ülkeler arasındaki farklılıkların nasıl ortaya çıkabileceğini anlamak için KMTK modellerinin nasıl kullanılabileceğini göstermekte olduğu belirtilmiştir. Bunun yanında ortak değişkenlerin modele dahil edilmesinin, örtük sınıfların tespit

edilmesinin yanı sıra bu gruplar arasındaki farklılıkların yorumlanmasını iyileştirme potansiyeline sahip olduğu gösterilmiştir.

DeMars ve Lau'nun (2011) araştırmalarının amacı, nispeten küçük madde kümelerinin DMF gösterdiği bağlamlarda sınıfların, madde parametrelerinin ve DMF etkilerinin geri kazanımının doğruluğunu değerlendirmektir. Veriler, iki sınıflı bir parametrelili lojistik (1PL) KMTK modeline uygun şekilde simüle edilmiştir. Simülasyon koşulları etki, DMF'li madde sayısı ve değiştirilmeyen (invariant) madde sayısı şeklinde $2 \times 2 \times 2$ tasarımıdır. Etki var ve yok şeklinde iki düzeyli, DMF'li madde sayısı dört ve sekiz olarak ayrıca değiştirilmeyen madde sayısı 10 ve 20 olarak belirlenmiştir. Genel olarak, çalışmadan elde edilen sonuçlara göre örtük DMF belirlenmesi için KMTK modellerinin kullanımının sorunlu olabileceği ortaya konulmuştur. Sınıf üyeliği iyileşmesi test edilen tüm koşullarda zayıf olmuştur. Ayırıcılık parametreleri sabit maddelerin yanı sıra grup etkisi olmadığında DMF maddeleri için de iyi tahmin edilmiştir. Ancak grup etkisi olduğunda, DMF maddeleri için ayırt edicilikler pozitif yanlı olmuştur. Grup etkisi olmadığında, DMF etki tahminleri pozitif yönde yanlı olma eğilimindedir. Genel olarak, daha az sayıda madde olması daha yanlı tahminler ve daha büyük standart hatalarla ilişkilendirilmiştir.

Chen ve Jiao (2014) çalışmalarında geleneksel gözlenen gruplara dayalı DMF yaklaşımının kullanılmasından kaynaklanan sorunları ele almış ve örtük sınıf üyeliğinin karakterizasyonunda etkili olan hem arka plan (background) hem de bilişsel ortak değişkenleri analiz etmişlerdir. Bu kapsamda PISA 2009 okuma değerlendirmesi ve öğrenci anketi analiz edilmiştir. Elde edilen sonuçlara göre örtük sınıf yaklaşımı ile belirlenen DMF maddelerinin sayısının gözlenen gruplar DMF yöntemi ile belirlenen daha fazla olduğunu göstermektedir. Bunun yanında tek bir gözlenen değişkenin birden fazla gözlenen değişkenin veya keşfedilmemiş (unexplored) gözlenen gruplama değişkenlerinin neden olduğu potansiyel DMF maddelerini taramak için yetersiz olabileceğini ve DMF analizinde sıklıkla incelenen gözlenen gruplama değişkenlerinin bazı maddeler için örtük DMF'nin gerçek nedeni olmadığını yansıtmakta olduğu belirtilmiştir. Ayrıca gözlenen grupların

madde performansı açısından homojen olmadığı çünkü fiilen avantajlı veya dezavantajlı grupları temsil etmedikleri ifade edilmiştir. Bu çalışma sonucunda örtük sınıf üyeliğini tahmin etmek ve örtük DMF'nin açıklanmasını kolaylaştırmak için bir dizi faktörün kullanılabileceğini göstermekte olduğu ve hem arka plan hem de bilişsel ortak değişkenlerin, DMF sergileyen örtük sınıfların özelliklerini birlikte tanımladığı görülmüştür. Tek bir gözlenen grup değişkeni ile örtük DMF kaynaklarını tam olarak öngörmek için yetersiz kaldığı ve örtük sınıf temelli DMF yaklaşımının, birden fazla değişkenin araya gelmesinden kaynaklanan potansiyel DMF maddelerinin taranması için olası bir yöntem olduğunu göstermiştir.

Gürkan (2021) tezinde belirli bir örnekte, uluslararası büyük ölçekli değerlendirme verilerinin analizinden elde edilen geleneksel tahminlere dayalı ülke içi korelasyonlar arasındaki varyansı araştırmak ve bu varyansın ne ölçüde ölçüm hatası miktarındaki heterojenlikten ve verilerin kümelenmiş yapısından kaynaklandığını belirlemeyi amaçlamıştır. Bu kapsamda hem verilerin kümelenmiş yapısının hem de heterojen ölçüm hatasının, ILSA'lara (International large-scale assessments-Uluslararası geniş ölçekli değerlendirmeler) katılan ülkeler arasında arka plan verileri ile yeterlilik ölçekleri arasında bildirilen korelasyonlar üzerindeki etkisini incelenmiştir. Bu bağlamda, rakip modelleme tekniklerinin çalışma özellikleri PISA 2012 verilerine yapılan uygulamalarla araştırılmış ve ülkeler arasında matematik öz yeterliliği ile matematik başarısı arasındaki korelasyonların tahminleri bu çalışmanın odak noktasını oluşturmuştur. Tezde, PISA 2012 veri tabanından elde edilen veriler kullanılarak ölçme hatası ve kümelenmenin istatistiksel çıkarımlar üzerindeki etkisi kapsamlı bir şekilde incelenmiştir. Araştırma sonuçlarına göre ilk dört aşamada gerçekleştirilen modellerin adım adım iyileştirilmesi, geleneksel korelasyonların eksik bir resim sunabileceğini ve sonuçta ortaya çıkan çıkarımların şaşırtıcı ve hatta yanıltıcı olabileceğini göstermiştir. Çok düzeyli madde tepki kuramı modellerinden elde edilen bulgular, ölçüm hatası miktarının ülkeler arasında farklılık gösterdiğini ve verilerdeki daha büyük miktardaki rahatsız edici (nuisance) korelasyon tahminlerinin zayıflamasına yol açtığını ortaya koymuştur. Araştırmada kullanılan ÇDÇBKMTK modelleri, PISA verilerinin

özelliklerinin gerektirdiği gibi, yalnızca ölçüm hatasını ve verilerdeki kümelenmeyi uygun şekilde hesaba katmakla kalmamış, aynı zamanda ülkelerdeki farklı alt popülasyonların olasılığını da dikkate almıştır.

Lee vd.'nin (2017 çalışmalarında çok düzeyli bir karma madde tepki modelinin gerekli olduğu durumları göstermek amacıyla, bir karma madde tepki modeli için çok düzeyli bir yapının göz ardı edilmesinin sonuçları araştırılmıştır. Çalışma kapsamında çok düzeyli veri yapısı bağlamında model seçim performansını ve parametre doğruluğunu etkileyebilecek üç simülasyon koşulu teorize edilmiştir. Simülasyon Çalışması 1, iç içe geçmiş bir veri yapısı mevcut olsa bile, tek düzey yerine çok düzeyli bir karma madde tepki modeline uymanın yanlılık, RMSE (Root Mean Squared Error) ve/veya SE (Standart Error- Standart Hata) oranı açısından genellikle faydalı olduğunu ancak her zaman faydalı olmadığını ortaya koymuştur. Bu nedenle, Simülasyon Çalışması 2'de, gerçekte iç içe geçmiş bir veri yapısı olmadığına çok düzeyli bir karma madde tepki modeli uygulamanın sonuçları araştırılmıştır. Simülasyon sonuçları KİK aralığı için (0,1 ile 0,3), daha yüksek seviyeli bir yuvalama yapısını göz ardı eden karma madde tepki modellerinin, küme sayısı 24'ten ve küme boyutu altıdan büyük olduğunda madde ayırt etme parametrelerinin daha az doğru tahminleri ve standart hataları ile sonuçlandığını göstermiştir. Ayrıca, sonuçlar, çok düzeyli yapıyı göz ardı eden bir karma madde tepki modelinin, küme sayısı ve küme boyutu sırasıyla en az 50 ise, Bayes bilgi kriterine (BIC) dayalı çok düzeyli karma madde tepki modeline göre seçilmediğini göstermiştir. Ayrıca tek düzeyli verilere gereksiz yere çok düzeyli bir karma madde tepki modeli uygulamanın sonuçları incelenmiş ve güven verici bir şekilde, tek düzeyli veriler bağlamında, çok düzeyli bir karma madde tepki modeli BIC tarafından seçilmemiştir ve küme boyutu en az 20 olduğunda kullanımı düzey içi madde parametre tahminlerini veya standart hataları bozmamıştır. Sonuç olarak da KMTK modellerinde çok düzeyli bir yapının göz ardı edilmesinin, özellikle çok sayıda küme ve küme boyutu olduğunda, daha az doğru model seçimi sonuçları, madde ayırt edicilik tahminleri ve bunların SE'leri ile sonuçlandığı ve çok düzeyli doğrusal modellemede, çok

düzeyleli modellemenin gerekliliğinin tipik olarak KİK ile gerekçelendirilebileceği ifade edilmiştir.

Dras'ın (2023) tezinde amaç, ÇDKMTK 2 PL modelinin hiyerarşik olarak yapılandırılmış, iki kategorili verilerde kapsamlı bir dizi koşulda DMF belirlenmesindeki etkinliğini araştırmaktır. DMF analizi, yinelemeli (iteratif) MTK tabanlı sıralı LR tekniği kullanılarak gerçekleştirilmiş ve odak grup, çok düzeyli bir karma modelinden örtük sınıfların tahmin edilmesiyle belirlenmiştir. DMF analizi için ÇDKMTK 2 PL modelinin performansı (a) gerçek sınıf yapısını geri kazanma doğruluğu, (b) üyelik sınıflandırmasının doğruluğu ve (c) DMF maddelerini ve Tip I hata oranlarını tespit etme hassasiyeti ile değerlendirilmiştir. Önceki araştırmalarda, KMTK modellerinin DMF belirleme amacıyla uygulanmasındaki birçok faydayı vurgulanmıştır, ancak bu çalışmanın kapsamı, model parametrelerini tahmin etmek için hesaplama açısından zorlu bir Bayezyen MCMC yöntemi kullanılarak kısıtlanmıştır. Bu çalışmanın sonuçları önceki çalışmalar KMTK modelinin DMF analizi için yeterli performans gösterdiği alternatif koşulları tanımlamış olsa da bu çalışmadan elde edilen bulgular, hiyerarşik olarak yapılandırılmış eğitim değerlendirme verileri bağlamında, bu çalışmada uygulanan yöntemin yaygın eğitim değerlendirme senaryoları arasında iyi bir genelleme yapmadığını göstermektedir.

Zhang (2017) tezinde ortak değişkenin ve kestirim yönteminin KMTK çerçevesinde örtük DMF'nin belirlenmesi üzerindeki etkisini araştırmak istemektedir. DMF'nin büyüklüğü, DMF türü, DMF maddelerinin oranı, grup etkisi ve ortak değişken ile örtük grup üyeliği arasındaki ilişki manipüle edilerek bir Monte Carlo simülasyon çalışması gerçekleştirilmiştir. Üretilen cevap verileri, ortak değişkenin dahil edilmesi ve kestirim yöntemi manipüle edilerek KMTK 2PL modeli kullanılarak analiz edilmiştir. Kestirim sonuçları, örtük grup yapısının iyileştirilmesi, model parametrelerinin iyileştirilmesi ve DMF'nin belirlenmesi açısından değerlendirilmiştir. Çalışmanın temel amacı, örtük DMF'nin KMTK modellemesinde MLE (Maximum Likelihood Estimation) ve MCMC (Markov Chain Monte Carlo) kestiriminin farklılıklarını ele almak ve DMF analizinde KMTK modellerinin

kullanımına ilişkin içgörü ile öneriler sunmaktır. Elde edilen sonuçlar, DMF maddelerinin belirlenmesinin örtük grupların ayrıştırılmasına ve yargıda bulunmak için DMF hakkında ne kadar bilginin mevcut olduğuna bağlı olduğunu göstermektedir. Ayrıca DMF etki büyüklüğü fazla olduğunda veya DMF maddelerinin büyük bir yüzdesi dahil edildiğinde gücün genellikle daha yüksek olduğu bulunmuştur. Bunun yanında tek tip DMF maddelerinin tek tip olmayan DMF maddelerine kıyasla daha iyi ayırt edildiği görülmüştür.

Cohen ve Bolt (2005) çalışmalarında cinsiyet DMF'si için madde tepki örüntülerini analiz ediyor ve ardından bir KMTK modeli kullanarak, DMF ile ilişkili gözlenen özelliğin, madde(ler) tarafından gerçekte avantajlı veya dezavantajlı olan örtük gruplarla genellikle nasıl çok zayıf bir ilişkiye sahip olduğunu göstermişlerdir. Daha sonra, DMF değerlendirmesine alternatif bir yaklaşım önererek, öncelikle DMF'ye katkıda bulunan birincil boyut(lar)ı tanımlamak için keşifsel bir karma model analizi kullanılmasını ve ikinci olarak da DMF'nin neden(ler)ini anlamak için bu boyutlarla ilişkili sınav katılımcısı özelliklerinin incelenmesini önermişlerdir. İlk çalışma için üniversite düzeyinde bir matematik seviye belirleme sınavının bir bölümündeki çoktan seçmeli maddeler kullanılmıştır. Her biri 500 erkek ve 500 kadından oluşan 1000 kişilik iki örneklem 5249 öğrenciden oluşan bir örneklemden rastgele seçilmiştir. Örneklem 1'den alınan yanıtlar DMF belirlenmesi ve madde parametrelerinin kestirimi için kullanılmıştır. Örneklem 2'den alınan yanıtlar örtük sınıf üyeliğinin kestirimi için kullanılmıştır. DMF için bir olabilirlik oranı testi DMF'yi her seferinde bir madde belirlemek için kullanılmıştır. İlk çalışmaya ait sonuçlar ile DMF'yi incelemeye yönelik alışlagelmiş yaklaşımın DMF'nin nedenlerini anlamaya fazla katkıda bulunmadığı gösterilmiştir. Grupların cinsiyete dayalı bir DMF analizine göre tanımlanmış olmasına rağmen, maddelerin farklı işlev gördüğü denekler cinsiyetlerine göre doğru bir şekilde tanımlanmamış olduğu görülmüştür. İkinci çalışmaya ait sonuçlar DMF analizini, karma modellerin ilk olarak bir tür DMF'nin mevcut olabileceği örtük grupları tanımlamak için nasıl kullanılabileceğini göstermek üzere genişletmiştir. Çalışma 2'deki gruplar söz konusu olduğunda, ikincil boyutun bilişsel bir yorumu vardır ve dolayısıyla

DMF'nin nedeni daha kolay yorumlanabilir. Sonuçlara göre karma modellerin bir avantajı, potansiyel olarak önemli sınav katılımcısı farklılıklarının belirlenmesine yardımcı olabilmesi ve böylece DMF'nin farklı örtük sınav katılımcısı alt popülasyonlarının yanıtlarını nasıl etkilediğine dair daha fazla içgörü sağlamasıdır.

İlgili araştırmalar incelendiğinde çoğunluğun simülatif veri ile gerçekleştiği görülmektedir. Araştırmaların geneli düzeylilik, boyutluluk ve ortak değişkenlerin analize dahil olması durumlarından bir tanesini incelemiştir. Bunun yanında çalışmaların çoğunluğu sadece Rasch modele göre analizlerini yapmışlardır. Alanda yapılanları geniş bir yelpazede bir araya getirip genişletecek bu çalışmada amaç cinsiyet ve evdeki kitap sayısı değişkenleri ile, karma modele dayalı KMTK, ÇDKMTK, ÇBÇDKMTK ve odÇBÇDKMTK modellerinin DMF'li madde belirleme ve DMF'nin olası kaynaklarını açıklayabilme performanslarını karşılaştırma olarak belirlenmiştir. Elde edilen sonuçların özellikle testlerde adilliği sağlama ve geçerlik kanıtlarını artırma bakımından literatüre katkıda bulunacağı düşünülmektedir. Ayrıca 54 farklı karma modelin yer aldığı araştırma ile bu alanda çalışma yapacak araştırmacılara DMF belirleme çalışmalarında ölçme aracının yapısına uygun DMF belirleme yönteminin seçiminin çalışmaların kalitesini artıracığı gerçeği göz önüne serilecektir. Son olarak da DMF belirleme sürecinin sadece madde işaretleme olmadığını ve DMF'ye sebep olan etkenleri bulmanın çok önemli olduğu düşüncesinin yer etmesi beklenmektedir.

Bölüm 3

Yöntem

Bu bölümde; araştırmanın türü, araştırmanın örnekleme, veri toplama süreci, veri toplama aracı ve verilerin analizi ile ilgili kısımlara yer verilmiştir.

Araştırmanın Türü

Boyutluluk, düzey ve ortak değişkenlerin dahil edildiği KMTK modellerinin kendi aralarında ve gözlenen grup değişkenlerine göre DMF analizi sonuçlarının karşılaştırıldığı bu çalışma betimsel araştırma türündedir. Betimleyici çalışmalar, belirli bir durumu mümkün olduğunca eksiksiz ve dikkatli bir şekilde tanımlar. (Fraenkel ve diğerleri, 2012).

Araştırmanın Örnekleme

Örneklem sayılarının araştırmalar kapsamında kullanılan gerçek ve simülatif verilere göre farklılaştığı görülmüş ve bu bilgiler ışığında örneklem sayısına karar verilmiştir. Tofighi ve Enders (2008) ortak değişkenlerin dahil edilmesinin olumsuz etkisinin örneklem sayısının 1000 olduğunda azalmaya başlamasının tesadüf olmayabileceğini belirtirlerken ortak değişkenlerin sınıf belirleme üzerindeki etkisinin örneklem büyüklüğüne bağlı olduğunu ve ortak değişkenleri dahil etmenin bazı faydaları olduğunu ifade etmişlerdir. Cho (2007) simülasyon çalışmasında 8000, gerçek veride 8362 kişilik örneklem üzerinde çalışırken, Finch ve Finch (2013) gerçek veri üzerinde 2553 kişilik örneklem, Karadavut vd. (2019) gerçek veri üzerinde 1372 kişilik örneklem, Li vd. simülasyon çalışmasında 2000 kişilik örneklem ve Dai (2009) simülasyon çalışmasında 1000 kişilik örneklem ile çalışmıştır. Maij-de Meij vd.'nin (2010) 1000, 5000 ve 25000 kişilik örneklem kullandıkları simülasyon ve 1917 kişilik örneklemden oluşan gerçek veri üzerindeki çalışmaları sonucunda örneklem büyüklüğü 1000 kişiye düşürüldüğünde, KMTK modeli madde parametre tahminlerinin varyans/kovaryans matrisini doğru bir şekilde tahmin edemeyeceği belirtilmiş ve ampirik örneğin gözlenen bir gösterge içeren örtük DMF belirleme yönteminin 2000 katılımcıdan oluşan örneklere uygulanabileceği gösterilmiştir. Diğer taraftan araştırmalardaki

örneklemelere ait büyüklük arttıkça modellerin kestirim süreleri uzamaktadır. Bu bilgiler ışığında örneklem sayısının kestirim süresini aşırı uzatmayacak kadar fazla olmasının hem model parametre tahminlerinin daha doğru tahmin edilmesini sağlayacağı hem de zamanı ekonomik kullanmaya katkı sunacağı Cho'nun (2007) çalışmasında olduğu gibi 8000 civarında olması istenmiştir. Araştırma örneklemini eTIMSS 2019 uygulamasındaki 1. kitapçığı cevaplayan 8106 öğrenci oluşturmuştur. Ancak araştırma kapsamında gözlenen gruplar için yapılacak analizde cinsiyet değişkeni kullanıldığı için cinsiyetini belirtmeyen 119 öğrenci analiz kapsamına alınmamış ve analizler 7987 öğrencinin cevapları kullanılarak yapılmıştır.

Veri Toplama Süreci

Misyonu, eğitim sistemleri içinde ve genelinde kaliteli öğretim ve öğrenimi geliştirmek için eğitim uygulamalarını, süreçlerini ve politikalarını daha iyi anlamak olan bunu geniş ölçekli çalışmalar yürüterek yapan Uluslararası Eğitim Başarılarını Değerlendirme Derneği (IEA) tarafından hazırlanıp uygulanan TIMSS uygulamasına ait 2019 yılı verilerine <https://www.iea.nl/> adresinden ulaşılmıştır.

Veri Toplama Aracı

TIMSS (Uluslararası Matematik ve Fen Eğilimleri Araştırması) dünya genelinde matematik ve fen bilimlerinde öğrenci başarısını ölçen uluslararası karşılaştırmalı bir çalışmadır. 1995'ten bu yana dört yıllık bir değerlendirme döngüsünde yürütülen TIMSS, dördüncü ve sekizinci sınıflarda öğrenci başarısını 1995, 1999, 2003, 2007, 2011, 2015, 2019 ve 2023 olmak üzere sekiz kez değerlendirmiş ve 24 yıllık eğilim ölçümleri (TIMSS 2023 uygulaması sonuçları henüz yayınlanmadığı için) biriktirmiştir. TIMSS, öğrencilerin matematik ve fen öğrendikleri ulusal, ev, okul ve sınıf bağlamları hakkında zengin bir dizi bilgi toplamaktadır. Bu bağlam verileri, matematik ve fen başarısı ile ilgili eğitim faktörleri hakkında uluslararası karşılaştırmalı perspektifler sağlamaktadır. TIMSS 2019 uygulamasına ait veriler içerisinde eTIMSS verileri tercih edilmiştir.

2019 yılında TIMSS, "eTIMSS" adı verilen kağıt-kalem değerlendirmesinin dijital bir versiyonunu sunarak bilgisayar tabanlı değerlendirmeye geçmeye başlamıştır. Dijital değerlendirmeye geçiş, "çağa ayak uydurmak" ve hem yapı temsilini hem de veri kullanımını artırmak için önemlidir. Tüm TIMSS ülkeleri dijital değerlendirme yapmaya hazır olmadığından, IEA geçişi iki değerlendirme döngüsü (TIMSS 2019 ve TIMSS 2023) boyunca uygulamaya karar vermiştir. TIMSS 2019'a katılan 64 ülkenin yarısından fazlası değerlendirmelerin "e" versiyonunu uygulamayı tercih ederken, geri kalan ülkeler TIMSS'i önceki değerlendirme döngülerinde olduğu gibi kağıt-kalem formatında uygulamıştır (paperTIMSS).

eTIMSS 2019 değerlendirmeleri, renkli grafikler ve etkileşimli özellikler içeren çeşitli teknoloji ile geliştirilmiş madde formatları içermektedir. Bu özellikler matematik ve fen çerçevelerinin kapsamını genişletmiş ve öğrenci katılımını teşvik etmiştir. Dijital uygulama modu, eTIMSS matematik sorularının önemli bir kısmının makine tarafından puanlanmasına da olanak sağlamıştır. eTIMSS 2019 değerlendirmeleri, IEA'nın araç oluşturma, çeviri ve uyarlama, doğrulama, öğrencilere ulaştırma, puanlama ve veri girişi için kullanılan çevrimiçi araçlar koleksiyonunu barındıran eAssessment System kullanılarak oluşturulmuş ve uygulanmıştır. Dijital uygulama modu, eTIMSS'in ekran ekran zamanlama verileri ve öğrencilerin başarı maddeleriyle etkileşimlerini incelemek için analiz edilebilecek ek süreç değişkenleri gibi öğrencilerin maddeler üzerinde nasıl çalıştıkları hakkında bilgi toplamasına olanak sağlamıştır (Fishbein ve diğerleri, 2021).

eTIMSS 2019 8. sınıf uygulamasına 22 ülke ve beş kıyaslama katılımcısı olan ülke veya şehir katılmıştır. Uygulamada 14 matematik ve 14 fen bilimleri dersi maddeleri ile meydana gelen toplam 28 tane blok bulunmaktadır. 28 bloğun farklı kombinasyonları ile 14 öğrenci kitapçığı oluşturulmuştur. Öğrenci kitapçıkları, iki matematik bloğu ve iki fen bilimleri bloğu olacak şekilde toplamda dört bloktan meydana gelmektedir (Mullis ve diğerleri, 2020). Ayrıca her kitapçıkta çoktan seçmeli (2-3-4 seçenek içeren), kısa cevaplı ve açık uçlu maddeler bulunmaktadır ve kitapçıklardaki madde sayıları ve türleri farklılaşabilmektedir.

Matematik alanında sayılar, cebir, geometri ve veri ile olasılık konuları yer alırken fen bilimleri alanında fizik, kimya, biyoloji ve yeryüzü bilimleri yer almaktadır. Bilişsel alan olarak ise bilme (knowing), uygulama (applying) ve akıl yürütme (reasoning) olmak üzere üç kategori belirlenmiştir. Araştırma kapsamında eTIMSS 2019 8. sınıflara ait 1. kitapçığın verileri kullanılmıştır. 1. kitapçıkta matematik ve fen bilimleri alanlarına ait 44'er madde yer almaktadır. Ancak dört seçenekli çoktan seçmeli kimya alanında bilme bilişsel alanında yer alan bir maddeye ait veri bulunmamaktadır. Toplam 87 maddeye ait madde türü, konu alanı ve bilişsel alanlara ait bilgiler Tablo 1'de verilmiştir.

Tablo 1

Kitapçık 1 için Madde Türü, Konu Alanı ve Bilişsel Alanlara Ait Bilgiler

Alan	Kitapçık 1	Matematik	Fen Bilimleri
Madde Türü	2 seçenekli çoktan seçmeli	5	9
	3 seçenekli çoktan seçmeli	3	3
	4 seçenekli çoktan seçmeli	18	16
	Açık uçlu	18	15
Konu Alanı	Sayılar/ Fizik	14	15
	Cebir/ Kimya	10	8
	Veri ve olasılık/ Biyoloji	9	14
	Geometri/ Yeryüzü Bilimi	11	6
Bilişsel Alan	Bilme	13	16
	Uygulama	20	14
	Akıl Yürütme	11	13

Tablo 1'e göre matematik bölümünde en fazla açık uçlu ve dört seçenekli madde yer alırken en az üç seçenekli madde yer almaktadır. Konu alanı olarak en fazla sayılar konusundan maddenin yer aldığı bölümde uygulama bilişsel alanına ait 20 madde bulunmaktadır. Fen bilimleri bölümünde ise en fazla dört seçenekli madde yer alırken en az

üç seçenekli madde yer almaktadır. Konu alanı olarak en fazla biyolojiye ait konuların yer aldığı bölümde bilme bilişsel alanına ait 16 madde bulunmaktadır. Ayrıca araştırma kapsamında çok boyutlu ve çok düzeyli olan veriye ortak değişken olarak eklenen evdeki kitap sayısı değişkeni eklenmiştir ve bu değişken beş kategorilidir.

Verilerin Analizi

Veri analizine başlamadan önce veri toplama aracına ilişkin boyutluluk ve düzeylilik açısından incelemeler yapılarak araştırmanın amacına uygun olacak biçimde araca son şekli verilmiştir. Ardından veri toplama aracına ve örnekleme ilişkin betimsel istatistikler hesaplanmıştır. Üçüncü olarak veri setine en uygun modeller belirlenmiştir. Son olarak da seçilen modellere ait betimsel istatistikler ve parametre tahminleri oluşturulmuştur.

1. Boyut ve Düzey İncelemesi

eTIMSS 2019 8. sınıf uygulaması 1. kitapçığında dört seçenekli çoktan seçmeli matematik ve fen bilimleri derslerine ait toplam 34 madde bulunmaktadır ve çalışma kapsamında bu kitapçığa ait verilerin kullanılması planlanmıştır. Ancak araştırma kapsamında boyutluluk ve düzeylilik şartlarının sağlanıp sağlanmadığının incelemesi yapılmış ve 14 maddenin araştırma verisinden çıkarılmasına karar verilmiştir. Maddelerin çıkarılma sebepleri çok boyutluluğu oluşturmak ve çok düzeyli veri yapısına uygun olan maddeler ile analizleri yapmaktır. Buna göre veriler 10 matematik ve 10 fen bilimleri olmak üzere 20 maddeden oluşmaktadır.

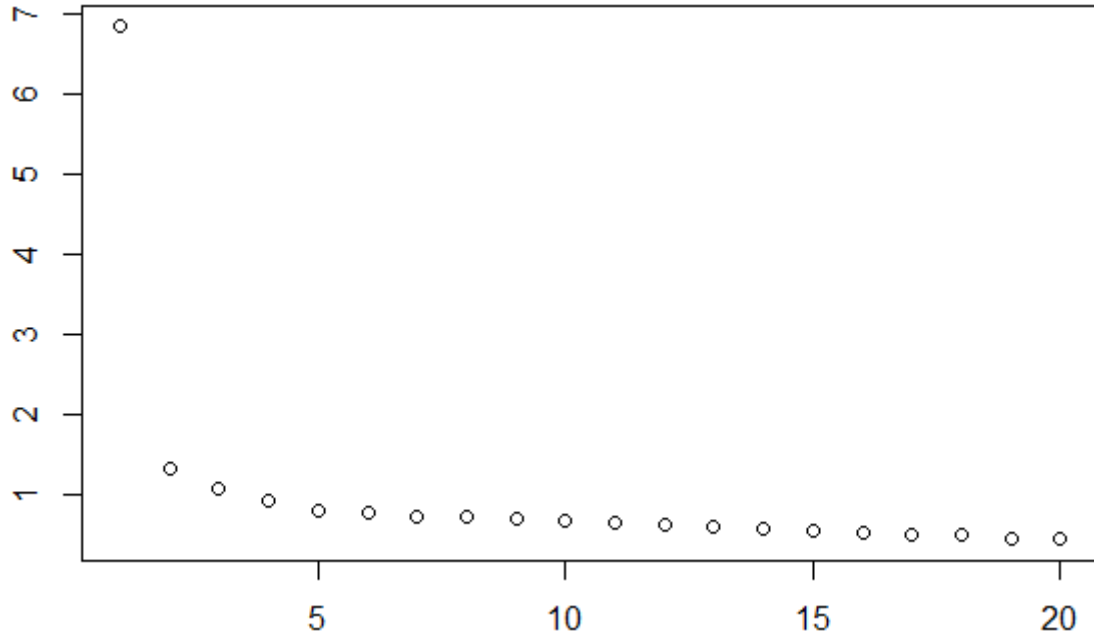
Maddelerin boyutluluk analizleri için açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) yapılmıştır. Araştırmada kapsamında yer alan maddeler 0-1 olarak puanlandıkları için boyutluluk incelemesinde ilk olarak tetrakorik korelasyon matrisine dayalı AFA kullanılmıştır. AFA öncesi veri yapısının analize uygun olup olmadığını belirlemek için Kaiser-Meyer-Olkin (KMO) testi ile Bartlett Küresellik testi uygulanmıştır. Ölçülecek yapının faktör analizine uygun halde olabilmesi için KMO değerinin 0,60 veya daha büyük değere sahip olması, Bartlett testinin ise anlamlı ($p<.05$) olması gerekir (Tabachnick & Fidell, 2013).

Analiz sonucunda KMO deęerinin (0,93) örnekleme yeterlięi üst seviyede ve Bartlett Küresellik testi sonucunun (0,00) anlamlı olduęu bulunmuştur. Bu yüzden veri yapısının faktörleşmeye uygun olduęu söylenebilir. AFA için R yazılımına ait “psych” paketinde (Revelle, 2023) bulunan ‘fa’ fonksiyonu kullanılmış ve özdeęerler hesaplanmıştır. Buna göre özdeęeri 1’den büyük olan ilk üç faktöre ait özdeęerler sırasıyla 6,84, 1,31 ve 1,06’dır. Bir veride faktör sayısına karar vermek için çeşitli alternatifler vardır. Bunlardan birisi özdeęerler aracılıęıyla karar vermektir. Kaiser (1960) kriteri olarak bilinen yöntemde göre özdeęerin 1’den büyük olması faktör sayısı için kriterdir. Buna göre incelendięinde faktör sayısının üç olduęu söylenebilir. Ancak birinci ve ikinci özdeęerin oranının yaklaşık beş olması (Lord, 1980) yapının tek boyutlu olabileceęini gösterir.

Faktör sayısı belirlemek için sık kullanılan ikinci yöntem Cattell (1966) tarafından geliştirilen yamaç birikinti grafięidir. Bu yöntemde göre dikey eksenden (özdeęer) yatay eksene (faktör sayısı) doğru eğimli bir düşüş olur fakat bir noktadan itibaren eğim gitgide azalır ve grafik düz bir biçimde devam eder. Bahsedilen nokta kırılma noktası olarak adlandırılır ve bu noktaya kadar olan sayı faktör sayısı olarak kabul edilir. 20 maddeye ilişkin hesaplanan özdeęerlere göre çizilen yamaç birikinti grafięi Şekil 3’te yer almaktadır.

Şekil 3

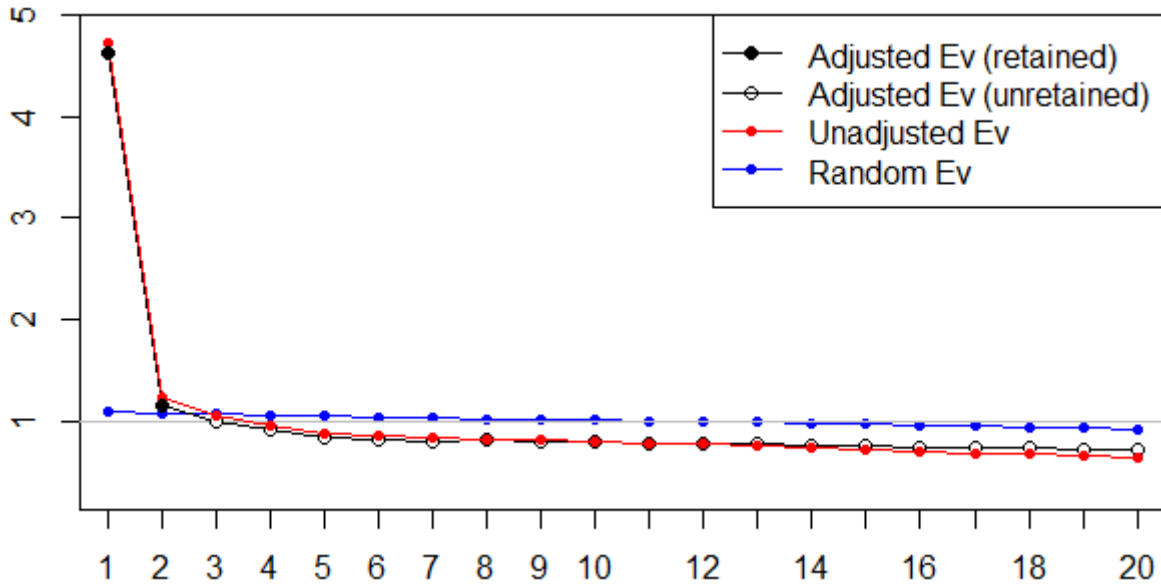
Veri Setini Oluşturan 20 Madde İçin Yamaç Birikinti Grafięi



Şekil 3'te ilk faktörden ikinci faktöre doğru keskin bir düşüş olduğu ve diğer faktörlerin ikinci faktöre yakın değerler aldıkları ve yatay bir seyirde gittikleri görülmektedir. Bu görünüm tek boyutluluğun varlığını göstermektedir. Kaiser kriteri ve yamaç birikinti grafiği tek boyutlu bir yapıyı göstermektedir fakat bu yöntemlerin güvenilir olmayabileceğini belirten çalışmalar (Koçak ve diğerleri, 2016; Revelle & Rocklin, 1979; Tucker ve diğerleri, 1969) mevcuttur. Bu sebeple özellikle karşılaştırma içeren çalışmalarda daha kararlı olup ve doğru sayıda faktör tespit ettiği belirtilen (Crawford ve diğerleri, 2010; Koçak ve diğerleri, 2016) Horn'un paralel analiz yöntemi ile de inceleme yapılmıştır. Horn'un (1965) geliştirdiği yöntemde gerçek veri setine paralel rastgele veri setleri oluşturulur. Rastgele üretilen veriye ait özdeğerin gerçek veriye ait özdeğerden büyük olduğu nokta faktör sayısını verir. R kütüphanesindeki "paran" paketi (Dinno, 2018) ile yapılan analize ait görsel Şekil 4'te gösterilmektedir.

Şekil 4

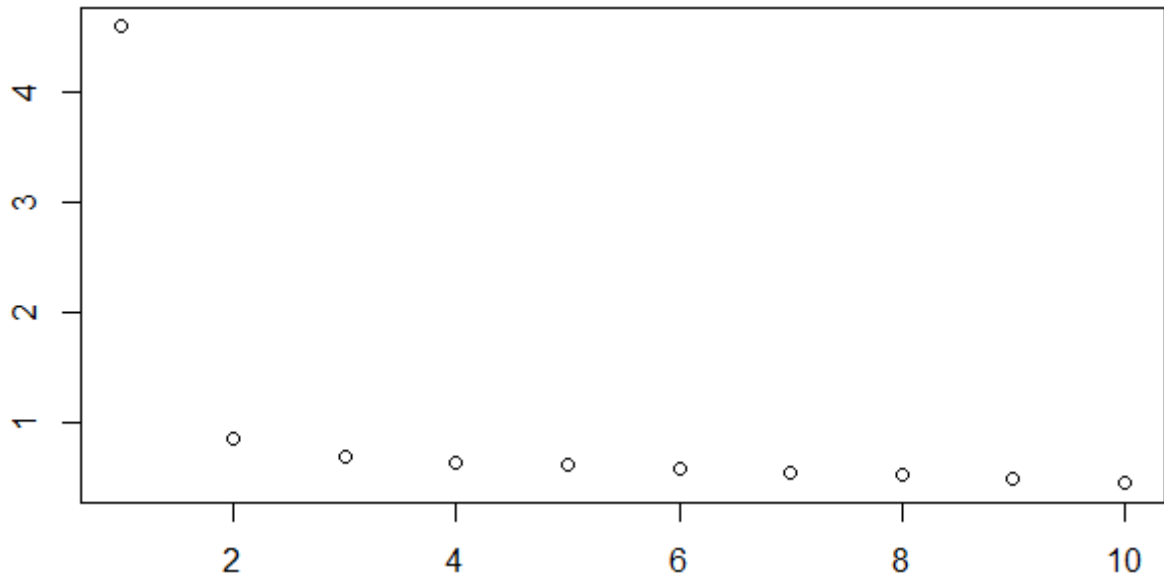
20 Maddelik Veri Setinde Boyutluluk Analizi Kapsamında Paralel Analiz Sonuç Görseli



Şekil 4'te yer alan düzeltilmiş özdeğer (korunmuş) ve düzeltilmemiş özdeğerlerden ikisi rastgele üretilen özdeğerlerden fazladır. Buna göre verinin faktör sayısının iki olduğu görülmektedir. Bu durum önceki iki yöntem ile çelişen bir sonuç ortaya çıkarmıştır. Bu sebeple seçilen 20 maddenin matematik ve fen bilimleri derslerine ait 10'ar maddesi ayrı ayrı birer yapı olarak ele alınıp incelenmiştir. Ayrıca hem 20 madde için hem de 10 matematik ve 10 fen bilimleri maddeleri için DFA yapılacak ve model uyum indekslerinin sonuçlarına göre en uygun yol izlenmiştir. Öncelikle 10 matematik maddesi için AFA öncesi veri yapısının analize uygun olup olmadığını belirlemek için Kaiser-Meyer-Olkin (KMO) testi ile Bartlett Küresellik testi uygulanmış ve KMO değerinin (0,89) örnekleme yeterliğinin üst seviyede, Bartlett Küresellik testi sonucunun (0,00) ise anlamlı olduğu bulunmuştur. Bu sebeple veri yapısının faktörleşmeye uygun olduğu söylenebilir. AFA için R yazılımına ait "psych" paketinde (Revelle, 2023) bulunan 'fa' fonksiyonu kullanılmış ve özdeğerler hesaplanmıştır. 10 matematik maddesine ait özdeğeri 1'den büyük olan sadece bir faktör vardır ve özdeğeri 4,60'tır. Kaiser kriterine göre özdeğeri 1'i (bir) geçen sadece bir faktör olduğu için veri setinin tek boyutlu olduğu görülmektedir. 10 maddelik matematik veri setine ilişkin hesaplanan özdeğerlere göre çizilen yamaç birikinti grafiği Şekil 5'te yer almaktadır.

Şekil 5

10 Maddelik Matematik Veri Seti İçin Yamaç Birikinti Grafiği

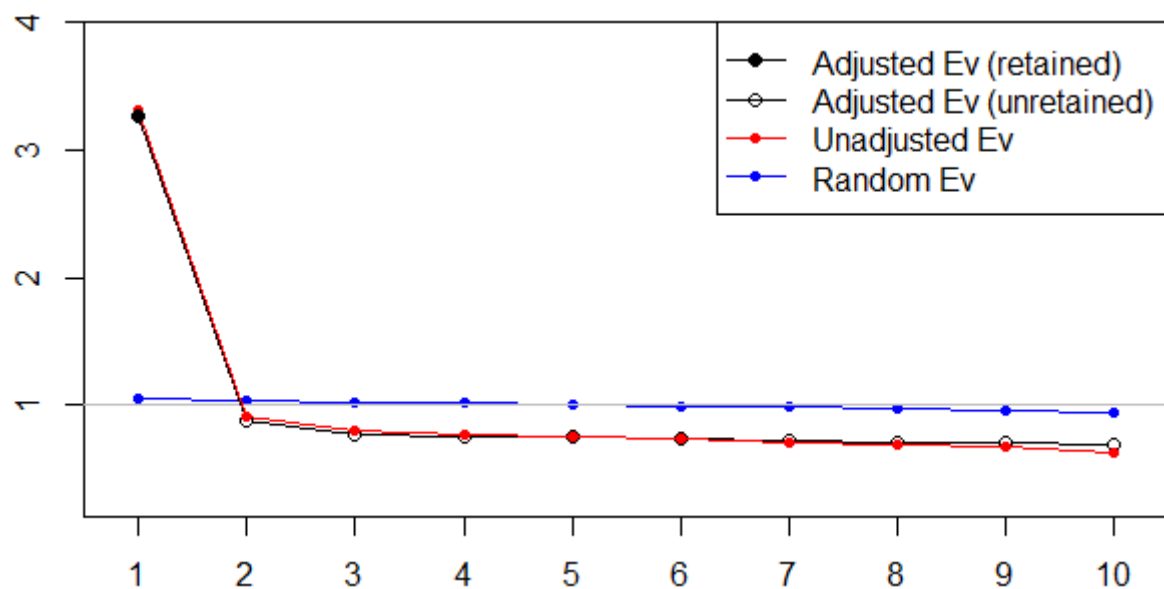


Şekil 5'te ilk faktörden ikinci faktöre doğru keskin bir düşüş olduğu ve diğer faktörlerin ikinci faktöre yakın değerler aldıkları ve yatay bir seyirde gittikleri görülmektedir. Bu görünüm tek boyutluluğun varlığını göstermektedir. Kaiser kriteri ve yamaç birikinti grafiğinin ardından Horn'un paralel analiz yöntemi uygulanmış ve sonuç Şekil 6'da gösterilmiştir.

Şekil 6

10 Maddelik Matematik Veri Setinde Boyutluluk Analizi Kapsamındaki Paralel Analiz

Sonuç Görseli

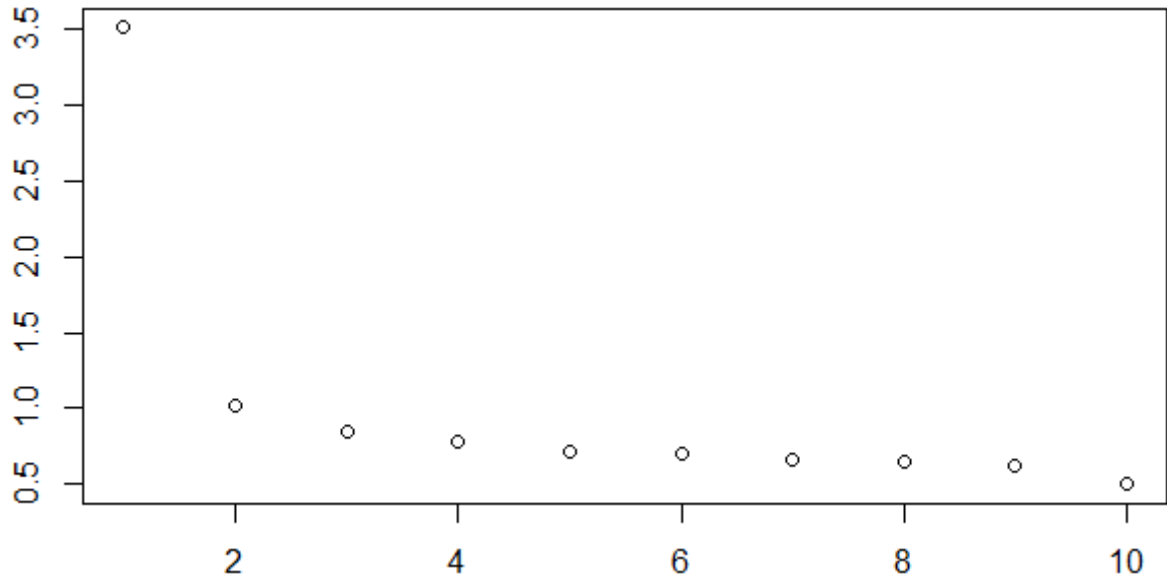


Şekil 6'da yer alan düzeltilmiş özdeğer (korunmuş) ve düzeltilmemiş özdeğerlerden ilki rastgele üretilen özdeğerlerden fazladır. Buna göre verinin faktör sayısının bir olduğu görülmektedir. Yapılan tüm boyutluluk analizlerinde 10 matematik maddesinin tek boyutlu olduğu görülmüştür.

Fen bilimleri dersine ait 10 maddelik veri setine için de matematik dersine ait 10 maddelik veri setine yapılan işlemler yapılmıştır. Buna göre öncelikle 10 fen bilimleri maddesi için AFA öncesi veri yapısının analize uygun olup olmadığını belirlemek için Kaiser-Meyer-Olkin (KMO) testi ile Barlett Küresellik testi uygulanmış ve KMO değerinin (0,83) örnekleme yeterliğinin üst seviyede, Bartlett Küresellik testi sonucunun (0,00) ise anlamlı olduğu bulunmuştur. Bu sebeple veri yapısının faktörleşmeye uygun olduğu söylenebilir. AFA için R yazılımına ait "psych" paketinde (Revelle, 2023) bulunan '*fa*' fonksiyonu kullanılmış ve özdeğerler hesaplanmıştır. 10 fen bilimleri maddesine ait özdeğeri 1'den büyük olan ik faktör vardır ve özdeğerleri sırasıyla 3,51 ve 1,01'dir. Kaiser kriterine göre özdeğeri 1'i (bir) geçen iki faktör olduğu için veri setinin iki boyutlu olduğu görülmektedir. Ancak birinci ve ikinci özdeğerin oranının yaklaşık dört olması (Lord, 1980) yapının tek boyutlu olabileceğini gösterir. 10 maddelik fen bilimleri veri setine ilişkin hesaplanan özdeğerlere göre çizilen yamaç birikinti grafiği Şekil 7'de yer almaktadır.

Şekil 7

10 Maddelik Fen Bilimleri Veri Seti İçin Yamaç Birikinti Grafiği

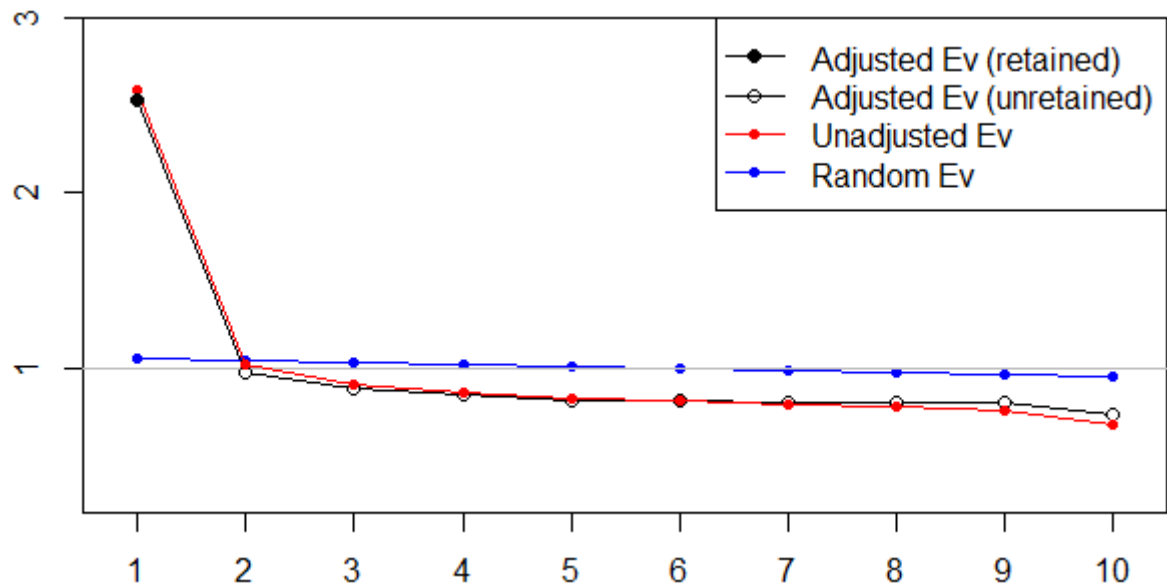


Şekil 7’de ilk faktörden ikinci faktöre doğru keskin bir düşüş olduğu ve diğer faktörlerin ikinci faktöre yakın değerler aldıkları ve yatay bir seyirde gittikleri görülmektedir. Bu görünüm tek boyutluluğun varlığını göstermektedir. Kaiser kriteri ve yamaç birikinti grafiğinin ardından Horn’un paralel analiz yöntemi uygulanmış ve sonuç Şekil 8’de gösterilmiştir.

Şekil 8

10 Maddelik Fen Bilimleri Veri Setinde Boyutluluk Analizi Kapsamında Paralel Analiz

Sonuç Görseli



Şekil 8'de yer alan düzeltilmiş özdeğer (korunmuş) ve düzeltilmemiş özdeğerlerden ilki rastgele üretilen özdeğerlerden fazladır. Buna göre verinin faktör sayısının bir olduğu görülmektedir. Yapılan tüm boyutluluk analizlerinde 10 fen bilimleri maddesinin tek boyutlu olduğu görülmüştür.

AFA sonrası 20 maddelik veri setinin tek boyutlu olması ve 10'ar maddelik iki farklı boyuttan oluşmuş olması durumlarının hangisinin daha doğru olduğunu belirlemek için DFA yapılmıştır. DFA sonucu elde edilen model uyum indeksleri aşağıda yer almaktadır. İlk olarak tek boyutlu bir yapı için elde edilen sonuçlar Tablo 2'de verilmiştir.

Tablo 2

20 Maddelik Veri Setinin Tek ve İki Boyutlu Yapı Sergilemesi Durumunda DFA Sonuçları

İndeks	Değer	
	Tek boyut	İki boyut
CFI	0,93	0,96
TLI (NNFI)	0,92	0,96
NFI	0,92	0,96
BIC	206331,47	205604,79
RMSEA	0,03	0,02
RMR	0,007	0,005
GFI	0,98	0,99
AGFI	0,97	0,98

CFI, GFI, AGFI, TLI ve NFI indeksleri uyum indeksleridir ve $\geq 0,90$ iyi uyumu, $\geq 0,95$ ise mükemmel uyumu gösterir (Hooper ve diğerleri, 2008). Buna göre tek boyutlu verinin modele iyi uyum sağladığı söylenebilir. İki boyutlu yapı için elde edilen uyum indekslerine ait değerler mükemmel uyumu göstermektedir. RMSEA ve RMR değerleri hata indeksleridir ve $\leq 0,05$ mükemmel uyumu, $\leq 0,08$ iyi uyumu gösterir (Jöreskog & Sörbom, 1993). Tek boyutlu veri hata indeksleri için mükemmel uyum sağlanmıştır. İki boyutlu veri için hata

indekslerine ait değerler de mükemmel uyumu göstermektedir. BIC değerinin ise küçük olması model veri uyumunun daha iyi olduğunu gösterir (de Ayala, 2009). Tek ve iki boyutlu modeller için model veri uyumunu test etmek amacıyla R kütüphanesinde yer alan “lavaan” (Rosseel, 2012) paketinden yararlanılmıştır. İki boyutlu yapı için elde edilen BIC değeri (205604,794) tek boyutlu yapı için elde edilen değerden (206331,47) daha küçüktür ve bu da model veri uyumunun daha iyi olduğunu göstermektedir. Ayrıca faktör sayısını belirlemek için kullanılan yöntemlerden Kaiser ve yamaç eğim yöntemlerine göre daha üstün olduğu belirtilen (Crawford ve diğerleri, 2010; Piccone, 2009) paralel analizin sonuçları iki boyutlu yapıyı desteklemektedir. Buna göre 20 maddeden oluşan veri setinin matematik ve fen bilimleri olmak üzere iki boyutlu yapıda olduğu söylenebilir.

Veri setinin çok düzeyli bir yapı sergileyip sergilemediğini belirlemek için R kütüphanesinde yer alan “misty” (Yanagida, 2023) paketindeki “icc” fonksiyonu ile her bir maddeye ait KİK değerleri elde edilmiştir. 20 maddeye ait KİK değerleri 0,05 ve 0,15 aralığında iken ve ortalama KİK değeri 0,10 olarak bulunmuştur. Bu sonuç ülke düzeyinin toplam varyansın %10’unu açıklayabileceğini göstermektedir. KİK değerinin 0 (sıfır) olması verinin bağımsız olduğu ve üst düzeylerde herhangi bir varyantlaşmanın olmadığını gösterir. Değer 1’e yaklaştıkça üst düzeyde değişimin çok olduğuna işaret eder. Muthen (1997) gruptaki birey sayısı 15’ten fazla olduğunda KİK değerinin 0,10’dan büyük olduğunda çok düzeyli modellemenin kesinlikle dikkate alınması gerektiğini, Julian (2001) ile Selig vd. (2008) KİK değerlerinin 0,10’dan düşük değerlerinde bile hiyerarşik yapının göz ardı edilmemesi gerektiğini önermişlerdir (aktaran Şen, 2022). Hox (2013) KİK değerinin 0,05’in altında olması durumunda ise ikinci düzey etkisinin düşük olduğunu belirtmiştir. Buna göre veri setinin çok düzeyli yapıya sahip olduğu görülmektedir.

2. Veri Setine Ait Betimsel İstatistikler

eTIMSS 2019 kapsamında yukarıda bahsedilen beş kıyaslama katılımcısı olan şehir veya ülke araştırma kapsamına alınmamış ve 22 ülkeye ait verilerle işlemler gerçekleştirilmiştir. Tablo 3’te 20 maddeye ait konu ve bilişsel alanlar gösterilmektedir.

Tablo 3*Matematik ile Fen Bilimleri Verilerine Ait 20 Madde İçin Konu Alanları ve Bilişsel Alanlar*

Madde Numarası	Alan Adı	Konu Alanı	Bilişsel Alanı
1	Matematik	Sayılar	Bilme
2	Matematik	Sayılar	Akıl Yürütme
3	Matematik	Cebir	Uygulama
4	Matematik	Cebir	Bilme
5	Matematik	Geometri	Uygulama
6	Matematik	Geometri	Uygulama
7	Matematik	Veri ve Olasılık	Uygulama
8	Matematik	Sayılar	Uygulama
9	Matematik	Cebir	Bilme
10	Matematik	Cebir	Bilme
11	Fen Bilimleri	Biyoloji	Bilme
12	Fen Bilimleri	Biyoloji	Bilme
13	Fen Bilimleri	Fizik	Bilme
14	Fen Bilimleri	Fizik	Bilme
15	Fen Bilimleri	Yeryüzü Bilimleri	Bilme
16	Fen Bilimleri	Biyoloji	Bilme
17	Fen Bilimleri	Biyoloji	Bilme
18	Fen Bilimleri	Kimya	Bilme
19	Fen Bilimleri	Yeryüzü Bilimleri	Akıl Yürütme
20	Fen Bilimleri	Yeryüzü Bilimleri	Uygulama

Tablo 3'e göre matematik bölümünde cebir konu alanından dört, sayılar konu alanından üç, geometri konu alanından iki ve veri ile olasılık konu alanından bir madde vardır. Fen bilimleri bölümünde ise biyoloji konu alanından dört, yeryüzü bilimleri konu alanından üç, fizik konu alanından iki ve kimya konu alanından bir madde vardır. eTIMSS

2019 katılımcısı 22 ülkeye ait katılımcı sayısı, 20 maddeden elde edilen ortalama puan ve katılımcıların cinsiyetleri Tablo 4'te gösterilmiştir.

Tablo 4

eTIMSS 2019 Katılımcısı 22 Ülkeye Ait Betimsel İstatistikler

Ülke	Katılımcı Sayısı	Ortalama Puan	Cinsiyet(K/E)
Birleşik Arap Emirlikleri (BAE)	1573	9,05	785/788
Şili	290	7,63	155/135
İngiltere	235	10,59	129/106
Finlandiya	342	10,89	156/186
Fransa	270	7,51	125/145
Gürcistan	242	7,85	115/127
Hong Kong	228	11,98	99/129
Macaristan	320	10,98	156/164
İsrail	261	9,87	137/124
İtalya	253	9,89	130/123
Kore	272	13,99	136/136
Litvanya	255	9,88	116/139
Malezya	498	10,09	248/250
Norveç	296	9,33	148/148
Portekiz	237	9,87	133/104
Katar	272	8,84	136/136
Rusya	277	11,68	132/145
Singapur	350	15,47	186/164
İsveç	283	10,23	140/143
Türkiye	284	10,35	156/128
Tayvan	375	14,90	174/171
Amerika Birleşik Devletleri (ABD)	604	10,58	272/332
Genel	7987	10,37	3964/4023

Tablo 4'e göre en çok katılımcının olduğu ülke BAE olurken en az katılımcı Hong Kong'tandır. Ülkelerdeki katılımcı cinsiyetlerine bakıldığında kız ve erkek sayıları birbirine yakın değerdedir. 20 maddelik veri setine ait merkezi eğilim ve yayılım ölçülerini içeren çarpıklık ve basıklığın raporlandığı betimsel istatistikler Tablo 5'teki gibidir.

Tablo 5

Veri Setine Ait Betimsel İstatistikler

Aritmetik Ortalama	10,37
Medyan	10
Mod	7
Standart Sapma	4,75
Varyans	22,61
Çarpıklık	0,18
Basıklık	-0,92
En küçük değer	0
En Büyük değer	20
Ranj	20

Tablo 5'te yer alan merkezi eğilim ölçüleri olan aritmetik ortalama, medyan ve mod incelendiğinde aritmetik ortalama ve medyanın moddan daha büyük ve birbirine yakın değerlere sahip oldukları görülmektedir. Tablo 5'e göre çarpıklık ve basıklık değerlerinin -1 ile +1 aralığında değiştiği görülmektedir. Tabachnick ve Fidell'e (2013) göre -1,5 ile +1,5 arasındaki bu değerler kabul edilebilir aralıktadır. Bu durumda puanların normalden aşırı bir sapma göstermediği söylenebilir.

Veri setine ait güvenilirlik katsayısının belirlenmesi için veri setinin çok boyutlu ve aşağıda belirtildiği üzere çok düzeyli yapısı dikkate alınarak birden fazla güvenilirlik belirleme yöntemi kullanılmıştır. Bu yöntemlerden ikisi veri setinin hiyerarşik yapısını dikkate alan McDonald'ın ω_h ve ω_t değerleridir (McDonald, 1999; Revelle & Zinbarg, 2009). McDonald (1999), bir testin genel faktör doyumunun bir tahmini olarak omega (hijerarşik) katsayısını

(ω_h) önermiştir. Zinbarg vd. (2005) McDonald'ın ω_h değerini Cronbach'ın α ve Revelle'in β değerleriyle karşılaştırmıştır ve ω_h değerinin en iyi tahmin olduğu sonucuna varmışlardır. Ayrıca Guttman'ın Lambda 6 (G6) değeri, her bir maddedeki varyansın diğer tüm maddelerin doğrusal regresyonu ile açıklanabilen miktarını dikkate alan bir güvenilirlik katsayısıdır. Konjenerik ve çok boyutlu ölçme durumları, Cronbach alpha kestirimlerinde yanlılıklara sebep olabilmektedir (Cho & Kim, 2015; Graham, 2006; Sijtsma, 2009). Cronbach alpha katsayısındaki yanlılık sebebiyle, Cronbach vd. (1965) alt boyutları olan veri setlerinden elde edilen bileşik puanların güvenilirliği için tabakalı (stratified) alpha katsayısının kullanımını tavsiye etmişlerdir. Osburn (2000) ile Revelle ve Zinbarg (2009) çok boyutlu yapılarda tabakalı alpha katsayısının gerçek güvenirlüğün iyi bir kestiricisi olduğunu belirtmişlerdir. R kütüphanesinde bulunan "psych" (Revelle, 2023) paketindeki "omega" fonksiyonu ve "sirt" paketinde yer alan "stratafiied.cronbach.alpha" fonksiyonu yardımıyla hesaplanan güvenilirlik değerleri iki boyutlu veri için 0,82 ile 0,87 değerleri arasında yer alırken matematik bölümü alt boyutu için 0,76 ile 0,90, fen bilimleri bölümü alt boyutu için 0,69 ile 0,77 değerleri arasında yer almaktadır. Güvenirlik katsayıları iç tutarlık anlamında yorumlandığında veri setinin iç tutarlığının yeterli düzeyde olduğu görülmektedir.

20 maddelik veri setinin MTK'nin yerel bağımsızlık varsayımını sağlayıp sağlamadığını belirlemek için Yen'in Q_3 istatistiğinden faydalanılmıştır. Yerel bağımsızlık kişinin testteki maddelere verdiği cevapların birbirinden bağımsız olması durumudur (Hambleton & Swaminathan, 1985). Yerel bağımsızlığın incelenmesinde faydalanılan istatistiklerden biri Yen'in Q_3 istatistiğidir. Q_3 istatistiği testte yer alan madde çiftlerine odaklanır ve her bir çift için üretilen artık değerlerin korelasyonu olarak belirtilir. Yen'in Q_3 istatistik değerinin 0,20'den büyük olmaması, testte bulunan ilgili madde çiftleri için yerel bağımsızlık varsayımının sağlandığına kanıt oluşturmaktadır (Demars, 2016). Buna göre araştırmada R kütüphanesindeki "irtGUI" (Yıldız, 2021) paketi kullanılarak Yen'in Q_3 istatistiği ile yerel bağımsızlık testi yapılmıştır. 20 madde için 190 madde çiftinde 2PL ve 3 PL modellerine göre Q_3 değerleri elde edilmiş ve bulgular incelendiğinde her iki model için

de bulunan istatistik değerlerinin 0,20'yi geçmediği görülmektedir. Bu durum yerel bağımsızlık varsayımının sağlandığını göstermektedir.

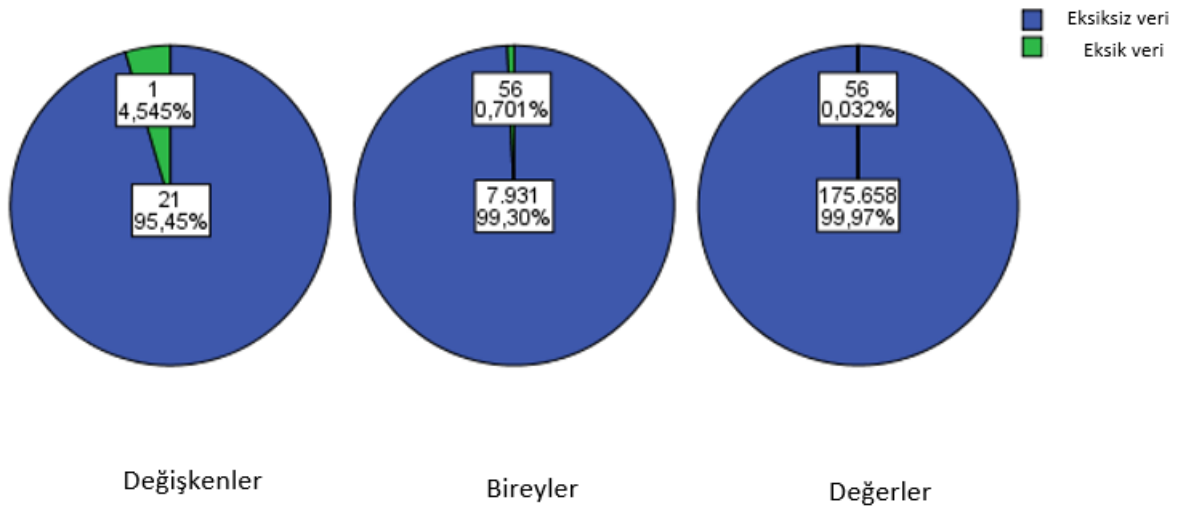
MTK'nin bir başka varsayımı madde karakteristik eğrisinin monoton artışıdır. Eğrinin monoton artışı Hambleton vd.'nin (1991) ifadesiyle, bireye ait yetenek düzeyinin artmasıyla bireyin ilgili maddeyi doğru cevaplama olasılığının artmasıdır. 20 maddelik veri seti için madde karakteristik eğrileri R kütüphanesindeki "irtGUI" (Yıldız, 2021) paketi kullanılarak elde edilmiştir. Madde karakteristik eğrileri incelendiğinde veri setindeki tüm maddelerde, bireylere ait yetenek düzeyleri arttığında ilgili maddeyi doğru cevaplama olasılıklarının da arttığı görülmektedir. Bu da madde karakteristik eğrisinin monoton artışı varsayımının sağlandığını gösterir.

Araştırma kapsamında çok boyutlu ve çok düzeyli olan veriye ortak değişken eklenerek analizler yapılmıştır. Evdeki kitap sayısı değişkeni ile birlikte bilgisayar veya tablet varlığı, internet varlığı, cep telefonu varlığı, aile eğitim düzeyi, "matematik derslerinde problemleri ne sıklıkla kendi başınıza çözersiniz?", "bu testin bir bilgisayar veya tablet üzerinde yapılmasını beğendiniz mi?", "fen bilimleri öğrenmekle ilgili bu ifadelere ne kadar katılıyorsunuz? fen bilimlerini seviyorum." ve "aşağıdaki terimlerin her birinin anlamını ne kadar iyi biliyorsunuz? kes ve yapıştır" gibi 26 farklı değişkenin öğrencilerin 20 maddeden aldıkları toplam puanı ile korelasyonuna bakılmış ve en yüksek korelasyon değerine sahip olan değişken evdeki kitap sayısı (0,32) olmuştur.

Verilere ait incelemeler yapıldığında Şekil 9'da yer alan ve ortak değişken olarak alınan evdeki kitap sayısına ait verilerde kayıp veri tespit edilmiştir. Evdeki kitap sayısı değişkeni için 56 bireye ait veriler kayıp olarak görünmektedir. Ayrıca hücre bazında bakıldığında kayıp veri oranı %0,032 olarak belirlenmiştir. Yapılan kayıp veri analizinde Little'nin (1988) MCAR (tamamen seçkisiz kayıp-TSK) testi sonucu manidar ($p < 0,05$) olarak bulunmuştur ve kayıpların tamamen seçkisiz kayıp olmadığı sonucuna ulaşılmıştır.

Şekil 9

20 Maddelik Veri Setine Ait Kayıp Veri Bilgileri



Karma modellere ait analizlerin gerçekleştirildiği Mplus 8 (Muthén & Muthén, 1998-2017) programı hem frekansçı hem de Bayesçi analiz kullanarak eksik verili modellerin tahmin edilmesini sağlar. Frekansçı analiz ile Mplus, sürekli, durdurulmuş (censored), ikili, sıralı kategorik (ordinal), sırasız kategorik (nominal), sayımlar veya bu değişken türlerinin kombinasyonları için Little ve Rubin (2002) tarafından sınıflanan kayıp türleri olan TSK, MAR (seçkisiz kayıp-SK) ve NMAR (seçkisiz olmayan kayıp-SOK) altında maksimum olabilirlik tahmini sağlar. SOK modellemesi, kategorik sonuçların kayıplığın göstergeleri olduğu ve kayıplığın sürekli ve kategorik gizli değişkenler tarafından tahmin edilebildiği durumlarda maksimum olabilirlik tahmini kullanılarak mümkündür (Muthen & Muthen, 2017). Bu sebeple araştırma kapsamında kullanılan ortak değişkenlerden kayıp olanları "9" olarak kodlanmış ve analizlere dahil edilmiştir. Ancak yapılan analizlerin karmaşıklığına kayıp verinin de eklenmesi parametre değerlerini olumsuz etkilemiş ve uzun olan analiz süresini daha da uzatmıştır. Bu sebeple veri setinde %0,7'lik kısma denk gelen evdeki kitap sayısını belirtmeyen 56 birey analiz kapsamından çıkarılmıştır.

3. Model Belirleme

Veri setinin çok boyutlu ve çok düzeyli bir yapıya sahip olduğu bilinmektedir ve boyut ile düzey sayısının en uygun olduğu modellerin seçimi yapılarak o modellere ilişkin DMF analizleri uygulanmıştır. Analizler Mplus ve R programları ile yapılmıştır ve KMTK modeline

ait analizlerden birinin kodu EK-E'de ÇDKMTK modeline ait analizlerden birinin kodu EK-F'de, ÇBÇDKMTK modeline ait analizlerden birinin kodu EK-G'de ve ortak değişkenin eklendiği ÇBÇDKMTK modeline ait analizlerden birinin kodu EK-H'de verilmiştir. Ayrıca örtük sınıfların belirlenmesinin ardından yapılan DMF analizleri için R kütüphanesinden "difR" (Magis ve diğerleri, 2010) paketinin "difMH" ve "difGMH" fonksiyonlarından yararlanılmıştır. MH yöntemi, diğer değişkenleri kontrol ederken ister nominal ister sıralı olsun, iki kategorik değişken arasındaki ilişkinin derecesini değerlendirmek için oldukça esnek bir metodoloji içermektedir. MH analitik yaklaşımlarının çok yönlülüğü, onları hem ikili hem de çoklu maddelerin farklı işlevlerinin değerlendirilmesinde çok popüler hale getirmiştir (Fidalgo & Madeira, 2008). İki örtük sınıfın bulunduğu durumlarda standart MH testi kullanılırken, ikiden fazla örtük sınıfın bulunduğu durumlarda Penfield (2001) tarafından ikiden fazla grupta kullanılması tavsiye edilen GMH kullanılmıştır. Anlamlı bir GMH sonucunun ardından, her bir örtük sınıf çifti için tek biçimli DMF'yi değerlendirmek üzere ikili MH testleri kullanılmıştır. GMH istatistiği, DMF olmayan sıfır hipotezi altında J-1 serbestlik derecesine sahip bir ki-kare değişkeni olarak dağıtılır.

Karma modellere dayalı parametre kestiriminde en uygun modeli belirlemek için en çok kullanılan üç bilgi ölçütü indeksi Akaike'nin (1974) bilgi kriteri (AIC), Schwarz'ın (1978) Bayes bilgisi kriteri (BIC) ve BIC'in örneklem büyüklüğü ayarlı versiyonudur (SABIC; Sclove, 1987). Kriterlere ait eşitlikler aşağıda verilmiştir.

$$AIC = -2\log L + 2p$$

$$BIC = -2\log L + p \cdot \ln(N)$$

$$SABIC = -2\log L + p \cdot \ln((N+24)/2)$$

Burada "L" olabilirlik fonksiyonunu gösterirken "p" kestirilecek parametre sayısını ve "N" örneklem büyüklüğünü göstermektedir. Değeri daha küçük olan AIC ve BIC ölçütleri daha iyi model uyumunu temsil etmektedir. AIC ve BIC sonuçlarının farklı olduğu durumlarda araştırmacıların tercihleri değişebilmektedir. AIC kriterinin tutarsızlık sebebiyle

sınıf sayısını olduğundan daha fazla kestirmeye eğilimli olduğu McLachlan ve Peel (2000) tarafından belirtilirken BIC ve SABIC kriterleri, olasılık fonksiyonuna daha fazla düzeltme uygulamakta olduğu ve örneklem büyüklüğü arttıkça, BIC kriteri tutarlı olarak doğru sınıfları seçmeye eğilimli olduğu Vrieze (2012) tarafından ifade edilmiştir. Li vd. (2009) ile Preinerstorfer ve Formann (2011), tek düzeyli KMTK modellerinde model seçimi için BIC değerinin AIC değerinden daha hatasız olduğunu öne sürmüşlerdir. Benzer şekilde, Şen vd. (2018) BIC değerinin ÇDKMTK modellerinin seçiminde daha doğru olduğunu öne sürmüştür. Cho (2007), Uyar (2015) ve Zhu (2013) ÇDKMTK ile uygun sınıflı modelin belirlenmesinde AIC değerini kullanırlarken Cho ve Cohen (2010), Choi vd. (2015), Li vd. (2009); Saatçioğlu (2021), Şen ve Toker (2021) çalışmalarında örtük sınıf sayısını belirlemek için BIC kriterini kullanmışlardır. Mevcut bilgiler ışığında araştırma kapsamında BIC değeri örtük sınıflar belirlemek için kullanılacak ölçüt olarak belirlenmiştir.

Model belirleme aşamasında öncelikle iki parametrelili lojistik (2PL) model kullanılarak tek boyut ve iki boyut için modeller oluşturulmuş ve düzeyin olmadığı modeller arasından en düşük BIC değerine sahip olan model ile düzeyin olduğu modeller arasından yine en düşük BIC değerine sahip olan model seçilmiştir. Model isimlendirmesinde “C” düzey içi “CB” düzeyler arasını göstermek üzere Tablo 6’da model isimlerine ait örnekler verilmiştir.

Tablo 6

Model İsimlerine Ait Örnekler ve Açıklamalar

Model Adı	Düzyey	Açıklama
C4	Tek	Tek düzeyde dört örtük öğrenci sınıfının yer aldığı modeldir.
CB1-C3	İki	İki düzeyin yer aldığı modelde birinci düzeyde üç örtük öğrenci sınıfı, ikinci düzeyde ise bir örtük ülke sınıfı yer almaktadır.
CB2-C2	İki	İki düzeyin yer aldığı modelde birinci düzeyde iki örtük öğrenci sınıfı, ikinci düzeyde ise iki örtük ülke sınıfı yer almaktadır.

Tablo 6'daki bilgilere 2PL ve 3PL model, tek ve iki boyut ile ortak değişkenlerin eklenme durumları dahil edilerek modeller genişletilmiştir. Örneğin CB2-C2 adlı modelin hem tek boyut 2PL ile 3PL hem iki boyut 2PL ile 3PL ve birinci ile ikinci düzeye ortak değişkenlerin eklendiği 2PL ile 3PL olmak üzere altı farklı şekli incelenmiştir. Araştırma kapsamında tek ve iki boyut, 2PL ve 3PL tek ve iki düzeye de ortak değişkenin eklendiği modeller olmak üzere 54 model analiz edilmiştir. Boyut, düzey, parametre ve ortak değişken durumlarına göre modellerden en iyi bilgi ölçütü indeksine sahip olanlar Tablo 7'de yer almaktadır.

Tablo 7

Bilgi Ölçütü İndeksi En İyi Olan Modeller

Modeller	LL	PS	AIC	BIC	SABIC
Tek Boyutlu C3 2PL	-97217,80	122	194679,61	195531,85	195144,16
Tek Boyutlu C3 3PL	-97196,38	122	194636,76	195489,00	195101,314
Tek Boyutlu CB2-C2 2PL	-94994,12	222	190432,24	191983,03	191277,56
Tek Boyutlu CB2-C2 3PL	-95014,32	243	190514,64	192212,13	191439,92
İki Boyutlu C2 2PL	-97222,81	84	194613,61	195200,39	194933,46
İki Boyutlu C3 3PL	-97120,18	143	195526,361	195525,298	195070,872
İki Boyutlu CB2-C2 2PL	-94942,28	201	190286,57	191690,67	191051,93
İki Boyutlu CB2-C2 3PL	-95317,06	162	190958,13	192089,79	191574,99
Ortak Değişkenli C2 2PL	-95928,29	83	192022,58	192602,30	192338,62
Ortak Değişkenli C2 3PL	-95920,982	105	192451,96	193785,45	192251,78
Ortak Değişkenli CB-2-C2 2PL	-93815,46	225	188080,91	189651,08	188936,75
Ortak Değişkenli CB-2-C2 3PL	-95136,269	185	190642,54	191934,87	191346,97
İki Ortak Değişkenli CB2-C2 2PL	-93790,817	226	188033,63	189610,78	188892,59
İki Ortak Değişkenli CB2-C2 3PL	-194534,29	186	189440,59	190738,60	190147,53

NOT: LL: Log likelihood değeri, PS: Parametre Sayısı, AIC: Akaike'nin Bilgi Kriteri, BIC: Bayes Bilgisi Kriteri, SABIC: Örneklem Büyüklüğü Ayarlı BIC.

Tablo 7'ye göre tek boyut ve tek düzeyli modeller (Tek Boyutlu C3 2PL ve Tek Boyutlu C3 3PL) haricinde neredeyse tamamında birinci düzey örtük sınıf sayısı iki olarak bulunmuştur. Yine tek boyut ve tek düzeyli modeller haricinde tüm model çiftlerinde 2PL modellerin 3PL modellerden göre veri setine daha uygun olduğu görülmektedir. 54 model içerisinde en iyi bilgi ölçütü indeksine sahip olan model "İki Ortak Değişkenli CB2-C2 2PL" modelidir.

4. Seçilen Modellere Ait Betimsel İstatistikler ve Parametre Tahminleri

Veri setine boyut, düzey ve ortak değişken bakımından en iyi uyumu gösteren modeller için birey sayısı, madde güçlük indeksi, puan ortalamaları ve standart sapmaları ile ülke düzeyine sahip modellerde örtük ülke sınıflarını gösteren tablolar oluşturulmuştur. Ancak model sayısının fazla olması sebebiyle sadece en iyi bilgi ölçütü indeksine sahip model olan "İki Ortak Değişkenli CB2-C2 2PL" modeline ve gözlenen değişken olan cinsiyete ait betimsel istatistikler verilmiştir. Tablo 8'de modele ait örtük sınıflara ayrılan bireylerin cinsiyet dağılımı yer almaktadır.

Tablo 8

İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 KMTK Modelinde Öğrenci ve Ülke Düzeyi Örtük Sınıfları için Nihai Sınıf Sayıları ve Oranları

Ülke Düzeyi Örtük Sınıfı	Öğrenci Düzeyi Örtük Sınıfı	
	1	2
1	1443 (764/679)	1048 (485/563)
2	3474 (1754/1720)	1966 (965/1001)

Tablo 8'e göre dört örtük öğrenci sınıfında da kız ve erkek öğrenci sayıları birbirine yakındır. Örtük ülke sınıflarında ise 2. örtük ülke sınıfı 1. örtük ülke sınıfının yaklaşık iki katı kadar öğrenciye sahiptir. Örtük ülke sınıfı 1'de yer alan örtük öğrenci düzeyi 1, örtük öğrenci düzeyi 2'nin yaklaşık 1,5 katı kadar iken örtük ülke sınıfı 2'de yer alan örtük öğrenci düzeyi

1, örtük öğrenci düzeyi 2'nin 1,5 katından da fazladır. Tablo 9'da seçilen modelin öğrenci düzeyindeki madde güçlük indeksleri verilmiştir. Örtük öğrenci sınıfları kodlanırken örtük ülke sınıfı ve örtük öğrenci sınıfı sırasıyla kodlanmıştır. Örneğin ÖÜ1-1 örtük ülke sınıfı 1 içinde yer alan örtük öğrenci sınıfı 1'i gösterirken ÖÜ2-2 örtük ülke sınıfı 2 içinde yer alan örtük öğrenci sınıfı 2'yi göstermektedir.

Tablo 9

İki Düzeye Ortak Değişkenin eklendiği 2 PL CB2-C2 KMTK Modeli Öğrenci Düzeyi için Nihai Modelin Madde Güçlük İndeksleri

Madde	ÖÜ1-1	ÖÜ1-2	ÖÜ2-1	ÖÜ2-2
	β_1	β_2	β_3	β_4
1	0,71	-3,88	3,07	-2,85
2	1,05	-0,51	2,05	-1,91
3	-0,09	-0,85	1,16	-0,63
4	-0,11	-1,68	0,64	-2,01
5	0,70	-1,39	4,83	-0,35
6	-0,46	-2,43	1,81	-4,88
7	0,62	-1,49	3,93	-1,12
8	0,10	-1,45	2,59	-2,03
9	-0,40	-1,53	0,13	-2,11
10	0,17	-0,59	1,38	-0,33
11	-0,91	-1,52	0,46	-0,91
12	0,15	-2,71	0,23	-1,56
13	1,55	0,05	2,06	-0,84
14	0,54	-4,74	0,92	-3,97
15	-0,28	-4,10	-0,22	-5,80
16	0,27	-1,05	0,28	-3,19
17	-0,08	-0,92	1,01	-1,63
18	0,32	-1,14	2,84	0,10

19	0,06	-2,82	0,45	-2,45
20	-0,46	-2,59	0,89	-0,79

Tablo 9'a göre matematik bölümüne ait 10 maddenin beş tanesinde ÖÜ1-2, beş tanesinde ise ÖÜ2-2 en düşük madde güçlüğü indeksine sahiptir. Fen bilimleri bölümünde ise altı maddede ÖÜ1-2, dört tanesinde ise ÖÜ2-2 en düşük madde güçlüğü indeksine sahiptir. 20 maddenin tamamında ise ÖÜ2-1 en yüksek madde güçlüğü indeksine sahiptir. Bu durum ÖÜ1-2 ve ÖÜ2-2'nin başarılı örtük sınıflar iken ÖÜ2-1'in en başarısız örtük sınıf olduğunu gösterir. Tablo 10'da seçilen modelde örtük ülke düzeyindeki madde güçlük indeksleri verilmiştir.

Tablo 10

İki Ortak Değişkenli 2 PL CB2-C2 KMTK Modeli Ülke Düzeyi için Nihai Modelin Madde Güçlük İndeksleri

Madde	ÖÜ1	ÖÜ2
	β_1	β_2
1	-0,43	0,28
2	0,11	0,28
3	-0,45	0,30
4	-0,62	-0,18
5	-0,22	0,98
6	-0,92	-0,05
7	-0,16	0,59
8	-0,46	0,32
9	-0,78	-0,50
10	-0,22	0,58
11	-1,20	-0,24
12	-0,56	-0,34

13	0,63	0,38
14	-0,38	-0,10
15	-0,85	-0,74
16	-0,37	-0,30
17	-0,49	-0,15
18	-0,28	1,05
19	-0,58	-0,24
20	-0,91	-0,01

Tablo 10'a göre 13. madde haricindeki tüm maddelerde ÖÜ1'in madde güçlük indeksi değeri ÖÜ2'den daha düşüktür. Bu durum ÖÜ1'in ÖÜ2'den daha başarılı olduğunu gösterir. Tablo 11'de seçilen modele ait örtük sınıfların puan ortalamaları ve standart sapmaları verilmiştir.

Tablo 11

İki Ortak Değişkenli 2 PL CB2-C2 ÇBÇDKMTK Modeli Örtük Sınıflara Ait Puan Ortalamaları ve Standart Sapmaları

Örtük Sınıf	Puan Ortalaması (Kız/Erkek)	Standart Sapma (Kız/Erkek)
ÖÜ1-1	9,44 (9,49/9,39)	3,92 (3,90/3,93)
ÖÜ1-2	15,74 (15,62/15,85)	3,45 (3,36/3,52)
ÖÜ2-1	6,97 (7/6,94)	2,83 (2,79/2,87)
ÖÜ2-2	14,26 (14,19/14,33)	2,79 (2,79/2,78)
ÖÜ1	12,09 (11,87/12,32)	4,85 (4,76/4,94)
ÖÜ2	9,61 (9,55/9,66)	4,49 (4,43/4,56)

Tablo 11'e göre en yüksek puan ortalamasına sahip örtük sınıf ÖÜ1-2 iken en düşük puan ortalamasına sahip örtük sınıf ÖÜ2-1'dir. Ayrıca ÖÜ1'in puan ortalaması da ÖÜ2'den daha yüksektir. Tablo 9 ve Tablo 10'da verilen madde güçlük indeksleri ile de bu sonuçlar

elde edilmiştir. Puan ortalamaları düşük olmasına rağmen ÖÜ1-1 ve ÖÜ2-1'in standart sapmaları aynı örtük ülkede yer aldıkları ÖÜ1-2 ve ÖÜ2-2' ye göre daha yüksektir. Tablo 12'de seçilen modeldeki ülke düzeyinde yer alan örtük ülke sınıflarındaki ülkeler verilmiştir.

Tablo 12

İki Ortak Değişkenli 2 PL CB2-C2 ÇBÇDKMTK Modeli Ülke Düzeyi Örtük Sınıflarda Yer Alan Ülkeler

Ülke Düzeyi Örtük Sınıflar	
Örtük Sınıf 1	Örtük Sınıf 2
Tayvan	Şili
Hong Kong	Fransa
Kore	Gürcistan
Malezya	Macaristan
Norveç	İsrail
Singapur	İtalya
İsveç	Litvanya
İngiltere	Portekiz
	Rusya
	BAE
	Türkiye
	ABD
	Finlandiya
	Katar

Tablo 12'ye göre ÖÜ1'de genel olarak başarılı Asya ve Avrupa ülkeleri yer almaktadır. ÖÜ2'de ise 10 Avrupa ülkesi ile ikişer Amerika ve Asya kıtası ülkeleri yer almaktadır. Tablo 13'te cinsiyet gözlenen değişkenine ait birey sayıları, kız ile erkek öğrenci oranları, puan ortalamaları ve standart sapmaları verilmiştir.

Tablo 13

Gözlenen Değişken Olan Cinsiyete Ait Sayılar, Oranlar, Puan Ortalamaları ve Standart Sapmaları

Cinsiyet	Birey Sayısı	Oran	Puan Ortalaması	Standart Sapma
Kız	3968	0,50	10,2	4,67
Erkek	3963	0,50	10,5	4,83

Tablo 13'e göre seçilen veri setindeki kız ve erkek öğrenci sayıları birbirine çok yakındır. Ayrıca puan ortalamalarında erkekler daha yüksek olsa da puanlar arasında anlamlı bir farklılaşma yoktur.

Bölüm 4

Bulgular, Yorumlar ve Tartışma

Bu bölümde araştırmanın alt problemlerine ait yapılan analizlerin sonuçlarına göre elde edilen bulgular ile literatür ve ilgili araştırmalar göz önüne alınarak yapılan yorumlar ve tartışmaya yer verilmiştir.

Birinci Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma

Cinsiyet ve evdeki kitap sayısı gözlenen değişkenleri için Mantel Haenszel yöntemine, Karma Madde Tepki Kuramına, Çok Düzeyli Karma Madde Tepki Kuramına, Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına ve ortak değişkenin dahil olduğu Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına göre yapılan analizlerde DMF'li madde sayıları ve etki büyüklüğü düzeyleri farklılaşmakta mıdır?

Birinci alt problemin çözümünde önce gözlenen değişkenlere ait DMF analizleri verilmiştir. Cinsiyet değişkeni için MH yöntemi ile yapılan DMF analizine ait sonuçlar Tablo 14'te yer almaktadır.

Tablo 14

Cinsiyet Değişkenine Göre Yapılan DMF Analizi Sonuçları

Madde	Ki-Kare	Alpha MH	Delta MH	Etki Büyüklüğü
1	3,34	1,10	-0,23	A
2	39,74*	0,71	0,79	A
3	31,20*	0,74	0,69	A
4	0,96	1,06	-0,13	A
5	0,63	1,04	-0,09	A
6	1,48	1,07	-0,15	A
7	9,42*	1,16	-0,40	A
8	10,96*	1,19	-0,41	A
9	15,00*	0,81	0,49	A

10	35,89*	0,73	0,75	A
11	4,97*	0,89	0,26	A
12	23,92*	1,29	-0,59	A
13	4,46*	1,11	-0,24	A
14	0,43	1,03	-0,07	A
15	4,18*	1,12	-0,26	A
16	0,00	0,99	0,00	A
17	1,48	0,94	0,14	A
18	1,63	0,93	0,15	A
19	2,64	1,09	-0,20	A
20	13,71*	1,20	-0,44	A

Not: A: İhmal edilebilir etki, B: Orta düzey etki, C: Büyük etki, etki düzeyi büyüklüklerini gösterir.

Tablo 14'e göre cinsiyet değişkeni için yapılan DMF analizinde manidarlık düzeyi 0,05 olarak kabul edildiğinde 2, 3, 7, 8, 9, 10, 11, 12, 13, 15 ve 20. maddeler DMF göstermektedir. Ancak etki büyüklüğüne bakıldığında tüm bu maddelerin A düzeyinde yani ihmal edilebilir düzeyde DMF'li olduğu görülmektedir. Tablo 15'te evdeki kitap sayısı değişkeni için GMH yöntemi ile yapılan DMF analizine ait sonuçlar verilmiştir.

Tablo 15

Evdeki Kitap Sayısı Değişkenine Göre Yapılan DMF Analizi Sonuçları

Madde	Ki-Kare	Alpha MH	Delta MH	Etki Büyüklüğü
1	2,15	1,10	-0,23	A
2	0,08	0,98	0,05	A
3	2,49	0,90	0,25	A
4	0,01	0,99	0,02	A
5	1,64	0,92	0,21	A
6	6,72*	1,18	-0,40	A

7	0,32	1,04	-0,09	A
8	0,01	1,01	-0,02	A
9	0,06	1,02	-0,04	A
10	6,92*	0,83	0,43	A
11	8,85*	0,83	0,44	A
12	0,31	0,96	0,08	A
13	0,15	0,98	0,06	A
14	6,80*	1,18	-0,40	A
15	9,95*	1,23	-0,48	A
16	1,75	1,09	-0,20	A
17	20,80*	0,75	0,68	A
18	1,46	0,92	0,20	A
19	21,51*	1,36	-0,73	A
20	0,69	0,95	0,12	A

Not: A: İhmal edilebilir etki, B: Orta düzey etki, C: Büyük etki, etki düzeyi büyüklüklerini gösterir.

Tablo 15'e göre evdeki kitap sayısı değişkeni için yapılan DMF analizinde manidarlık düzeyi 0,05 olarak kabul edildiğinde 6, 10, 11, 14, 15, 17 ve 19. maddeler DMF göstermektedir. Ancak etki büyüklüğüne bakıldığında tüm bu maddelerin A düzeyinde yani ihmal edilebilir düzeyde DMF'li olduğu görülmektedir. Bu sebeple de ikili karşılaştırmalara ait sonuçlar tabloda verilmemiştir. Gözlenen değişkenler olan cinsiyet ve evdeki kitap sayısı değişkenlerinde 2, 3, 7, 8, 9, 12 ve 15. maddeler DMF'li bulunmuştur ve etki büyüklükleri de A düzeyindedir. Literatürde yer alan birçok araştırmada (Aydemir, 2023; Bayram, 2024; Ceylan, 2022; Gürdil, 2023; Hasaıçebi, 2021 vb.) DMF analizleri gözlenen gruplara dayalı olarak yapılmaktadır. Gözlenen grupların homojen yapıda olduğu varsayılarak yapılan bu analizlerin hatalı çıkarımlara sebep olabileceği çeşitli araştırmalarda (Cho & Cohen 2010; Maij-de Meij, Kelderman & Flier, 2010; McLachlan & Peel, 2000; Samuelsen, 2005 vb.) belirtilmiştir. Modellere ait DMF analizlerine ilişkin sonuçlar karşılaştırma yapma açısından

katkı sağlayacaktır. Tablo 16'da en iyi bilgi kriteri indekslerine sahip iki düzeye ortak değişkenin eklendiği ülke düzeyinde 2 PL CB2-C2 modeli için ÖÜ1-1 ile ÖÜ1-2 örtük sınıflarına ait MH ile yapılan DMF analizi sonuçları yer almaktadır.

Tablo 16

İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 Öğrenci Düzeyinde KMTK Modelinde ÖÜ1'e Ait DMF Analizi Sonuçları

Madde	Ki-Kare	Alpha MH	Delta MH	Etki Büyüklüğü
1	28,23*	2,33	-1,983	C
2	1,79	1,27	-0,556	A
3	22,71*	0,41	2,073	C
4	0,14	0,93	0,168	A
5	0,02	0,98	0,054	A
6	7,23*	1,64	-1,160	B
7	3,18	1,34	-0,688	A
8	19,03*	0,48	1,732	C
9	5,00*	0,64	1,059	B
10	39,68*	0,32	2,649	C
11	25,51*	0,38	2,287	C
12	8,52*	1,62	-1,136	B
13	29,18*	0,37	2,3105	C
14	16,66*	1,90	-1,5114	C
15	1,97	1,29	-0,5893	A
16	2,07	1,26	-0,5491	A
17	18,36*	0,49	1,681	C
18	12,09*	0,56	1,375	B
19	7,33*	1,59	-1,0882	B
20	2,43	0,79	0,5624	A

Not: A: İhmal edilebilir etki, B: Orta düzey etki, C: Büyük etki, etki düzeyi büyüklüklerini gösterir.

Tablo 16'da DMF analizinde manidarlık düzeyi 0,05 olarak kabul edildiğinde 2, 4,5, 7, 15, 16 ve 20. maddeler haricindeki tüm maddelerin DMF gösterdiği görülmektedir. Etki büyüklüklerine bakıldığında DMF gösteren maddelerden 6, 9, 12, 18 ve 19. maddelerin B düzeyinde diğer sekiz maddenin ise C düzeyinde DMF'li olduğu görülmektedir. Tablo 17'de iki düzeyde ortak değişkenin modele eklendiği ülke düzeyinde 2 PL CB2-C2 modeli için ÖÜ2-1 ile ÖÜ2-2 örtük sınıflarına ait MH ile yapılan DMF analizi sonuçları yer almaktadır.

Tablo 17

İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 Öğrenci Düzeyinde KMTK Modelinde ÖÜ2'ye Ait DMF Analizi Sonuçları

Madde	Ki-Kare	Alpha MH	Delta MH	Etki Büyüklüğü
1	192,91*	3,66	-3,050	C
2	170,89*	3,26	-2,779	C
3	5,73*	1,26	-0,540	A
4	122,05*	2,85	-2,463	C
5	84,42*	2,36	-2,020	C
6	155,38*	3,31	-2,811	C
7	204,33*	3,90	-3,200	C
8	192,63*	3,69	-3,065	C
9	64,58*	2,23	-1,884	C
10	0,57	0,93	0,170	A
11	30,50*	1,72	-1,274	B
12	28,40*	1,62	-1,131	B
13	4,36*	1,22	-0,4647	A
14	82,70*	2,33	-1,9843	C
15	146,87*	4,00	-3,2571	C
16	17,10*	1,45	-0,8736	A

17	45,30*	1,85	-1,4438	B
18	30,06*	1,67	-1,2068	B
19	55,15*	1,96	-1,5831	C
20	14,87*	1,41	-0,8139	A

Not: A: İhmal edilebilir etki, B: Orta düzey etki, C: Büyük etki, etki düzeyi büyüklüklerini gösterir.

Tablo 17'de DMF analizinde manidarlık düzeyi 0,05 olarak kabul edildiğinde 10. madde haricindeki tüm maddelerin DMF gösterdiği görülmektedir. Etki büyüklüklerine bakıldığında DMF gösteren maddelerden 3,13,16 ve 20. maddelerin A düzeyinde, 11, 12, 17 ve 18. maddelerin B düzeyinde diğer 11 maddenin ise C düzeyinde DMF'li olduğu görülmektedir. Tablo 18'de iki düzeyde ortak değişkenin modele eklendiği ülke düzeyinde 2 PL CB2-C2 modeli için MH ile yapılan DMF analizi sonuçları yer almaktadır.

Tablo 18

İki Düzeye Ortak Değişkenin Eklendiği 2 PL CB2-C2 Ülke Düzeyinde KMTK Modeline Ait DMF Analizi Sonuçları

Madde	Ki-Kare	Alpha MH	Delta MH	Etki Büyüklüğü
1	0,36	1,06	-0,13	A
2	2,78	1,20	-0,43	A
3	84,75*	2,43	-2,08	C
4	0,15	0,96	0,09	A
5	25,22*	1,58	-1,08	B
6	3,72	1,24	-0,50	A
7	0,74	1,09	-0,20	A
8	23,40*	1,63	-1,15	B
9	0,02	0,98	0,04	A
10	40,26*	1,86	-1,46	B
11	36,75*	1,87	-1,47	B

12	21,54*	0,62	1,13	B
13	13,80*	0,65	1,03	B
14	78,51*	0,40	2,13	C
15	42,64*	0,52	1,55	C
16	30,94*	0,56	1,36	B
17	5,78*	0,78	0,60	A
18	76,34*	2,72	-2,35	C
19	44,40*	0,50	1,62	C
20	40,60*	2,03	-1,67	C

Not: A: İhmal edilebilir etki, B: Orta düzey etki, C: Büyük etki, etki düzeyi büyüklüklerini gösterir.

Tablo 18’de DMF analizinde manidarlık düzeyi 0,05 olarak kabul edildiğinde 1, 2, 4, 6, 7 ve 9. maddeler haricindeki tüm maddelerin DMF gösterdiği görülmektedir. Etki büyüklüklerine bakıldığında DMF gösteren maddelerden 17. maddenin A düzeyinde, 5, 8, 10, 11, 12, 13 ve 16. maddelerin B düzeyinde diğer altı maddenin ise C düzeyinde DMF’li olduğu görülmektedir. Farklı koşullarda (boyut-düzyey-ortak değişken) veri setine en uygun modeller için yapılan öğrenci düzeyindeki DMF analizlerinde her modele ait analizde ortak olarak DMF gösteren maddeler ile Tablo 19 oluşturulmuştur.

Tablo 19

Veri Setine En Uygun Modeller Kullanılarak Yapılan Öğrenci Düzeyindeki DMF Analizi

Sonuçları

Model	DMF’li Madde Sayısı	DMF’li Maddeler ve Etki Büyüklüğü Düzeyleri			B ve C düzeyi DMF’li madde sayısı
		A	B	C	
Cinsiyet	11	2,3,7,8,9,10,11 12,13,15,20	-	-	-

Evdeki Kitap Sayısı	13	1,2,3,6,7,8,9,12 14,15,16,17,19	-	-	-
Tek Boyutlu C3	17	2,5,6,16,18	8	1,3,7,9,10,11,13 14,15,17,19	12
Tek Boyutlu CB2-C2	19	-	-	3,12,16,17,19,20	19
İki Boyutlu C2	17	8,11,17	2,5,7,20	1,3,4,9,10,12 14,15,16,19	14
İki Boyutlu CB2-C2	19	-	-	1,2,3,4,5,6,7,8,9 12,14,17,18,19,20	19
Ortak Değişkenli C2	15	6,8,9,18	7,16,17	1,3,10,11,13,14,15,19	11
Ortak Değişkenli CB-2-C2	19	-	-	1,2,4,5,6,7,8,9,10,11,12,13,14 ,15,16,18,19,20	19
İki Ortak Değişkenli CB-2-C2	12	-	12,18	1,8,14	12

Tablo 19'a göre gözlenen gruplara göre yapılan analizlerde cinsiyet için 11 maddede evdeki kitap sayısı için 13 maddede DMF'li madde bulunmaktadır. Bu maddelerin tamamı A düzeyinde etki büyüklüğüne sahiptir ve KMTK ile oluşturulan örtük sınıflar için yapılan tüm DMF analizlerindeki madde sayılarından azdır. Çalışmada elde edilen bu sonuç aynı zamanda DMF analizi için KMTK modellerinin kullanımına dayanan önceki araştırmalarla da (Cho & Cohen, 2010; Cohen & Bolt, 2005; Cohen ve diğerleri, 2005; Samuelsen, 2005) tutarlıdır. Düzeyin ihmal edildiği tek ve iki boyutlu modellerde (Tek boyutlu C3-İki Boyutlu C2) 17'şer madde DMF'li bulunmuştur. Yedi madde ortak olarak C düzeyi etki büyüklüğüne sahip olmasına rağmen DMF göstermeyen maddeler ile A ve B düzeyi etki büyüklüğüne sahip maddeler ortak değildir. B ve C düzeyi DMF'li madde sayıları 12 ve 14'tür. Samuelsen (2005), gözlenen gruplar ve örtük sınıflar tamamen örtüşmediğinde, gözlenen grupların kullanımının DMF'nin belirlenmesinde Tip I hata oranlarının artmasına, güç kaybına ve

DMF'nin büyüklüğünün olduğundan az tahmin edilmesine neden olacağını belirtmektedir. Bununla birlikte, gözlenen gruplar ve örtük sınıflar arasındaki örtüşme miktarı azaldıkça, DMF içeren maddeleri doğru bir şekilde belirleme gücü de azalacağı ve DMF'nin gerçek büyüklüğünün giderek daha fazla gizleneceğini ifade etmiştir. Son olarak bu durumun maddelerin başarılı bir şekilde tespit edilmesini zorlaştıracağını vurgulamıştır.

Boyutun ihmal edildiği modellerde (Tek boyutlu C3- Tek Boyutlu CB2-C2) üç madde ortak olarak C düzeyi etki büyüklüğüne sahiptir. Bunun yanında C3 için DMF'li olmayan iki madde CB2-C2 modelinde C düzeyi etki büyüklüğüne sahiptir. B ve C düzeyi DMF'li madde sayıları 12 ve 19'dur. Ortak değişkenin olmadığı iki boyutlu modellerde (İki boyutlu C2- İki Boyutlu CB2-C2) sırasıyla 17 ve 19 madde DMF'li bulunmuştur. C düzeyi etki büyüklüğüne sahip yedi madde ortaktır. B ve C düzeyi DMF'li madde sayıları 14 ve 19'dur.

Ortak değişkenin modele dahil edildiği çalışmalarda (Ortak Değişkenli C2, Ortak Değişkenli CB2-C2) farklı sayılarda (5 ve 1) DMF'li olmayan madde bulunmaktadır. DMF'li üç madde ortak olarak C düzeyi etki büyüklüğüne sahiptir. İki düzey için de ortak değişkene sahip çok boyutlu çok düzeyli model (iki ortak değişkenli CB2-C2) ile sadece öğrenci düzeyinde ortak değişkenin yer aldığı çok boyutlu çok düzeyli model (ortak değişkenli CB2-C2) karşılaştırıldığında 12 maddenin ortak olarak DMF'li olduğu görülmektedir. DMF'li üç madde ortak olarak C düzeyi etki büyüklüğüne sahiptir. B ve C düzeyi DMF'li madde sayıları 19 ve 12'dir. Her iki düzeye de ortak değişkenin eklenmesi ile B ve C etki büyüklüğüne sahip madde sayısında azalma görülmüştür. Cho ve Cohen'in (2010) çalışmasında ortak değişken eklenmesi DMF'li madde sayısını değiştirmemiştir. DMF kapsamında yapılan çalışmalar çoğunlukla yöntemleri karşılaştırma üzerine yapılmıştır (Örneğin Aydemir, 2023; Bayram, 2024; Erdoğan, 2021; Hasançebi, 2021 vb.). Araştırma alt probleminin tartışılabileceği literatür sınırlı olduğu için ilk olarak en ilgili çalışma ardından kısmen ilgili çalışmalar hakkında bilgiler aşağıda verilmiştir.

Finch ve Finch (2013) ÇBÇDKMTK modeli üzerine yaptıkları araştırmada öğrenci düzeyinde örtük sınıfların kendi içlerindeki DMF'li maddeleri vermemiş ve "diğer seviyenin

belirli sınıflarına bakılmaksızın, öğrenci ve okul seviyesindeki DMF sonuçları” olarak tek tablo şeklinde göstermiştir. Öğrenci ve okul düzeyinde madde güçlük indekslerinin verildiği tablolarda öğrenci düzeyi için yedi matematik ve 12 okuma maddesi DMF’li bulunurken okul düzeyi için bir matematik ve üç okuma maddesi DMF’li bulunmuştur.

Yalçın (2018) örtük sınıflar aracılığıyla DMF belirlemeyi amaçladığı çalışmasında örtük sınıfları belirledikten sonra MH yöntemi ile “örtük grup yaklaşımına dayalı KMTK modelini” karşılaştırmıştır. MH yönteminin gözlenen gruplara dayalı yöntem olarak ifade edildiği karşılaştırmada Mplus çıktısı olan estimate değerinin (madde ayırt edicilik indeksine eşit olan değer bkz. Şen ve diğerleri, 2020) standart hata değerine bölünmesiyle elde edilen değer aracılığıyla örtük grup yaklaşımına dayalı olarak DMF analizi sonuçları elde etmiştir. Elde edilen sonuçlara göre MH ile yapılan analizde 10 madde, örtük grup yöntemiyle dört madde DMF’li bulunmuştur ve DMF’li bulunan maddelerden iki tanesi ortaktır. Farklı şartlarda (boyut-düzyey-ortak değişken) veri setine en uygun modeller için yapılan ülke düzeyindeki DMF analizi sonuçlarına göre Tablo 20 oluşturulmuştur.

Tablo 20

Veri Setine En Uygun Modeller Kullanılarak Yapılan Ülke Düzeyindeki DMF Analizi

Sonuçları

Model	DMF’li Madde Sayısı	DMF’li Maddeler ve Etki Büyüklüğü Düzeyleri		
		A	B	C
Tek Boyutlu CB2-C2	11	1,8,18	10,12	3,11,14,15,16,19
İki Boyutlu CB2-C2	13	1,5,7,16	8,9,15,19	3,10,11,12,14
Ortak Değişkenli CB-2-C2	17	10,12,16	2,17,19	1,3,4,5,6,7,8,10,11,18,20
İki Ortak Değişkenli CB-2- C2	14	17	5,8,10,11,12,13,16	3,14,15,18,19,20

Tablo 20'deki sonuçlara göre veri setine en uygun 2 PL modellerde DMF'li en az 11 madde bulunmaktadır ve B ile C düzeyi etki büyüklüğüne sahip en az sekiz madde bulunmaktadır. Ortak değişken bulunmayan iki modelde (Tek Boyutlu CB2-C2 ve İki Boyutlu CB2-C2) A etki büyüklüğü düzeyinde birer ve C etki büyüklüğü düzeyinde üçer madde ortaktır. Ayrıca önemli etki büyüklüğü düzeyleri kabul edilen B ve C düzeyi DMF'li maddelerin yedi tanesi ortaktır. Ortak değişkenin yer aldığı modellerde ise B ile C düzeyi DMF'li olan sırasıyla 14 ve 13 madde vardır ve 11'er madde ortak olarak B ile C düzeylerinde bulunmaktadır. Ayrıca C etki büyüklüğü düzeyinde üçer madde ortaktır.

Cho ve Cohen'in (2010) çalışmalarında gözlenen gruplara göre DMF analizi yapmalarına rağmen ÇDKMTK modeli ile örtük grupların belirlenmesi sonrası tespit edilen DMF'li madde sayısının, gözlenen grupları kullanan olağan DMF analizinden beklenenden daha fazla olduğu vurgulamışlardır. Bunun nedeni olarak da örtük sınıf yaklaşımının örtük sınıflar arasındaki farklılıkları en yüksek düzeye çıkararak daha fazla sayıda DMF maddesi ve örtük sınıflar arasında madde güçlüklerinde daha büyük farklılıklar ortaya çıkarması gösterilmiştir.

Cohen ve Bolt'un (2005) çalışmalarında ilk olarak olabilirlik oranı yöntemi ile cinsiyete dayalı DMF belirlenmiş ve elde edilen parametre değerleri ile 3PL lojistik model kullanılarak örtük gruplar oluşturulup cinsiyetin grupları doğru ayırıcı bir değişken olmadığı gösterilmiştir. Daha sonra geniş örneklem üzerinde karma modele göre DMF analizi yapılmıştır. Araştırma sonunda örneklemere özgü olsa da Çalışma 1 ve 2'den elde edilen sonuçlar ile ikincil önemsiz boyutların (secondary nuisance dimensions) sınav katılımcılarının matematik testi maddelerine verdikleri yanıtlar üzerinde etkili olduğu belirtilmiştir.

Gürdil (2023) simülatif ve gerçek veri kullanarak Tek Boyutlu ve Çok Boyutlu Madde Tepki Kuramı kestirimlerinin DMF'yi kontrol altına almaya yönelik etkilerini incelediği çalışmasında tek ve çok boyutlu kestirimlere ait Tip I hata ve istatistiksel güç

oranları arasında, çok boyutlu kestirim lehine manidar farklılıklar olduğu sonucuna ulaşmıştır. Gerçek veri üzerinde lojistik regresyon, SIBTEST ve Lord'un χ^2 tekniklerinin Tip I hata ve istatistiksel güç oranlarına ait incelemeleri sonucunda her ikisinde de manidar farklar bulmuştur. Fakat farka dair ayrıntılandırma yapılmamıştır. Ayrıca çalışmada DMF'li madde ya da etki büyüklüğü gibi bilgiler verilmemiştir.

İkinci Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma

Karma Madde Tepki Kuramına, Çok Düzeyli Karma Madde Tepki Kuramına ve Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına göre yapılan analizlerde 2PL ve 3PL modellere göre en uygun olan örtük sınıf modeli, DMF'li madde sayıları ve etki büyüklükleri farklılaşmakta mıdır?

İkinci alt problemin çözümünde farklı koşullarda (boyut-düzye-ortak değişken-model) veri setine en uygun modeller için yapılan DMF analizi sonuçlarına göre Tablo 117 oluşturulmuştur. Tablo 21'de öğrenci düzeyinde DMF'li madde sayısı, DMF'li maddeler ve etki büyüklüğü düzeyleri gösterilmiştir.

Tablo 21

Veri Setine En Uygun Modeller için Öğrenci Düzeyinde DMF'li Madde Sayısı, DMF'li Maddeler ve Etki Büyüklüğü Düzeyleri

Model	DMF'li Madde Sayısı	DMF'li Maddeler ve Etki Büyüklüğü Düzeyleri (Örtük öğrenci sınıf çiftlerinin her ikisinde aynı düzeyde olan maddeler)			B ve C düzeyi DMF'li madde sayısı
		A	B	C	
Tek Boyutlu C3 2PL	17	2,5,6,16,18	8	1,3,7,9,10,11,13 14,15,17,19	12
Tek Boyutlu C3 3PL	19	-	-	3,12,16,17,19,20	13

Tek Boyutlu CB2-C2 2PL	19	-	-	3,12,16,17,19,20	19
Tek Boyutlu CB2-C2 3PL	19	-	-	11,14,15,18	13
İki Boyutlu C2 2PL	17	8,11,17	2,5,7,20	1,3,4,9,10,12 14,15,16,19	14
İki Boyutlu C3 3PL	17	9,19	4,10,11,15,16,17,20	1,2,5,6,7,8,12,13	15
İki Boyutlu CB2-C2 2PL	19	-	-	1,2,3,4,5,6,7,8,9 12,14,17,18,19,20	19
İki Boyutlu CB2-C2 3PL	17	16	-	1,2,3,4,6,7,8,9,10,11	12
Ortak Değişkenli C2 2PL	17	5,9,16,18,20	6	1,3,7,8,9,10,11,13,14 15,17,19	12
Ortak Değişkenli C2 3PL	15	6,8,9,18	7,16,17	1,3,10,11,13,14,15,19	11
Ortak Değişkenli CB-2-C2 2PL	19	-	-	1,2,4,5,6,7,8,9,10,11,12 13,14,15,16,18,19,20	19
Ortak Değişkenli CB-2-C2 3PL	18	-	-	1,4,6,7,8,9,11,12 14,15,18,19	18
İki Ortak Değişkenli CB2-C2 2PL	12	-	12,18	1,8,14	12
İki Ortak Değişkenli CB2-C2 3PL	8	11,16	2	-	2

Tablo 21'deki sonuçlara göre düzeyin ihmal edildiği tek boyutlu modellerde (Tek Boyutlu C3 2PL-3PL) DMF'li olmayan maddelerde ortak yoktur. DMF'li maddelerde ise 14 ortak madde mevcuttur ve C düzeyi etki büyüklüğüne sahip altı madde ortaktır. B ve C

düzeyi DMF'li madde sayıları 12 ve 13'tür. Düzeyin var olduğu tek boyutlu modellerde (Tek Boyutlu CB2-C2, 2PL-3PL) ortak DMF'li madde olmamasına rağmen B ve C düzeyi etki büyüklüğüne sahip madde sayıları 19 ve 13'tür. Düzeyin ihmal edildiği iki boyutlu modellerde (İki Boyutlu C2 2PL- C3 3PL) ise DMF'li madde sayısı eşittir. DMF'li olmayan bir ortak madde vardır ve DMF'li maddelerin 15 tanesinde ortak madde mevcuttur. B düzeyi etki büyüklüğüne sahip bir ortak madde var iken C düzeyi etki büyüklüğüne sahip iki madde ortaktır. B ve C düzeyi DMF'li madde sayıları 14 ve 15'tir. İki boyutlu ve iki düzeyli modellerde (İki Boyutlu CB2-C2, 2PL-3PL) DMF'li olmayan bir madde ortak iken 12 madde ortak olarak DMF'li bulunmuştur. Ortak DMF'li maddelerden C düzeyi etki büyüklüğüne sahip sekizer madde bulunmaktadır.

Öğrenci düzeyi ortak değişkene sahip modellerden tek boyutlu modellerde DMF'li 15 madde ortak iken DMF'li olmayan üç madde ortaktır. Ortak DMF'li maddelerden ikişer tanesi A düzeyi etki büyüklüğüne sahip iken C düzeyi etki büyüklüğüne sahip sekizer madde bulunmaktadır. Öğrenci düzeyi ortak değişkene sahip modellerden iki düzeyli modellerde ise 17 madde ortak olarak DMF'li bulunmuştur. Ortak DMF'li maddelerden C düzeyi etki büyüklüğüne sahip 12'şer madde her iki modelde de bulunmaktadır. Öğrenci ve ülke düzeyi ortak değişkene sahip modellerde DMF'li sekiz madde ortak iken DMF'li olmayan bir madde ortaktır. Ortak DMF'li madde bulunmamakla birlikte B ve C düzeyi DMF'li madde sayıları 12 ve 2'dir. Tablo 22'de ise iki düzeyli modellerin ikinci düzeyi olan ülke düzeyinde DMF'li madde sayıları ve etki büyüklüğü düzeyleri verilmiştir.

Tablo 22

Veri Setine En Uygun İki Düzeyli Modeller için Ülke Düzeyinde DMF'li Madde Sayısı, DMF'li Maddeler ve Etki Büyüklüğü Düzeyleri

Model	DMF'li Madde Sayısı	DMF'li Maddeler ve Etki Büyüklüğü Düzeyleri	B ve C düzeyi DMF'li madde sayısı
-------	---------------------	---	-----------------------------------

		A	B	C	
Tek Boyutlu CB2-C2 2PL	11	1,8,18	10,12	3,11,14,15,16,19	8
Tek Boyutlu CB2-C2 3PL	16	13,14	17,20	1,2,3,4,5,6,7,8,9 10,11,18	14
İki Boyutlu CB2-C2 2PL	13	1,5,7,16	8,9,15,19	3,10,11,12,14	9
İki Boyutlu CB2-C2 3PL	15	1,4,6,8,17,18	2,9,12	3,5,10,14,16,19	9
Ortak Değişkenli CB-2-C2 2PL	17	10,12,16	2,17,19	1,3,4,5,6,7,8,10,11,18,20	14
Ortak Değişkenli CB-2-C2 3PL	14	2,4,5,17,20	10,11,18	3,12,14,15,16,19	9
İki Ortak Değişkenli CB2-C2 2PL	14	17	5,8,10,11,12,13,16	3,14,15,18,19,20	13
İki Ortak Değişkenli CB2-C2 3PL	18	3,5,8,10,11,13	1,4,6,7,9	2,12,14,15,16,17,19	12

Tablo 22'ye göre ülke düzeyi tek boyutlu modellerde yedi madde ortak olarak DMF'li bulunmuştur. Ortak DMF'li maddelerden C düzeyi etki büyüklüğüne sahip ikişer madde bulunmaktadır. İki boyutlu modellerde ise DMF'li olarak bulunmayan bir ortak madde vardır. Ortak olarak DMF'li bulunan madde sayısı ise 12'dir. Bu 12 maddeden A ve B düzeyi etki büyüklüğüne sahip birer madde, C düzeyi etki büyüklüğüne sahip üçer madde ortak olarak bulunmuştur. Öğrenci düzeyi ortak değişkenin yer aldığı iki modelde DMF'li olarak bulunmayan bir ortak madde vardır. Ortak olarak DMF'li bulunan 12 maddenin birer tanesi

C düzeyi etki büyüklüğüne sahiptir. İki düzey için de ortak değişkenin yer aldığı modellerde ise 12 madde ortak olarak DMF'li bulunmuştur. Ortak DMF'li maddelerden üçer C düzeyi etki büyüklüğüne sahiptir ve B ile C düzeyi etki büyüklüğüne sahip madde sayıları sırasıyla 13 ve 12'dir.

Öğrenci ve ülke düzeylerindeki modellere ait incelemelerde 2PL ve 3 PL modeller arasında bir uyum olmadığı görülmektedir. Öğrenci düzeyinde parametre sayısının artması ile DMF'li madde sayısında azalma meydana gelmiştir. Ülke düzeyinde ise parametre sayısının artması ile DMF'li madde sayısında artma olmuştur. Öğrenci düzeyinde 2 PL modeller ile 3 PL modellere göre daha fazla sayıda DMF'li madde bulunmuştur. Aynı zamanda B ve C düzeyi etki büyüklüğüne sahip DMF'li madde sayısında da 2 PL modeller daha fazladır. Ülke düzeyinde genel olarak 2 PL modellerde 3 PL modellere göre daha az sayıda DMF'li madde yer almaktadır. Ancak B ve C düzeyi etki büyüklüğüne sahip DMF'li maddelerde durum tersi şekildedir. Simülatif veriler ile yapılan iki çalışmada (Chen, 2015; Li ve diğerleri, 2009) ise KMTK modellerinde örtük sınıf sayısı arttığında model uyumundan (1PL, 2PL, 3PL) bağımsız bir şekilde madde parametrelerinin kestirim doğruluğunda düşüş görüldüğü sonucuna ulaşılmıştır.

Üçüncü Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma

Veriye uyarlanan en uygun modellere göre ortaya çıkan örtük sınıflarda cinsiyet, ülke ve madde parametrelerine ilişkin özelliklerin dağılımı ve bulgular farklılaşmakta mıdır?

Üçüncü alt problemin çözümünde veri setine uygun modellerin belirlenmesi ve bu modellere ait bazı betimsel istatistiklerin eldesi gerekmektedir. Boyut, düzey ve ortak değişkene göre yapılan analizler sonucunda veriye en uygun modeller Tablo 23'te verilmiştir.

Tablo 23

Veri Setine En Uygun Modellere Ait Bilgi Ölçütü İndeksleri

Modeller	LL	AIC	BIC	SABIC
Tek Boyut 2PL CB2-C2	-94994,12	190432,24	191983,03	191277,56
İki Boyut 2PL CB2-C2	-94942,28	190286,57	191690,67	191051,93
Birinci Düzeyde Ortak Değişkenli 2PL CB2-C2	-93815,46	188080,91	189651,08	188936,75
İki Düzeyde Ortak Değişkenli 2PL CB2-C2	-93790,82	188033,63	189610,78	188892,59

Tablo 23'e göre BIC değerleri baz alınarak seçilen modellerden iki boyutlu model tek boyutlu modelden, öğrenci düzeyine ortak değişkenin eklendiği model iki boyutlu ve iki düzeyliden ve iki düzeye de ortak değişkenlerin eklendiği iki boyutlu iki düzeyli 2 PL model bir ortak değişkenli modelden verinin yapısına daha uygundur. Çoktan seçmeli bir veri setinde 3 PL modelin daha iyi uyum sağlayacağı tahmin edilmesine rağmen genel olarak 2 PL modeller bu veri setine uyumlu bulunmuştur. Bu çalışmanın aksine Jensuttiwetchakul vd. (2016) çalışmalarında geliştirilen 3 PL modelin okul örtük sınıfının sınıflandırılması ve ölçülmesinin etkinliğini artırdığını belirtmişlerdir. Grup büyüklükleri önemli ölçüde farklılık gösterdiğinde bile, örtük sınıflar için DMF tespiti DMF maddelerini belirlemede iyi performans göstermektedir (Maij-de Meij ve diğerleri, 2010). Çok boyutlu kestirime ait sonuçlar tüm DMF belirleme yöntemlerinde tek boyutlu kestirimden daha düşük Tip I hata ve daha yüksek istatistiksel güç oranları elde edilmiştir (Gürdil, 2023). Finch ve Finch de (2013) yaptıkları analiz sonuçlarının, çok boyutlu modelin DMF'nin doğası hakkında tek boyutlu ayrı modellere göre daha eksiksiz bilgi sağladığını göstermekte olduğunu belirtmişlerdir.

Tüm modellerde hem birinci hem de ikinci düzey örtük sınıflara dağılan erkek ve kız öğrenci sayıları birbirine çok yakındır. Bu sebeple bu veri setine uygun olarak örtük sınıfların oluşumunda cinsiyetin etkisinin olmadığı söylenebilir. Bu alanda yapılan birçok çalışmada (Ayan, 2011; Demir, 2013; Taşdelen Teker, 2014; Tiryaki, 2019 vb.) cinsiyet ilk analiz edilen değişken olmaktadır. Bu durum her araştırma için yanlış yapılmıştır anlamına gelmemekle

birlikte öncelikle araştırmanın amacının irdelenip ardından veri setinin betimsel incelemesinin yapılması sonucu verilecek kararın önemini göstermektedir. Tablo 119'da yer alan bütün modeller hem 1. düzey hem de 2. düzeyde iki örtük sınıftan oluşmaktadır ve örtük sınıflar arasında toplam puanlarda farklılıklar mevcuttur. Bu sebeple de ülkelerin karşılaştırıldığı Tablo 24'te örtük sınıflar puanı yüksek ve düşük olan şekilde isimlendirilmiştir.

Tablo 24

Seçilen İki Düzeyli Modellerden Oluşturulan Ülkelere Ait Sınıflar

Model	Puanı Yüksek Örtük Sınıf	Puanı Düşük Örtük Sınıf
Tek Boyut 2PL CB-C2	Singapur, Tayvan, Hong Kong, Kore, Türkiye, İngiltere, Katar, Portekiz BAE, Gürcistan, İsrail, Şili, Malezya, ABD	İtalya, Litvanya, Fransa, Rusya, İsveç, Norveç, Macaristan, Finlandiya
İki Boyut 2PL CB-C2	Singapur, Tayvan, Hong Kong, Kore, Türkiye, İngiltere, Katar, Portekiz BAE, Gürcistan, İsrail, Şili, Malezya, ABD, Litvanya, Rusya	Finlandiya, Fransa, Macaristan, İtalya, Norveç, İsveç
Ortak Değişkenli 2PL CB2-C2	Tayvan, Finlandiya, Hong Kong, Kore, Norveç, Singapur, İsveç	Şili, Fransa, Gürcistan, Macaristan, İsrail, İtalya, Litvanya, Malezya, Portekiz, Rusya, BAE, Türkiye, ABD, İngiltere, Katar
İki Düzeyde Ortak Değişkenli 2PL CB2-C2	Tayvan, Hong Kong, Kore, Malezya, Singapur, İsveç, Norveç İngiltere	Şili, Finlandiya, Fransa, Macaristan, İsrail, İtalya, Litvanya, Katar, Portekiz, Rusya, Gürcistan, Türkiye, ABD, BAE

Tablo 24'e göre modellerdeki puanı yüksek olan örtük sınıfların tamamında Tayvan, Kore, Hong Kong ve Singapur bulunmaktadır. Modellerdeki puanı düşük olan örtük sınıfların tamamında ise Fransa, İtalya ve Macaristan bulunmaktadır. Tablo 25'te tek boyutlu modellerde en iyi uyum değerine sahip olan model ile tüm modeller arasında en iyi uyum değerine sahip olan iki düzeye ortak değişkenin eklendiği modelin madde güçlükleri karşılaştırılmıştır.

Tablo 25

Tek Boyutlu CB2-C2 2 PL Modele ve İki Düzeye Ortak Değişkenin Eklendiği Modele Ait Madde Güçlük İndeksleri

Madde No	Tek Boyutlu CB2-C2 2PL Model Öğrenci Düzeyleri				İki Düzeye Ortak Değişkenin Eklendiği 2PL Model Öğrenci Düzeyleri			
	ÖÜ1-1	ÖÜ1-2	ÖÜ2-1	ÖÜ2-2	ÖÜ1-1	ÖÜ1-2	ÖÜ2-1	ÖÜ2-2
	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
1	2,04	0,99	-4,60	-5,86	3,07*	0,71	-2,85	-3,88
2	2,10	1,48	4,05	0,05	2,05	1,05	-1,91	-0,51
3	0,77*	1,31	-4,45	1,40	1,16	-0,09	-0,63	-0,85
4	0,69	1,88	2,65	3,12	0,64	-0,11	2,01	1,68
5	1,95	0,33	-3,70	-1,38	4,83	0,70	0,35	1,39
6	0,58	1,51	1,50	5,80	1,81	-0,46	4,88	2,43
7	3,30	0,75	4,67	-1,60	3,93	0,62	1,12	1,49
8	1,65	-0,98	-3,62	-4,57	2,59*	0,10	-2,03	-1,45
9	0,15	2,41	0,64	2,07	0,13	-0,40	2,11	1,53
10	1,32*	1,13	3,97	1,47	1,38	0,17	0,33	0,59
11	0,20*	3,81	3,31	-0,08	0,46	-0,91	0,91	1,52
12	1,63*	5,47	0,09	-1,46	0,23*	0,15	1,56	2,71
13	3,18	3,04	4,11	1,54	2,06	1,55	0,84	0,05
14	1,98*	3,09	0,28	-1,20	0,92*	0,54	3,97	-4,74
15	0,35*	4,83	-1,28	-2,64	-0,22	-0,28	5,80	-4,10
16	0,57*	1,79	0,36	-0,89	0,28	0,27	3,19	-1,05
17	1,43	1,93	2,48	-1,02	1,01	-0,08	1,63	-0,92
18	2,78	0,44	1,75	-0,31	2,84*	0,32	0,10	-1,14
19	1,54*	4,76	-0,17	-1,30	0,45	0,06	-2,45	-2,82
20	1,69	2,29	0,89	-1,40	0,89	-0,46	-0,79	-2,59

Tablo 25'te madde güçlükleri her iki model için de puanı en düşük olan sınıftan en yüksek olana doğru sıralanmıştır. İki modele ait madde güçlükleri karşılaştırıldığında veriye uyum sağlayan modellerin arasında madde güçlük indekslerinde uyum olmadığı görülmektedir. Bunun sebebi de örtük sınıfların oluşturulmasında düzey, boyut, ortak değişken gibi etkenlerin rol almasıdır. Bu durum DMF'li maddeleri de farklılaştırmaktadır. Yanlarında işaret olan maddeler modellerde yer alan her iki örtük düzeyde de B ve C düzeyi DMF'li olanlardır. 12 ve 14. maddeler her iki model için de ortak olarak DMF'lidir. Örneğin hem 12. hem de 14. maddenin güçlük indeksi tek boyutlu model için başarı düzeylerine göre olması gerekenden farklı şekilde değişmiştir.

Dördüncü Alt Probleme İlişkin Bulgular, Yorumlar ve Tartışma

Ortak değişkenlerin dahil olduğu Çok Boyutlu Çok Düzeyli Karma Madde Tepki Kuramına göre yapılan DMF analizlerinde DMF'nin kaynağına belirlemeye yönelik zengin içerikli çıkarımlar yapılabilen midir?

Dördüncü alt problemin cevaplanması için veri setine en iyi uyumu gösteren iki düzeyine de ortak değişkenin dahil edildiği CB2-C2 modeli kullanılmıştır. CB2-C2 modeline göre oluşturulan Tablo 26'da örtük sınıflara göre birey sayıları matematik ve fen bilimleri derslerine ait ortalamalar, bu dersler arasındaki korelasyon ve evdeki kitap sayısı (EKS) değişkenine ait ortalamalar verilmiştir.

Tablo 26

Örtük Sınıflara Ait Birey Sayıları, Ortalamalar ve Korelasyon Değerleri ile EKS

Ortalamaları

Örtük Sınıf	Model	Birey Sayısı	Matematik Ortalaması	Fen Bilimleri Ortalaması	Korelasyon Değerleri	Genel Ortalama	EKS Ortalaması
Öğrenci Düzeyi	ÖÜ1-1	1443	4,63	4,81	0,36	9,44	2,29
	ÖÜ1-2	1048	8,00	7,75	0,45	15,75	3,60
	ÖÜ2-1	3474	2,94	4,03	0,18	6,97	2,30

	ÖÜ2-2	1966	7,07	7,19	0,20	14,26	3,42
Ülke Düzeyi	ÖÜ1	2491	6,05	6,04	0,59	12,09	2,84
	ÖÜ2	5440	4,43	5,17	0,56	9,60	2,71

Tablo 26'ya göre EKS değişkenine ait ortalamalar incelendiğinde öğrenci başarıları ile doğru orantılı ve örtük sınıfların oluşumunda etkili olduğu görülmektedir. Öğrenci düzeyinde en başarılı örtük sınıf hem matematik hem de fen bilimleri dersi için ÖÜ1-2'dir. Ardından ÖÜ2-2 ve ÖÜ1-1 gelirken en başarısız örtük sınıf ÖÜ2-1'dir. ÖÜ1 ülke düzeyi örtük sınıfı ÖÜ2 ülke düzeyi örtük sınıfına göre daha başarılıdır. Üç örtük sınıfta fen bilimleri başarıları matematiğe göre fazla iken ÖÜ1-2 modelinde matematik başarıları daha fazladır. Bu çalışmada öğrenci düzeyi örtük sınıflar için her iki derste de başarılı ve başarısız olanlar ayrılmıştır. Ancak Finch ve Finch'in (2013) çalışmasında öğrenci düzeyi üç alt boyut yer almakta ve örtük sınıf 1 hem dil hem de matematik sınavlarında yüksek puanlar alırken örtük sınıf 2 bireyler her iki test için de nispeten düşük ortalamalara sahiptir. Buna karşılık, örtük sınıf 3'teki bireyler dil testinde en düşük, matematikte ise en yüksek ortalama puana sahip olmuştur. Bu şekilde İki madde setine dayalı örtük sınıfların ayrı ayrı tahmin edilmesini, yapıların çok boyutlu modele birlikte dahil edildiğinde ortaya çıkan ile aynı çok yönlü örüntüyü ortaya koymadığını belirtmişlerdir. Tablo 27'de ikinci düzey olan ülke düzeyindeki örtük sınıflarda yer alan öğrencilerin dağılımı verilmiştir. Parantez içlerinde yer alan sayılar birinci düzey örtük sınıflarda yer alan birey sayılarını göstermektedir.

Tablo 27

Ülke Düzeyi Örtük Sınıflardaki Öğrencilerin Örtük Sınıflardaki Sayıları

Ülke Düzeyi Örtük Sınıflar	
Örtük Sınıf 1 (ÖÜ1-1/ÖÜ1-2)	Örtük Sınıf 2(ÖÜ2-1/ÖÜ2-2)
Tayvan (181/163)	Şili (238/50)
Hong Kong (158/70)	Finlandiya (176/165)
Kore (94/178)	Fransa (197/72)

Malezya (383/114)	Gürcistan (176/63)
Norveç (142/146)	Macaristan (142/178)
Singapur (181/169)	İsrail (153/107)
İsveç (133/146)	İtalya (145/108)
İngiltere (171/62)	Litvanya (152/103)
	Portekiz (150/87)
	Katar (198/69)
	Rusya (124/153)
	BAE (1138/425)
	Türkiye (154/129)
	ABD (331/257)

Tablo 27'ye göre ortalama puanı fazla olan örtük ülke sınıfındaki ülkelerden Tayvan, Hong Kong, Kore ve Singapur TIMMS'de en başarılı ülkeler arasında yer almaktadır. Ayrıca Norveç ve İngiltere matematik bölümünde ortalamanın üzerinde yer almaktadırlar. Ancak Malezya'nın bu örtük sınıfta neden yer aldığıyla ilgili eldeki çıktılarla bir açıklama yapılamamaktadır. Ortalama puanı daha az olan örtük sınıftaki ülkeler genel olarak ortalamanın altında yer almaktadır. Fakat Finlandiya, Rusya ve Macaristan gibi hem matematik hem de fen bölümlerinde ortalama üzerinde yer alan ülkeler de bu gruptadır. Düzeyin eklendiği KMTK modelleri ile ülkeler arası karşılaştırma ve bağıl değerlendirme yapılabilmektedir. Gürkan (2021) çalışmasında gelişmiş modelleme tekniklerinin yalnızca verilerin özellikleri göz önüne alındığında daha uygun olmakla kalmayıp, aynı zamanda ülkeler arasındaki ilişki kalıpları hakkında daha fazla bilgi sağladığını ifade etmiştir. Tablo 28'de ÖÜ2-2 öğrenci düzeyleri örtük sınıflarında ortak olarak yer alan B ve C düzeyi etki büyüklüğüne sahip DMF'li maddeler, bu maddelerin ait olduğu test bölümleri, konu alanları ve bilişsel alanları verilmiştir. DMF'li maddelerden odak grup lehine olanlar (+), referans grup lehine olanlar (-) işaretleri ile gösterilmiştir. İlk sıradaki işaret ÖÜ1 içerisinde yer alan örtük öğrenci sınıfları arasındaki karşılaştırma için iken ikinci sıradaki işaret ÖÜ2 içerisinde

yer alan örtük öğrenci sınıfları arasındaki karşılaştırma içindir. Örneğin 1. maddede her iki örtük ülke sınıflarında yer alan örtük öğrenci sınıfları için madde referans grup lehine DMF'li bulunmuştur.

Tablo 28

Öğrenci düzeyi B ve C Düzeyi DMF'li Maddelere Ait Test Bölümleri Konu Alanları ve Bilişsel Alanlar

B ve C düzeyi etki büyüklüğüne sahip DMF'li maddeler		Test Bölümü	Konu Alanı	Bilişsel Alan
Madde No	Odak Grup			
1	-/-	Matematik	Sayılar	Bilme
6	-/-	Matematik	Geometri	Uygulama
8	+/-	Matematik	Sayılar	Uygulama
9	+/-	Matematik	Cebir	Bilme
11	+/-	Fen Bilimleri	Biyoloji	Bilme
12	-/-	Fen Bilimleri	Biyoloji	Bilme
14	-/-	Fen Bilimleri	Fizik	Bilme
17	+/-	Fen Bilimleri	Biyoloji	Bilme
18	+/-	Fen Bilimleri	Kimya	Bilme
19	-/-	Fen Bilimleri	Yeryüzü Bilimleri	Akıl Yürütme

Tablo 28'e göre matematik ve fen bilimleri bölümlerde ait maddelerde dört ve altı madde B ile C düzeyi etki büyüklüğüne sahip DMF göstermektedir. Matematikte iki madde sayılar konusunda iken bir madde cebir bir madde ise geometri konusundadır. Bilişsel alan olarak ise iki tanesi uygulama alanında iken iki tanesi bilme alanındadır. Fen bilimlerinde üç madde biyoloji birer madde fizik, kimya ve yeryüzü bilimleri konu alanlarındadır. Ayrıca maddelerin biri hariç tamamı bilme bilişsel alanına aittir. Buna göre öğrencilerin özellikle bilme bilişsel alanında oneli farklılıklara sahip oldukları görülmektedir. Bu farklılığın

giderilebilmesi için de çeşitli öğretim planları yapılmalıdır. Cho ve Cohen (2010) ile Jensuttiwetchakul vd. (2016) ÇDKMTK ile ilgili çalışmalarında modelin öğrenci düzeyinin, soruları yanıtlamaya yönelik tepki stratejilerinde farklılık gösteren örtük sınıfların var olup olmadığını belirleme fırsatı sağlayacağını belirtmişlerdir. Ayrıca okul düzeyindeki bilgilerin, örtük sınıflar arasındaki müfredat veya pedagojik farklılıklardan kaynaklanabilecek olası farklılıkları ortaya çıkarmak için kullanılabileceğini vurgulamışlardır. Odak ve referans gruplar için yapılan incelemede ÖÜ1 içerisinde yer alan örtük öğrenci sınıflarında beş madde odak grup lehine DMF gösterirken kalan beş madde ise referans grup lehine DMF göstermiştir. ÖÜ2 içerisinde yer alan örtük öğrenci sınıflarında tüm maddeler referans grup lehine DMF göstermiştir. Tablo 29'da ise ülke düzeyi B ve C düzeyi etki büyüklüğüne sahip DMF gösteren maddelere ait bilgiler yer almaktadır.

Tablo 29

Ülke düzeyi B ve C Düzeyi DMF'li Maddelere Ait Test Bölümleri Konu Alanları ve Bilişsel Alanlar

B ve C düzeyi etki büyüklüğüne sahip DMF'li maddeler		Test Bölümü	Konu Alanı	Bilişsel Alan
Madde No	Odak Grup			
3	-	Matematik	Cebir	Uygulama
5	-	Matematik	Geometri	Uygulama
8	-	Matematik	Sayılar	Uygulama
10	-	Matematik	Cebir	Bilme
11	-	Fen Bilimleri	Biyoloji	Bilme
12	+	Fen Bilimleri	Biyoloji	Bilme
13	+	Fen Bilimleri	Fizik	Bilme
14	+	Fen Bilimleri	Fizik	Bilme
15	+	Fen Bilimleri	Yeryüzü Bilimleri	Bilme
16	+	Fen Bilimleri	Biyoloji	Bilme
18	-	Fen Bilimleri	Kimya	Bilme

19	+	Fen Bilimleri	Yeryüzü Bilimleri	Akıl Yürütme
20	-	Fen Bilimleri	Yeryüzü Bilimleri	Uygulama

Tablo 29'a göre ülke düzeyinde incelendiğinde matematik ve fen bilimleri bölümlerde ait maddelerde dört ve dokuz madde B ile C düzeyi etki büyüklüğüne sahip DMF göstermektedir. Matematikte iki madde cebir konusunda iken birer madde sayılar ve geometri konusundadır. Bilişsel alan olarak ise üç tanesi uygulama alanında iken bir tanesi bilme alanındadır. Fen bilimlerinde üçer madde biyoloji ve yeryüzü bilimleri, iki madde fizik ve bir madde de kimya konu alanındadır. Ayrıca maddelerin ikisi hariç tamamı bilme bilişsel alanına aittir. Özellikle fen bilimleri bölümünde hem öğrenci hem de ülke düzeyi incelemede bilme bilişsel alanında önemli farklılaşma olduğu görülmektedir. Bu durumu her ülke kendi ölçeğinde inceleyip gereken önlemleri alırsa testte adillik adına önemli bir adım atılmış olur. ÇBÇDKMTK modelleri, oluşturulan ikinci düzeylerle (örneğin okul düzeyi) hedef okulla aynı örtük sınıfta olmayan diğer okullarla ilişkili özelliklere kıyasla, söz konusu okulla ilişkili öğrenci ve okul düzeyindeki özelliklerin açıklamalarını kullanarak bir hedef okulun farklı madde performansını tanımlamayı mümkün kılar. Bu tanımlama daha sonra okullara, okullarının sonuçlarını kendi örtük sınıflarındaki ve diğer örtük sınıflardaki diğer okullarla karşılaştırabilecekleri bir çerçeveye sağlamak için kullanılabilir (Cho & Cohen, 2010). Finch ve Finch'in (2013) çalışmalarının sonucunda tek bir sınav oturumunda birden fazla yapı değerlendirildiğinde, ÇBÇDKMTK'nın birden fazla boyutu aynı anda inceleyerek hem Düzey 1'de (sınava giren kişi) hem de Düzey 2'de (okul) tek biçimli DMF hakkında daha fazla bilgi sağlayabileceğini göstermekte olduğunu ifade etmişlerdir. Bunun yanında birden fazla boyutun dahil edilmesi, öğrencilerin öğrendikleri ve değerlendirildikleri gerçek dünya bağlamlarını yansıtacak şekilde, birden fazla yapıdaki göreceli yeterliliklerini eş zamanlı olarak kullanarak sınava giren ve okul sınıflarının daha eksiksiz bir karakterizasyonunu sağladığı belirtilmiştir. Odak ve referans gruplar için yapılan altı madde odak grup lehine DMF gösterirken yedi madde ise referans grup lehine DMF göstermiştir.

Tablo 28 ve Tablo 29 öğrenci ve ülke düzeyi adına önemli çıktılar sağlamaktadır. Bu da kullanılan yöntemin bir başarısıdır. Uyar (2015) simülatif veri kullandığı tezinde verilerin ÇDKMTK modeline uygun yapıya sahip olması durumunda ÇDKMTK'nın gücünün gözlenen gruplarla DMF belirlemeye göre daha yüksek olduğunu bulmuştur. Bu sebeple araştırmacı uygun model kullanıldığında uzmanların B ve C düzeyinde DMF gösteren maddeleri yorumlamasının daha kolay olacağını ve maddelerin yanlılık gösterme nedenlerinin daha nesnel olarak belirlenmesini sağlayacağını ifade etmiştir. Çalışma kapsamında veri yapısına en uygun model ortak değişkenlerin eklendiği ÇBÇDKMTK modeli olduğu için maddelerin DMF gösterme sebepleri daha kolay anlaşılabilirliği vurgulanmıştır. Finch ve Finch (2013) ÇBÇDKMTK'nın çok boyutlu değerlendirme verileri mevcut olduğunda ölçme uzmanının araç kutusuna benzersiz bir katkı sağladığı inandıklarını belirtmişlerdir.

Model oluşturulurken ortak değişkenin seçimi de önemli bir etken olarak öne çıkar ve DMF'nin doğasının anlaşılmasında katkı sunar. DMF etki büyüklüğünün yüksek ve ortak değişken ile örtük değişken arasındaki ilişkinin güçlü ve örtük sınıf oranının eşit olduğu durumda ortak değişkenli karma modelin sınıflama doğruluğunun karma model ile aynı seviyeye ulaştığı Sırgancı (2019) tarafından belirlenmiştir. Bu durumda doğru seçilen ortak değişken ile DMF hakkında daha fazla bilgi sağlanıp sınıflama doğruluğu korunmuş olabilir.

Maij-de Meij vd. (2010) KMTK modellerinin hem grup içi hem de gruplar arası heterojenliği tanımladığını ve bu şekilde bazı maddelerin DMF'nin gerçek kaynağıyla ilişkili farklı işlevler gösterebileceği homojen alt grupları belirleyebileceğini ifade etmiştir. DMF'nin doğasını anlamak ve DMF kaynaklarını ortaya çıkarmaya çabalamak elde edilen çıktıların operasyonel olarak işe koşulabilecek şekilde olması sebebiyle kıymetlidir. Ceylan'ın (2022) araştırmasında DMF süreçleri tersten işletilmiş ancak DMF'li olarak yazılan maddelerle DMF belirleme yöntemlerinin işaretlediği maddeler farklılık göstermiştir. Bunun sebebi de bireylerin cinsiyet, ırk yaşanan yere göre sınıflandırıldığında homojenliğin sağlanamamasıdır. Kırsal kesimde yaşayan her bireyin kente ait lokasyon bilgilerinden

mahrum olduđunun öngörüsü ile sınıflanması analiz sonuçlarını etkilemiştir. Bu araştırma örtük sınıfların oluşturulup analizlerin örtük sınıflar üzerinden gerçekleştirilmesinin önemini göstermektedir. Ayrıca DMF'li maddelerin önsel ya da sonsal olarak kesin bir şekilde belirlenmesinin pek mümkün olmadığı inancı bu tür çalışmalarla pekişmektedir. Bu yüzden DMF'nin kaynaklarını önceden belirleme süreci Zumbo'ya (2007) göre DMF analizinin üçüncü neslinin de (hem nitel hem de nicel analiz süreçleri kullanımı) ötesine geçmiş olabilir.

Bölüm 5

Sonuç ve Öneriler

Bu bölümde araştırmadan elde edilen sonuçlara ve bu sonuçlara ilişkin geliştirilen önerilere yer verilmiştir.

Sonuçlar

Bu araştırmanın amacı cinsiyet ve evdeki kitap sayısı olarak belirlenen değişkenler ile oluşturulan gözlenen gruplar ile KMTK, ÇDKMTK, ÇBÇDKMTK ve odÇBÇDKMTK modelleri ile oluşturulan örtük gruplar için DMF'li maddeleri belirleme ve DMF'nin olası kaynaklarını açıklayabilme performanslarını karşılaştırmaktır. Bu kapsamda gözlenen değişkenlere ait gruplar için MH yöntemiyle, KMTK modellerine ait düzey, boyut ve ortak değişkenin dahil edilmesiyle elde edilen örtük gruplar için MH ve GMH yöntemleriyle DMF analizi yapılmıştır. Ayrıca gözlenen ve örtük gruplara ait betimsel istatistikler hesaplanıp karşılaştırmalar yapılmıştır.

Gözlenen grup değişkenleri olan cinsiyet ve evdeki kitap sayısı değişkenlerine göre yapılan DMF analizi sonucunda B ve C düzeyi etki büyüklüğüne sahip madde bulunmamıştır. KMTK ile oluşturulan örtük sınıflar için yapılan DMF analizine göre ise 5 ile 19 arasında değişen B ve C düzeyi etki büyüklüğüne sahip DMF'li maddeler bulunmuştur. Bu durum gözlenen değişkenler kullanılarak yapılan analizlerin bireyler arasında var olan farklılıkları ortaya çıkaramayabileceğini göstermektedir. Sadece gözlenen gruplar ile çalışmak ya da basmakalıp yargılarla bireyleri sınıflandırmak hatalı bulgular ortaya koyabilir.

Seçilen veri setinde KMTK 2PL modeller ile oluşturulan örtük sınıflar için boyut ve düzey eklenmesi, gözlenen gruplara göre yapılan analiz sonuçlarına göre çıktıları oldukça farklılaştırırsa da kendi içlerinde büyük farklılıklar oluşturmamaktadır. Analizlere ortak değişkenlerin eklenmesi ile DMF'li madde sayısı aralığı büyümüştür. Ayrıca B ile C düzeyi etki büyüklüğüne sahip madde sayısı, özellikle her iki düzeye de ortak değişkenin eklenmesi

ile azalmıştır. Bunun sebebi ise analizler veri setine bağlı olduğu için net olarak açıklanamaz.

2PL ve 3 PL modeller için DMF'li madde sayıları, etki büyüklüğü düzeyleri ve ortak DMF'li maddeler değişmiştir. Bu durum model parametre seçiminin etkisini göstermektedir. Öğrenci düzeyi modellerde parametre sayısının artışına bağlı olarak özellikle iki düzeyde ortak değişkenin yer aldığı modellerde DMF'li madde sayısında azalma meydana gelmiştir. Ülke düzeyinde ise parametre sayısının artması ile yine özellikle ortak değişkenin yer aldığı modellerde DMF'li madde sayısında artma görülmüştür.

Öğrenci düzeyinde 2 PL modeller aracılığıyla elde edilen DMF'li madde sayısı 3 PL modeller ile elde edilenlere göre daha fazla sayıdadır. Ayrıca B ve C düzeyi etki büyüklüğüne sahip DMF'li madde sayısı olarak da 2 PL modellerde 3PL modellerden sayı olarak daha fazladır. Ülke düzeyinde 2 PL modeller ile yapılan DMF analizi sonuçlarında 3 PL modellere göre daha az sayıda DMF'li madde yer almaktadır. Bununla birlikte B ve C düzeyi etki büyüklüğüne sahip DMF'li toplam madde sayılarında 2 PL modeller daha fazladır.

odÇBÇDKMTK modelleri ile öğrenciler arası farklılıkların incelenmesinin yanında ülkeler arası kıyaslamalar yapılabilmekte, eksiklikler veya başarılı alanlar fark edilebilmekte ve karşılaştırmalı değerlendirmeler yürütülebilmektedir. Bu çıktılar ülkelere müfredat güncelleme, öğretim programını düzenleme, bu çalışma özelinde matematik ve fen bilimlerine ait öğretim yöntem ve tekniklerini gözden geçirme imkânı sağlayabilir. Öğrenci düzeyinde elde edilen çıktılarda matematik ve fen bilimleri bölümlerinde belli konu alanlarındaki farklılaşmalar göze çarpmakta olup bunların çoğunluğu bilgi düzeyi bilişsel alanına aittir. Bu çıkarım ile ilgili tüm paydaşlar kendi düzenlemelerini yapmalıdır. Örneğin bu farklılaşmanın kapsam geçerliği sağlanamaması sebebiyle olup olmadığı test geliştiricileri tarafından, bilginin (doğru olarak) aktarılıp aktarılmadığı ya da uygun öğrenme/öğretme stillerinin kullanılıp kullanılmadığı öğretmenler tarafından, öğretim

programında, ilgili bilgiye ait kazanımın yeri, ilgili ders süresi, diğer kazanımlarla ilişkilendirilmesi gibi hususlar ise alan uzmanları tarafından tartışılabilir.

Odak ve referans grup incelemesi yapıldığında ÖÜ1 için örtük öğrenci sınıflarında matematik bölümündeki DMF'li olan dört maddenin ikisinin odak, ikisinin ise referans grup lehine olduğu görülmektedir. Dört maddenin ikisi bilme düzeyinde iken ikisi uygulama düzeyindedir. Fen bilimleri bölümündeki DMF'li olan altı maddenin üçü odak, üçü ise referans grup lehinedir. Bilişsel alanlara göre ÖÜ1'de yer alan gruplarda bilme ve uygulama alanında önemli bir farklılaşma olmamıştır. Sadece fen bilimleri bölümündeki akıl yürütme alanındaki bir madde referans grup lehinedir. ÖÜ2 için örtük öğrenci sınıflarında matematik ve fen bilimleri bölümlerindeki DMF'li olan 10 maddenin tamamının referans grup lehine olduğu görülmektedir. Ülke düzeyinde yer alan iki örtük ülke sınıfına ait örtük öğrenci sınıflarından ÖÜ1 içerisinde yer alanlarda önemli bir farklılaşma görülmemektedir. Ancak testteki maddeler ÖÜ2 içerisinde yer alan referans grup lehine odak gruba göre bariz şekilde farklılaşmıştır.

Ülke düzeyinde yapılan odak ve referans grup incelemesinde ortalaması düşük olan örtük ülke sınıfı odak grup olarak seçilmiş ve DMF'li olarak belirlenen 13 maddenin altı tanesinin odak grup lehine olduğu görülmüştür. Matematik bölümünde yer alan dört DMF'li maddenin tamamı referans grup lehinedir ve üç tanesi uygulama, bir tanesi bilme bilişsel alanına aittir. Fen bilimleri bölümünde yer alan dokuz maddenin altı tanesi odak grup lehinedir ve bu maddelerin beş tanesi bilme bilişsel alanına aittir. Referans grup lehine olan üç maddenin bir tanesi uygulama bilişsel alanında iken iki tanesi bilme bilişsel alanındadır. Örtük ülke sınıfları karşılaştırıldığında uygulama düzeyinde yer alan tüm maddelerin referans grup lehine farklılaştığı görülmektedir. Bu duruma sebep olan etkenlerden biri referans grup olarak alınan Örtük ülke sınıfı 1'de yer alan ülkelerin eğitime önem veren test performansları yüksek ve müfredatı uygulama odaklı Asya ülkeleri ile Avrupa ülkeleri olması ve derslerde uygulamaya ağırlık verilmesi olabilir.

Ortak deęişken olarak analizlere dahil edilen evdeki kitap sayısının örtük ülke sınıflarında yer alan örtük öğrenci sınıflarının oluşumunda etkili olduğu söylenebilir. Evdeki kitap sayısı fazla olan örtük öğrenci sınıflarının daha yüksek test ortalamalarına sahip olduğu görülmektedir. Ayrıca en başarılı örtük öğrenci sınıfı ile örtük ülke sınıfının uygulama bilişsel alanına yönelik DMF’li maddelerde avantajlı durumda oldukları görülmektedir.

İkinci düzey olan ülkelere ait DMF sonuçlarının test bölümleri, konu alanları ve bilişsel alanlar için incelendięi son bölümde DMF’li olarak belirlenen maddelerin üçte ikisinin matematik bölümüne ait olduğu görülmüştür. Ayrıca matematik bölümüne ait maddelerin yarısı da uygulama düzeyi bilişsel alanında farklılaşmıştır. Fen bilimleri bölümüne ait tüm DMF’li maddelerin de bilgi düzeyi bilişsel alanında olduğu önemli bir bulgudur ve bu tür bilgilerin ülkeler arasında farklılaşmanın tespitinin yapılmasını kolaylaştırarak DMF’nin kaynağına ulaşmayı sağlamış olduğu söylenebilir.

Öneriler

Bu bölümde araştırmanın sonuçlarına ve yeni araştırmalar için araştırmacılara yönelik önerilere yer verilmiştir.

Araştırmanın Sonuçlarına Yönelik Öneriler

Bir testte geçerlięi sağlama adına yapılması gereken çalışmalardan biri olan DMF analizi sonucu DMF kaynaklarını bulmak testte adillięi sağlama adına önemlidir. Bu aşamada KMTK modelleri kullanarak örtük sınıfların belirlenmesi önem arz etmektedir. DMF analizi yapılmak istenen çalışmalarda sadece gözlenen grupları kullanarak analiz yapmak yerine örtük sınıfları ortaya çıkarıp analizi yapmak daha az hata yapılmasını ve daha isabetli bulguların elde edilmesini sağlayacaktır. Ayrıca gözlenen gruplara göre ortaya çıkmayabilen ama var olan farklılıklar bu sayede tespit edilebilir. Bu yüzden daha zor ve zaman alıcı olsa da KMTK modelleri çalışmalarda kullanılabilir.

TIMSS tarzı uluslararası bazda karşılaştırma imkânı sunan çalışmalar için kapsamlı analizlerin yapılması detaylı sonuçların elde edilmesini sağlayarak farklılaşmanın

giderilmesinde her düzeydeki paydaş için ayrı ayrı ele alınması gereken çıktılar sunacaktır. Bu yüzden çalışmalara boyut, düzey ve ortak değişken gibi faktörler katılarak analiz zenginleştirilebilir.

Araştırma kapsamında örtük ülke sınıfları arasında ortaya çıkan farklılık ülkelere ait ilgili kurumlarca dikkatle incelenebilir ve farklılığa sebep olan etkenler ortaya çıkarılabilir. Bu çalışma bağlamında ortak değişken olarak alınan EKS değişkeninin yerine başka değişkenler de analizlere dahil edilerek farklılaşmanın eTIMSS'in doğasından kaynaklanma durumu incelenebilir.

Araştırmacılara Yönelik Öneriler

Elde edilen bulgulardaki farklılaşma araştırmacıların veri yapısını dikkatli bir şekilde incelemeleri gerektiğini göstermektedir. Çünkü modelin farklılaşması örtük sınıflardaki bireyleri ve ikinci düzey örtük sınıflardaki ülkeleri farklılaştırmaktadır.

KMTK modellerinde model karmaşıklıkça Mplus programında analiz süresi oldukça uzamaktadır. İki ortak değişkenin eklendiği CB2-C2 2PL modele ait analiz 13. nesil işlemci ve 16 GB RAM bulunduran bir dizüstü bilgisayarda yaklaşık 40 saat sürmüştür. Bu sebeple modeli oluştururken ve veri setini belirlerken bu durum göz özüne alınmalıdır.

Mplus ile yapılan analizlerde modele bağlı olarak madde parametrelerinde olağan sınırların oldukça dışında değerler çıktı alınabilmektedir. Bu sebeple analize başlarken başlangıç değeri ataması yapılarak iterasyon sayısı artırılmalıdır.

Bu çalışmada analiz yükünün fazla olması sebebiyle 2 PL ve 3 PL modeller kullanılmıştır. Araştırma sonucunda ise 2PL ve 3PL modeller için elde edilen değerlerin farklı olduğu görülmüştür. Bu yüzden farklı parametrelerin işe koşularak 1 PL, Rasch özellikle de 4 PL modeller ile model karşılaştırmaları yapıp DMF'li maddeler irdelenebilir.

Ortak değişken seçimi örtük sınıfların oluşumunda etkili olmuştur. Bu sebeple ortak değişkenin ilgilenilen değişkenlerle yüksek korelasyon sağlaması analizlerin doğruluğunu artırabileceği için ortak değişken seçimine dikkat edilmelidir.

Arařtırma kapsamında eTIMSS verileri kullanılmıřtır ve maddeler paylařılmamaktadır. Bu yzden DMF'li olarak belirlenen maddelere ait inceleme yapılamamaktadır. Maddelerine eriřilebilen bir testi uygulamak madde zerinde inceleme imkânı saęlayacaktır.

Arařtırmada drt seęenekli oktan seęmeli maddeler kullanılmıřtır. TIMSS'de yer alan kısa cevaplı, farklı sayıda seęenekli ve aık ulu maddeler ile de KMTK modelleri kullanılarak DMF alıřmaları yapılabilir.

ok boyutluluęu saęlamada bu alıřma zeline matematik ve fen bilimleri maddeleri kullanılmıřtır. Farklı derslerin ya da farklı boyutları len bir yapının KMTK modelleri ile incelenmesi rtk sınıflar arasında gz nnde olmayan iliřkileri ortaya ıkarabilir.

Kaynaklar

- Abedlazeed, Nabeel, Exploring DIF: Comparison of CTT and IRT Methods (September 6, 2010). OIDA International Journal of Sustainable Development, Vol. 1, No. 7, pp. 11-46, 2010, Available at SSRN: <https://ssrn.com/abstract=1672735>
- Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, 29, 67-91.
- Ackerman, T. A. (1994). Using multidimensional item response theory to understand what items and tests are measuring. *Applied Measurement in Education*, 7, 255-278.
- Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice*, 22(3), 37–53. <https://doi.org/10.1111/j.1745-3992.2003.tb00136.x>
- Akbař, D. & Kahraman, N. (2019). Bireysel farklılıkların kategorik deęişkenler olarak modellenmesinde örtük sınıf analizi kullanımı için uygulama kılavuzu: psikolojik dayanıklılık örneęi. *Akdeniz Eęitim Arařtırmaları Dergisi*, 13(29), 356-382. 10.29329/mjer.2019.210.19
- Allalouf, A., Hambleton, R. K., & Sireci, S. G. (1999). Identifying the causes of DIF in translated verbal items. *Journal of Educational Measurement*, 36(3), 185–198.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (1985). *Standards for educational and psychological testing*. American Educational Research Association.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (1999). *Standards for educational and psychological testing*. American Educational Research Association.

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Angoff, W. H. (1993). Perspective on differential item functioning methodology. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 3–24). Lawrence Erlbaum.
- Atar, B. (2010). Basit Doğrusal Regresyon Analizi ile Hiyerarşik Doğrusal Modeller Analizinin Karşılaştırılması, *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, Kış 2010, 1(2), 78-84*.
- Ayan, C. (2011). PISA 2009 fen okuryazarlığı alt testinin değişen madde fonksiyonu açısından incelenmesi. Yayınlanmamış Yüksek Lisans Tezi, Hacettepe Üniversitesi, Ankara.
- Bilir, M. K. (2009). *Mixture item response theory-MIMIC Model: Simultaneous estimation of differential item functioning for manifest groups and latent classes*. Unpublished doctoral dissertation. Florida State University.
- Bolt, D. M., Cohen, A. S., & Wollack, J. A. (2002). Item parameter estimation under conditions of test speededness: Application of a mixture Rasch model with ordinal constraints. *Journal of Educational Measurement, 39*, 331-348.
- Boughton, K., & Yamamoto, K. (2007). A HYBRID model for test speededness. In M. von Davier & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models* (pp. 147-156). New York, NY: Springer.
- Camilli, G. (1992). A conceptual analysis of differential item functioning in terms of a multidimensional item response model. *Applied Psychological Measurement, 16*, 129-147.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. California: Sage.

- Cattell, R. B. (1966). The Scree Plot Test for the Number of Factors. *Multivariate Behavioral Research*, 1, 140-161. http://dx.doi.org/10.1207/s15327906mbr0102_10
- Ceylan, N. (2022). Uzman görüşüne göre yanlı olduğu belirlenen maddelerin değişen madde fonksiyonu belirleme yöntemleri ile tutarlılığının incelenmesi. Yayınlanmamış Yüksek Lisans Tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Chen, H. F., & Jin, K. Y. (2018). Applying logistic regression to detect differential item functioning in multidimensional data. *Frontiers in Psychology*, 9, 1302.
- Chen, Q., Kwok, O.-M., Luo, W., & Wilson, V. L. (2010). The impact of ignoring a level of nesting structure in multilevel growth mixture models: A Monte Carlo study. *Structural Equation Modeling*, 17, 570-589.
- Cheng, C.-P., Chen, C.-C., & Shih, C.-L. (2020). An Exploratory Strategy to Identify and Define Sources of Differential Item Functioning. *Applied Psychological Measurement*, 44(7-8), 548-560. <https://doi.org/10.1177/0146621620931190>
- Cho, E., & Kim, S. (2015). Cronbach's coefficient alpha: Well-known but poorly understood. *Organizational Research Methods*, 18, 207-230.
- Cho, S. J., (2007). A multilevel mixture irt model for dif analysis. Unpublished Doctoral Dissertation, University of Georgia.
- Cho, S.-J., & Cohen, A. S. (2010). A Multilevel Mixture IRT Model With an Application to DIF. *Journal of Educational and Behavioral Statistics*, 35(3), 336-370. <https://doi.org/10.3102/1076998609353111>
- Cho, S. J., Suh, Y., & Lee, W. Y. (2015). An NCME instructional module on latent dif analysis using mixture item response models. *educational measurement: issues and practice*.
- Choi, H. J. (2010). A model that combines diagnostic classification assessment with mixture item response theory models. Unpublished doctoral dissertation, University of Georgia, Athens, Georgia.

- Choi, Y. J., Alexeev, N. & Cohen, A. S. (2015) Differential Item Functioning Analysis Using a Mixture 3-Parameter Logistic Model With a Covariate on the TIMSS 2007 Mathematics Test, *International Journal of Testing*, 15:3, 239-253, DOI: [10.1080/15305058.2015.1007241](https://doi.org/10.1080/15305058.2015.1007241)
- Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. an ncmie instructional module. *Educational Measurement: Issues and Practice*, 17(1), 31-44.
- Cohen, A.S., & Bolt, D.M. (2005). *A mixture model analysis of Differential item functioning. Journal of Educational Measurement*, 42, (2), 133-148.
- Cohen, A. S., Gregg, N., & Deng, M. (2005). The role of extended time and item content on a high-stakes mathematics test. *Learning Disabilities Research and Practice*, 20, 225-233.
- Crawford, A. V., Green, B. S., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of paralel analysis methods for determining the number of factors. *Educational and Psychological Measurements*, 70(6), 885-901.
- Cronbach, L. J., Schonemann, P., & McKie, D. (1965). Alpha coefficients for stratified-paralel tests. *Educational and Psychological Measurement*, 25, 291-312.
- Dai, Y. (2009). A mixture Rasch model with a covariate: a simulation study via Bayesian Markov Chain Monte Carlo Estimation. Unpublished doctoral dissertation, University of Maryland, College Park, Maryland.
- Dai, Y. (2013). A mixture Rasch model with a covariate: A simulation study via Bayesian Markov chain Monte Carlo estimation. *Applied Psychological Measurement*, 37, 375-396.
- De Ayala, R. J. (2009). *The theory and practice of Item Response Theory*. The Guilford Press.

- De Ayala, R. J., Kim, S.-H., Stapleton, L. M., & Dayton, C. M. (2002). *Differential item functioning: A mixture distribution conceptualization. International Journal of Testing, 2*, 243–276.
- De Ayala, R. J., & Santiago, S. Y. (2017). An introduction to mixture item response theory models. *Journal of School Psychology, 60*, 25-40.
- De Boeck, P., Cho, S.-J., & Wilson, M. (2011). Explanatory secondary dimension modeling of latent differential item functioning. *Applied Psychological Measurement, 35*, 583–603.
- De Jong, M.G., Steenkamp, J.B.E.M. (2010). Finite Mixture Multilevel Multidimensional Ordinal IRT Models for Large Scale Cross-Cultural Research. *Psychometrika* **75**, 3–32 <https://doi.org/10.1007/s11336-009-9134-z>
- DeMars, C. (2016). *Madde Tepki Kuramı* (H. Kelecioğlu, Cev.). Nobel Yayıncılık. (2010).
- DeMars, C. E., & Lau, A. (2011). Differential Item Functioning Detection With Latent Classes: How Accurately Can We Detect Who Is Responding Differentially? *Educational and Psychological Measurement, 71*(4), 597-616. <https://doi.org/10.1177/0013164411404221>
- Demir, S. (2013). *PISA 2009 matematik okuryazarlığı alt testinde bulunan maddelerinin Mantel-Haenszel, SIBTEST ve Lojistik Regresyon yöntemleri ile değişen madde fonksiyonunun incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi, Abant İzzet Baysal Üniversitesi, Bolu.
- Dinno A (2018). *_paran: Horn's Test of Principal Components/Factors_*. R package version 1.5.2, <<https://CRAN.R-project.org/package=paran>>.
- Doğan, N., & Öğretmen, T. (2008). Değişen madde fonksiyonunu belirlemede Mantel-Haenszel, ki-kare ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim, 33*(148), 100-112.

- Dras, L. (2023). *Multilevel mixture irt modeling for the analysis of differential item functioning*. Unpublished Doctoral dissertation, Brigham Young University.
- Embretson, S. E. (2007). Mixed Rasch models for measurement in cognitive psychology. In von Davier, M. & Carstensen C. H. (Eds.). *Multivariate and Mixture Distribution Rasch Models: Extensions and Applications*. New York, NY: Springer.
- Ercikan, K. (2002). Disentangling sources of differential item functioning in multilanguage assessments. *International Journal of Testing*, 2, 199–215.
- Fidalgo, A. M. & Madeira, J. M. (2008). Generalized mantel-haenszel methods for differential item functioning detection. *Educational and Psychological Measurement*, 68(6), 940-958.
- Finch, W.H. & French, B.F. (2018). *Educational and Psychological Measurement* (1st ed.). Routledge. <https://doi.org/10.4324/9781315650951>
- Finch, W. H., & Hernández Finch, M. E. (2013). Investigation of specific learning disability and testing accommodations based differential item functioning using a multilevel multidimensional mixture item response theory model. *Educational and Psychological Measurement*, 73(6), 973-993.
- Fishbein, B., Foy, P., & Yin, L. (2021). *TIMSS 2019 User Guide for the International Database* (2nd ed.). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/timss2019/international-database/>
- French, B. F., & Finch, W. H. (2010). Hierarchical logistic regression: Accounting for multilevel data in DIF detection. *Journal of Educational Measurement*, 47, 299-317.
- French, B. F., & Finch, W. H. (2013). Extensions of Mantel–Haenszel for Multilevel DIF Detection. *Educational and Psychological Measurement*, 73(4), 648-671. <https://doi.org/10.1177/0013164412472341>

- Fukuhara, H. & Kamata, A. (2011). A Bifactor Multidimensional Item Response Theory Model for Differential Item Functioning Analysis on Testlet-Based Items. *Applied Psychological Measurement* 35(8) 604–622.
- Gelin, M. N. (2005). Type I error rates of the DIF MIMIC approach using Joreskog's covariance matrix with ML and estimation (Doctoral dissertation). Retrieved from <https://open.library.ubc.ca/cIRcle/collections/ubctheses/831/items/1.0054355>
- Gnaldi, M., Bacci, S. & Bartolucci, F. A multilevel finite mixture item response model to cluster examinees and schools. *Adv Data Anal Classif* 10, 53–70 (2016). <https://doi.org/10.1007/s11634-014-0196-0>
- Gómez-Benito, J. ve Navas-Ara, M. J. (2000). A Comparison of χ^2 , RFA and IRT Based Procedures in the Detection of DIF. *Quality and Quantity*, 34(1), 17-31.
- Gök, B., Kelecioğlu, H. ve Doğan, N. (2010). Değişen madde fonksiyonunu belirlemede Mantel-Haenzsel ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 35(156), 3-16.
- Graham, J. M. (2006). Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability. *Educational and Psychological Measurement*, 66(6), 930–944.
- Hagenaars, J. A., & McCutcheon, A. L. (Eds.). (2002). *Applied latent class analysis*. Cambridge University Press.
- Gürdil, H. (2023). *Değişen madde fonksiyonunu kontrol altına almada çok boyutlu madde tepki kuramı modellerinin kullanımı*. Yayınlanmamış doktora tezi, Ankara Üniversitesi, Ankara.
- Gürkan, G. (2021). *From ols to multilevel multidimensional mixture irt: a model refinement approach to investigating patterns of relationships in pisa 2012 data*. Yayınlanmamış doktora tezi, Boston, Amerika Birleşik Devletleri.
- Hambleton, R. K. (2006). Good practices for identifying differential item functioning. *Medical Care*, 44(11), 182-188.

- Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory: Principles and applications*. Kluwer, Nijhoff Publishing.
- Hambleton, R. K., Swaminathan, R., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Sage Publication.
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum.
- Hooper, D., Coughlan, J., Mullen, M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.
- Hox, J. J. (2013). Multilevel Regression and Multilevel Structural Equation Modeling. In T. D. Little (Ed.), *The Oxford Handbook of Quantitative Methods* (pp. 281-294). New York: Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199934898.013.0014>
- Jensuttiwetchakul, P., Kanjanawasee, S. & Ngudgratoke, S. (2016). A Development of The 3PL MMM-IRT Model for Identifying Latent Class. *Procedia - Social and Behavioral Sciences* 217 (2016) 719 – 728
- Jeon, M. (2019). A specialized confirmatory mixture IRT modeling approach for multidimensional tests. *Psychological Test and Assessment Modeling*, 61(1), 91-123.
- Jiao, H., Kamata, A., Wang, S., & Jin, Y. (2012). A multilevel testlet model for dual local dependence. *Journal of Educational Measurement*, 49, 82-100.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International; Lawrence Erlbaum Associates, Inc.

- Kaiser, H. F. (1960). The applications of electronic computer to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Kalayciođlu, D. B., & Berberođlu, G. (2011). Differential Item Functioning Analysis of the Science and Mathematics Items in the University Entrance Examinations in Turkey. *Journal of Psychoeducational Assessment*, 29(5), 467-478.
<https://doi.org/10.1177/0734282910391623>
- Kane, M. T. (2006). Validation. In R. B. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education/Praeger.
- Kanovský, M., Halamová, J., Zuroff, D. C., Troop, N. A., Gilbert, P., Shahar, B., ... & Kupeli, N. (2020). A Multilevel Multidimensional Finite Mixture Item Response Model to Cluster Respondents and Countries. *European Journal of Psychological Assessment*.
- Kaplan, D., & Keller, B. (2011). A note on cluster effects in latent class analysis. *Structural Equation Modeling*, 18, 525-536.
- Karadavut, T., Cohen, A. S., & Kim, S. H. (2019). Mixture Rasch model with main and interaction effects of covariates on latent class membership. *International Journal of Assessment Tools in Education*, 6(3), 362-377.
- Kelderman, H., & Macready, G. B. (1990). The use of loglinear models for assessing differential item functioning across manifest and latent examinee groups. *Journal of Educational Measurement*, 27(4), 307-327.
- Koçak, D., Çokluk, Ö. ve Kayri, M. (2016). Faktör sayısının belirlenmesinde MAP testi, paralel analiz, K1 ve yamaç birikinti grafiđi yöntemlerinin karşılaştırılması. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 13(1), 330-359.
- Lee, W. Y., Cho, S. J., & Sterba, S. K. (2018). Ignoring a Multilevel Structure in Mixture Item Response Models: Impact on Parameter Recovery and Model Selection. *Applied psychological measurement*, 42(2), 136-154.
<https://doi.org/10.1177/0146621617711999>

- Li, F., Cohen, A., Bottge, B., & Templin, J. (2015). A Latent Transition Analysis Model for Assessing Change in Cognitive Skills. *Educational and Psychological Measurement*, 76(2), 181-204.
- Li, F., Cohen, A. S., Kim, S. H., & Cho, S. J. (2009). Model selection methods for mixture dichotomous IRT models. *Applied Psychological Measurement*.
- Li, T. (2014). *Different approaches to covariate inclusion in the mixture Rasch model*. Unpublished Doctoral dissertation, University of Maryland, College Park.
- Li, T., Jiao, H., & Macready, G. B. (2016). Different Approaches to Covariate Inclusion in the Mixture Rasch Model. *Educational and Psychological Measurement*, 76(5), 848-872. <https://doi.org/10.1177/0013164415610380>
- Lord, F. M. (1980). *Applications of Item response theory to practical testing problems*. Erlbaum.
- Magis, D., Beland, S., Tuerlinckx, F., De Boeck, P. (2010). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods*, 42, 847-862.
- Maij-de Meij, A. M., Kelderman, H., & van der Flier, H. (2008). Fitting a mixture item response theory model to personality questionnaire data: Characterizing latent classes and investigating possibilities for improving prediction. *Applied Psychological Measurement*, 32, 611–631.
- Maij-de Meij, A. M., Kelderman, H., & van der Flier, H. (2010). Improvement in detection of differential item functioning using a mixture item response theory model. *Multivariate Behavioral Research*, 45(6), 975-999.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. L. Erlbaum Associates, Mahwah, N.J.
- McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York, NY: John Wiley.

- Meyer, J. P. (2010). A mixture Rasch model with item response time components. *Applied Psychological Measurement, 34*, 521-538.
- Mislevy, R. J., & Verhelst, N. D. (1990). Modeling item responses when different subjects employ different solution strategies. *Psychometrika, 55*, 195-215.
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. Boston College, TIMSS & PIRLS International Study Center. <https://timssandpirls.bc.edu/timss2019/international-results/>
- Muthén, L.K. ve Muthén, B.O. (1998-2017). *Mplus User's Guide*. Eighth Edition. Los Angeles, CA: Muthén & Muthén.
- Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5*, 343-355.
- Park, J., & Yu, H.-T. (2016). The impact of ignoring the level of nesting structure in nonparametric multilevel latent class models. *Educational and Psychological Measurement, 76*, 824-847.
- Pektaş, S. (2018). Değişen madde fonksiyonu belirleme yöntemlerinin test parametreleri kestirimlerine, karar çalışmalarına, g ve phi katsayılarına etkisi. Yayınlanmamış Doktora Tezi. Gazi Üniversitesi, Ankara.
- Penfield, R. D. (2001). Assessing differential item functioning across multiple groups: A comparison of three Mantel–Haenszel procedures. *Applied Measurement in Education, 14*, 235-259.
- Penfield, R. D., & Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26, pp. 125–167). Amsterdam, The Netherlands: North-Holland.

- Piccone, A.V. (2009). A comparison of three computational procedures for solving the number of factors problem in exploratory factor analysis. Published Doctor of Philosophy dissertation. University of Northern Colorado.
- Potenza, M. T., & Dorans, N. J. (1995). DIF assessment for polytomously scored items: A framework for classification and evaluation. *Applied Psychological Measurement*, 19(1), 23–37. <https://doi.org/10.1177/014662169501900104>
- Preinerstorfer, D., & Formann, A. K. (2011). Parameter recovery and model selection in mixed Rasch models. *British Journal of Mathematical and Statistical Psychology*, 65(2), 251-262. doi: 10.1111/j.2044-8317.2011.02020.x
- Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical Linear Models. Applications and Data Analysis Methods (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Reckase, M. D. (1985). The difficulty of items that measure more than one ability. *Applied Psychological Measurement*, 9, 401-412.
- Reckase, M. D. (2009). *Multidimensional item response theory*. New York, NY: Springer.
- Revelle, W. (2023). *Psych: Procedures for psychological, psychometric, and personality research. (Version 2.3.3)*. <https://cran.r-project.org/web/packages/psych/psych.pdf>
- Revelle, W. & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega and the glb: comments on Sijsma. *Psychometrika*, 74(1):145–154.
- Rijmen, F., De Boeck, P. (2005). A relation between a between-item multidimensional IRT model and the mixture rasch model. *Psychometrika* **70**, 481–496. <https://doi.org/10.1007/s11336-002-1007-7>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement*, 14, 271–282.

- Saatçiođlu, F. M. (2021). Karma mtk modellerinin farklı kořullarda parametre kestirimine ve sınıflama dođruluđuna etkisinin simülatif ve gerçek veri ile incelenmesi. Yayınlanmamış Doktora Tezi. Gazi Üniversitesi, Ankara.
- Samuelson, K. (2005). Examining differential item functioning from a latent class perspective (Unpublished doctoral dissertation). University of Maryland, College Park.
- Samuelsen, K. (2008). Examining differential item function from a latent class perspective. In G. Hancock & K. Samuelsen (Eds.) *Mixture Models in Latent Variable Research*. Greenwich, CT: Information Age Publishing, Inc.
- Scheuneman, J. (1979). A method of assessing bias in test items. *Journal of Educational Measurement*, 16, 143–152.
- Sen, S. (2014). Robustness of mixture IRT models to violations of latent normality. Unpublished doctoral dissertation. University of Georgia, Athens, Georgia.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74, 107-120.
- Sireci, S. G., Fitzgerald, C., & Xing, D. (1998). Adapting credentialing examinations for international uses [Paper presentation]. Annual meeting of the American Educational Research Association, San Diego, CA.
- Smit, A., Kelderman, H., & van der Flier, H. (1999). Collateral information and mixed Rasch models. *Methods of Psychological Research Online*, 4, 1-13.
- Şen, S. (2022). Mplus ile yapısal eşitlik modellemesi uygulamaları. Ankara: Nobel.
- Şen, S., Cohen, A. S., & Kim, S. H. (2018). Model selection for multilevel mixture Rasch models. *Applied Psychological Measurement*, 43(4), 1-18. doi: 10.1177/0146621618779990

- Şen, S., Cohen, A. S., & Kim, S. (2020). *A Short Note on Obtaining Item Parameter Estimates of IRT Models with Bayesian Estimation in Mplus*, 11(3), 266-282. doi: 10.21031/epod.693719.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
- Taşdelen Teker, G. (2014). *Madde takımlarının güvenilirlik ve değişen madde fonksiyonu üzerine etkisi*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi, Ankara.
- Tay, L., Diener, E., Drasgow, F., & Vermunt, J. K. (2011). Multilevel Mixed-Measurement IRT Analysis: An Explication and Application to Self-Reported Emotions Across the World. *Organizational Research Methods*, 14(1), 177-207. <https://doi.org/10.1177/10944281110372674>
- Tiryaki, F. (2019). *PISA 2015 öğrenci tutum anketlerinin değişen madde fonksiyonu ve ölçme değişmezliğinin incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi. Ankara Üniversitesi, Ankara.
- Tofighi, D., & Enders, C. (2008). *Identifying the correct number of classes in a growth mixture model*. In G. Hancock & K. Samuelsen (Ed.), *Advances in latent variable mixture models* (pp. 317–341). Greenwich, CT: Information Age Publishing, Inc.
- Tucker, L.R., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika*, 34(4), 421-459.
- Uyar, Ş. (2015). *Ortak maddelerin değişen madde fonksiyonu gösterip göstermemesi durumunda test eşitlemeye etkisinin farklı yöntemlerle incelenmesi*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi, Ankara.
- Uysal Saraç, M. (2022). *Karma rasch modele ortak değişken ekleme yaklaşımlarının sınıflama ve kestirim doğruluğu açısından karşılaştırılması*. Yayınlanmamış Doktora Tezi. Gazi Üniversitesi, Ankara.

- Vermunt, J. K., & Magidson, J. (2005). *Technical guide for Latent GOLD 4.0: Basic and advanced*. Belmont MA: Statistical Innovations Inc.
- von Davier, M., Rost, J., & Carstensen, C. H. (2007). Introduction: Extending the Rasch model. In M. von Davier & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models: Extensions and applications* (pp. 1–12). New York, NY: Springer.
- von Davier, & Yamamoto, K. (2007). Mixture- distribution and hybrid Rasch models. In M. von Davier, & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models: Extensions and applications* (pp. 99-115). New York: Springer Verlag.
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, 17, 228-243.
- Yalcin, S. (2018). Determining differential item functioning with the mixture item response theory. *Eurasian Journal of Educational Research* 74,187-206.
- Yanagida T (2023). `_misty`: Miscellaneous Functions 'T. Yanagida'. R package version 0.5.2, <<https://CRAN.R-project.org/package=misty>>.
- Yao, L. & Li, F. (2015). A DIF detection procedure in multidimensional item response theory framework using MCMC technique. *Int. J. Quantitative Research in Education*, Vol. 2, Nos. 3/4 285-304.
- Yıldız, H. (2021). `irtGUI`: Item Response Theory Analysis with a Graphic User Interface_. R package version 0.2, <<https://CRAN.R-project.org/package=irtGUI>>.
- Zhang, Y. (2017). *Detection of latent differential item functioning (dif) using mixture 2pl irt model with covariate*. Unpublished doctoral dissertation. University of Pittsburgh. Pittsburgh.

Zhu, X. (2013). *Distinguishing continuous and discrete approaches to Multilevel Mixture IRT Models: A Model Comparison Perspective*. Doctoral Dissertation. University of Maryland

Zinbarg, R. E., Revelle, W., Yovel, I., and Li, W. (2005). Cronbach's α , Revelle's β , and McDonald's ω^2 : Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70(1), 123–133.

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.

EK A: KMTK Modeline Ait Mplus Kodlarından Biri

TITLE: TEZ_08.02_L0-G3 ;

VARIABLE: NAMES ARE country m1-m20 sum gender book meanbook meangender;

USEVARIABLES= m1-m20;

CATEGORICAL= m1-m20;

CLASSES=c(3);

DATA: FILE= C:\Users\OMERDOGAN\Desktop\yeni_covbook_2.txt;

ANALYSIS: TYPE = MIXTURE;

ALGORITHM = INTEGRATION;

PROCESSORS = 2;

MODEL:

%OVERALL%

f BY m1-m20;

[f@0];f@1;

%c#1%

f BY m1-m20;

[m1\$1-m20\$1]

[f@0];f@1;

%c#2%

f BY m1-m20;

[m1\$1-m20\$1]

[f@0];f@1;

%c#3%

f BY m1-m20;

[m1\$1-m20\$1]

[f@0];f@1;

OUTPUT: TECH1 TECH4 TECH8;

PLOT: TYPE=PLOT2;

SAVEDATA:file is TEZ_08.02_L0-G3.txt ;

save is cprob;

format is free;

EK B: ÇDKMTK Modeline Ait Mplus Kodlarından Biri

TITLE: TEZ tek boyut 27.04 L2-G2;

VARIABLE: NAMES ARE country m1-m20;

USEVARIABLES= m1-m20;

CATEGORICAL= m1-m20;

CLASSES=cb(2) c(2);

BETWEEN=cb;

CLUSTER =country;

DATA: FILE= C:\Users\OMERDOGAN\Desktop\TEZ20madde.txt;

ANALYSIS: TYPE = TWOLEVEL MIXTURE;

ALGORITHM = INTEGRATION;

PROCESSORS = 2;

MITERATIONS=1000;

MODEL:

%WITHIN%

%OVERALL%

f BY m1-m20;

%cb#1.c#1%

f BY m1-m20*;

[f@0];f@1;

%cb#1.c#2%

f BY m1-m20*;

[f@0];f@1;

%BETWEEN%

%OVERALL%

fb BY m1-m20;

fb@1;

%cb#1.c#1%

fb BY m1-m20*;

[m1\$1-m20\$1];

[fb@0];

%cb#1.c#2%

fb BY m1-m20*;

[m1\$1-m20\$1];

[fb@0];

%cb#2.c#1%

fb BY m1-m20*;

[m1\$1-m20\$1];

[fb@0];

%cb#2.c#2%

fb BY m1-m20*;

[m1\$1-m20\$1];

[fb@0];

OUTPUT: TECH1 TECH4 TECH8;

PLOT: TYPE=PLOT2;

SAVEDATA:file is TEZ_27.04_L2-G2_TEK.txt;

save is cprob;

format is free;

EK C: ÇDÇBKMTK Modeline Ait Mplus Kodlarından Biri

TITLE: TEZ 15.04_L2-G3_3PL_dim2;

VARIABLE: NAMES ARE country m1-m20;

USEVARIABLES= m1-m20;

CATEGORICAL= m1-m20(3pl);

CLASSES=cb(2) c(3);

BETWEEN=cb;

CLUSTER =country;

DATA: FILE= C:\Users\OMERDOGAN\Desktop\TEZ20madde.txt;

ANALYSIS: TYPE = TWOLEVEL MIXTURE;

ALGORITHM = INTEGRATION;

Integration = MONTECARLO(500);

PROCESSORS = 2;

MITERATIONS=1000;

MODEL:

%WITHIN%

%OVERALL%

f1 BY m1-m10;

f2 BY m11-m20;

[f1@0];f1@1;

[f2@0];f2@1;

%cb#1.c#1%

f1 BY m1-m10;

[f1@0];f1@1;

f2 BY m11-m20;

[f2@0];f2@1;

%cb#1.c#2%

f1 BY m1-m10;

[f1@0];f1@1;

f2 BY m11-m20;

[f2@0];f2@1;

%cb#1.c#3%

f1 BY m1-m10;

[f1@0];f1@1;

f2 BY m11-m20;

[f2@0];f2@1;

%BETWEEN%

%OVERALL%

fb1 BY m1-m10;

fb2 BY m11-m20;

fb1@1;

fb2@1;

%cb#1.c#1%

fb1 BY m1-m10;

[m1\$2-m10\$2](a1-a10);

[fb1@0];fb1@1;

fb2 BY m11-m20;

[m11\$2-m20\$2](a11-a20);

[fb2@0];fb2@1;

%cb#1.c#2%

fb1 BY m1-m10;

[fb1@0];fb1@1;

[m1\$2-m10\$2](a1-a10);

fb2 BY m11-m20;

[fb2@0];fb2@1;

[m11\$2-m20\$2](a11-a20);

%cb#1.c#3%

fb1 BY m1-m10;

[fb1@0];fb1@1;

[m1\$2-m10\$2](a1-a10);

fb2 BY m11-m20;

[fb2@0];fb2@1;

[m11\$2-m20\$2](a11-a20);

MODEL PRIORS:

a1-a10~N(1.386,1);

a11-a20~N(1.386,1);

OUTPUT: TECH1 TECH4 TECH8;

PLOT: TYPE=PLOT2;

SAVEDATA:file is TEZ_15.04_L2-G3_3PL_dim2.txt;

save is cprob;

format is free;

EK D: odÇDÇBKMTK Modeline Ait Mplus Kodlarından Biri

TITLE: TEZ_03.06_2OD_L2-G2-3PL_dim2;

VARIABLE: NAMES ARE country m1-m20 sum gender book meanbook ;

USEVARIABLES= m1-m20 book meanbook;

CATEGORICAL= m1-m20(3pl);

CLASSES=cb(2) c(2);

WITHIN=book;

BETWEEN=cb meanbook;

CLUSTER =country;

DATA: FILE= C:\Users\OMERDOGAN\Desktop\yeni_covbook_2.txt;

ANALYSIS: TYPE = TWOLEVEL MIXTURE;

STARTS=250 50;

ALGORITHM = INTEGRATION;

INTEGRATION=MONTECARLO(500);

PROCESSORS = 2;

MITERATIONS=1000;

MODEL:

%WITHIN%

%OVERALL%

f1 BY m1-m10;

f2 BY m11-m20;

c ON book;

%cb#1.c#1%

f1 BY m1-m10;

f2 BY m11-m20;

[f1@0];

[f2@0];

%cb#1.c#2%

f1 BY m1-m10;

f2 BY m11-m20;

[f1@0];

[f2@0];

%BETWEEN%

%OVERALL%

fb1 BY m1-m10;

fb2 BY m11-m20;

fb1@1;fb2@1;

cb ON meanbook;

%cb#1.c#1%

fb1 BY m1-m10;

[m1\$2-m10\$2](a1-a10);

[fb1@0];

fb2 BY m11-m20;

[m11\$2-m20\$2](a11-a20);

[fb2@0];

%cb#1.c#2%

fb1 BY m1-m10;

[m1\$2-m10\$2](a1-a10);

[fb1@0];

fb2 BY m11-m20;

[m11\$2-m20\$2](a11-a20);

[fb2@0];

%cb#2.c#1%

fb1 BY m1-m10;

[m1\$2-m10\$2](a1-a10);

[fb1@0];

fb2 BY m11-m20;

[m11\$2-m20\$2](a11-a20);

[fb2@0];

%cb#2.c#2%

fb1 BY m1-m10;

[m1\$2-m10\$2](a1-a10);

[fb1@0];

fb2 BY m11-m20;

[m11\$2-m20\$2](a11-a20);

[fb2@0];

OUTPUT: TECH1 TECH4 TECH8;


PLOT: TYPE=PLOT2;

SAVEDATA:file is TEZ_03.06_2OD_L2-G2-3PL_dim2.txt;

save is cprob;

format is free;

EK E: Araştırma Etik Komisyon İzin Muafiyeti Formu

	Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması/Araştırma Etik Komisyon İzin Muafiyeti Formu	F46
06 / 06 / 2024		
Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Ana Bilim Dalı Başkanlığına		
Araştırma Başlığı	ORTAK DEĞİŞKENLİ ÇOK BOYUTLU ÇOK DÜZEYLİ KARMA MODEL İLE DMF İNCELEMESİ	
Yukarıda başlığı/konusu verilen araştırma 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 veya ruh sağlığına müdahale içermemektedir. 4. Anket, ölçek (test), mülakat, odak grup çalışması, gözlem, deney, görüşme gibi teknikler kullanılarak katılımcılardan veri toplanmasını gerektiren nitel ya da nicel yaklaşımlarla yürütülen araştırmalar niteliğinde değildir. 5. Diğer kişi ve kurumlardan temin edilen veri kullanımını (kitap, belge vs.) gerektirmektedir. Ancak bu kullanım, diğer kişi ve kurumların izin verdiği ölçüde Kişisel Bilgilerin Korunması Kanuna riayet edilerek gerçekleştirilecektir. <p>Çalışmada kullanacağım veriler: (x) Kamusal erişime açık : https://www.iea.nl/studies/iea/timss adresinden alınmıştır. () Özel izin ve onaya tabi (buraya yazınız): () Üretilmiş veri (buraya yazınız): () Diğer (buraya yazınız):</p> <p>Yükseköğretim Kurumları Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre çalışmamın yürütülebilmesi için herhangi bir Etik Komisyondan/Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.</p> <p>Gereğini saygılarımla arz ederim.</p> <p style="text-align: right; margin-right: 100px;">Ömer DOĞAN</p>		
Araştırmacı Bilgileri		
Adı Soyadı	Ömer DOĞAN	
Öğrenci İse No	N20140443	
Ana Bilim Dalı	Eğitim Bilimleri	
Programı	Eğitimde Ölçme ve Değerlendirme	
Statüsü	<input type="checkbox"/> Yüksek Lisans <input checked="" type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr. <input type="checkbox"/> Diğer	
Danışman Görüşü ve Onayı*		
Prof. Dr. Burcu ATAR		
*Tez ve tezden üretilen yayınlarda gerekli		
Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Beytepe Yerleşkesi, 06800, Çankaya / ANKARA Telefon: 0(312) 297 85 72 Belgegeçer: 0(312) 297 85 66 e-Ağ: http://ebe.hacettepe.edu.tr/ e-Posta: ebe@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 yararlanması 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.

...../...../2024

(İmza)

Ömer DOĞAN

EK-G: Doktora Tez Çalışması Orijinallik Raporu

09/07/2024

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

Tez Başlığı : Ortak Değişkenli Çok Boyutlu Çok Düzeyli Karma Model ile DMF İncelemesi

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
09/07/2024	112	157299	27/06/2024	%8	2397456213

Uygulanan filtreler:

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

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

Ad Soyadı: Ömer DOĞAN

Öğrenci No.: N20140443

Ana Bilim Dalı: Eğitim Bilimleri

İmza

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü: Y.Lisans Doktora Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Prof. Dr. Burcu ATAR

EK-H: Dissertation Originality Report

09/07/2024

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

Thesis Title:... DIF Analysis with A Multidimensional Multilevel Mixture Model with Covariates

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
09/07/2024	112	157,299	27/06/2024	8%	2397456213

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: Ömer DOĞAN

Student No.: N20140443

Department: Educational Sciences

Program: Educational Measurement and Evaluation

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

Signature

ADVISOR APPROVAL

APPROVED
Prof. Dr. Burcu ATAR

EK-I: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

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

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

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

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

..... / /

(imza)

Ömer DOĞAN

"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 tezinerişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6.2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3 şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir*. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü 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.

