

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ FAKÜLTESİ

ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
ÖLÇME VE DEĞERLENDİRME PROGRAMI

PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ KURAMI
MODELLERİNDEN ÇEŞİTLİ FAKTÖRLERE GÖRE ELDE EDİLEN MADDE
VE YETENEK KESTİRİMLERİNİN KARŞILAŞTIRILMASI

DOKTORA TEZİ

EZGİ MOR DİRLİK

Ankara, Haziran, 2017

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ FAKÜLTESİ

ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
ÖLÇME VE DEĞERLENDİRME PROGRAMI

PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ KURAMI
MODELLERİNDEN ÇEŞİTLİ FAKTÖRLERE GÖRE ELDE EDİLEN MADDE
VE YETENEK KESTİRİMLERİNİN KARŞILAŞTIRILMASI

DOKTORA TEZİ

EZGİ MOR DİRLİK

Danışman: Prof. Dr. Nizamettin Koç

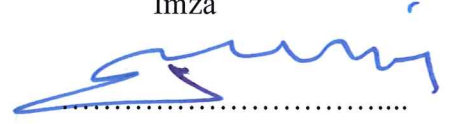
Ankara, Haziran, 2017

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

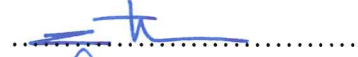
Ezgi Mor Dirlik'in hazırladıđı "Parametrik ve Parametrik Olmayan Madde Tepki Kuramı Modellerinden eřitli Fakt¼rlere G¼re Elde Edilen Madde ve Yetenek Kestirimlerinin Karşılaştırılması" bařlıklı bu alıřma j¼rimiz tarafından Eđitimde ¼lme ve Deđerlendirme Anabilim Dalı/Programı'nda Doktora Tezi olarak kabul edilmiřtir.

İmza


Bařkan : Prof. Dr. Nizamettin KO (Danıřman)



¼ye : Prof. Dr. Ezel TAVŐANCIL



¼ye : Prof. Dr. H¼lya KELECİOđLU



¼ye : Prof. Dr. R. N¼khet DEMİRTAŐLI



¼ye : Do. Dr. İsmail KARAKAYA



ONAY

Bu tez Ankara ¼niversitesi Lisans¼st¼ Eđitim-¼đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından .../.../20... tarihinde uygun g¼r¼lm¼ř ve enstit¼ Y¼nteim Kurulunca .../.../20.. tarihinde kabul edilmiřtir.

Prof. Dr. İsmail G¼ven

Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

ETİK BİLDİRİM

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

Ezgi MOR DİRLİK

ÖZET

PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ KURAMI MODELLERİNDEN ÇEŞİTLİ FAKTÖRLERE GÖRE ELDE EDİLEN MADDE VE YETENEK KESTİRİMLERİNİN KARŞILAŞTIRILMASI

MOR DİRLİK, Ezgi

Doktora, Ölçme ve Değerlendirme Anabilim Dalı

Tez Danışmanı: Prof. Dr. Nizamettin KOÇ

Haziran 2017, xvii + 209

Bu araştırmada, Madde Tepki Kuramı kapsamında yer alan iki farklı yaklaşım olan parametrik ve parametrik olmayan madde tepki kuramı modellerinden kestirilen madde ve yetenek parametrelerinin test uzunluğu, örneklem büyüklüğü ve maddelerin psikometrik nitelikleri faktörlerine göre karşılaştırılması amaçlanmıştır.

Temel araştırma modelinde yer alan araştırmanın verileri Uluslararası Matematik ve Fen Eğilimleri Araştırması'nın (Trends of International Mathematic and Science-TIMSS) 2011 yılındaki uygulamasından temin edilmiştir. Söz konusu uygulamanın matematik alt testlerinden tek boyutluluğu en yüksek düzeyde sağlayan bir kitapçık belirlenmiş ve bu kitapçığı yanıtlayan ve TIMSS 2011 başarı sırasında ilk 20'de yer alan ülkelerin verileri kullanılarak 7242 kişilik bir veri seti oluşturulmuştur. Oluşturulan temel veri setinden çalışmanın amaçları ve incelemeye alınan faktörlerin koşulları gözetilerek 16 ayrı veri seti oluşturulmuştur. Örneklem büyüklüğü faktörü incelenirken 500, 1000 ve 3000 kişilik veri setleri oluşturulmuş ve bu veri setleri beş, 15 ve 25 maddelik test uzunluğu koşulları ile çaprazlanarak PMTK ve POMTK'ya göre madde ve yetenek parametreleri kestirilmiştir. Değişen madde psikometrik nitelikleri faktörü altında ise, madde güçlük ve ayırt edicilik parametresine göre düşük ve yüksek değerlere sahip olan maddeler belirlenerek dört ayrı veri seti oluşturulmuş ve bu veri setlerinden birey yetenek kestirimleri yapılmıştır.

Araştırma soruları kapsamında elde edilen sonuçlar incelendiğinde, her iki yaklaşıma göre farklı örneklem büyüklüğünden kestirilen madde parametreleri arasında yüksek ve manidar bir uyum olduğu belirlenmiştir. Test uzunluğu ve örneklem büyüklüğü koşullarının birlikte incelendiği veri setleri için ise PMTK ve POMTK modellerinden kestirilen madde parametreleri arasında yüksek ve manidar ilişkiler olduğu bulunmuştur.

Aynı yaklaşıma göre kestirilen yetenek parametreleri arasındaki ilişkiler incelendiğinde ise, PMTK ile beş ve 15 maddelik veri setlerinden kestirilen yetenekler arasında manidar bir ilişki bulunamamışken, 25 maddelik veri setleri için tüm örneklemelerden kestirilen yetenek parametreleri arasında yüksek ve manidar ilişkiler olduğu belirlenmiştir. POMTK ile kestirilen yetenekler arasında ise tüm örneklem büyüklüğü ve test uzunluğu koşullarında manidar ve yüksek ilişkiler olduğu bulunmuştur. Ayrıca PMTK ve POMTK'dan kestirilen yetenekler arasındaki ilişkiler incelendiğinde ise, PMTK'dan kestirilen yeteneklerin yalnızca madde sayısının 25 ya da örneklem büyüklüğünün 3000 olduğu durumlarda POMTK ile tutarlı sonuçlar verdiği bulunmuştur. Çalışma kapsamında yetenek parametrelerine etkisi incelenen bir diğer faktör maddelerin güçlük ve ayırt edicilik düzeyleridir ve bu faktöre göre madde güçlük ve ayırt edicilik parametrelerinin düşük ve yüksek olarak gruplandırıldığı dört veri setinden, PMTK ve POMTK modellerinden kestirilen yetenekler arasında yüksek ve manidar ilişkiler bulunmuştur. Özetle her iki yaklaşıma göre farklı koşullar altında kestirilen madde ve yetenek parametreleri birbirleri ile yüksek ve manidar ilişkiler göstermiştir. Dolayısıyla PMTK modellerinin varsayımları veri setleri için yeterli düzeyde karşılanmadığında, incelenen koşullara uygun durumlarda POMTK yaklaşımının tercih edilebileceği sonucuna ulaşılmıştır.

Anahtar Kelimeler: Parametrik Madde Tepki Kuramı, Parametrik Olmayan Madde Tepki Kuramı, Mokken Ölçekleme, örneklem büyüklüğü, test uzunluğu.

SUMMARY

THE COMPARISON OF ITEM AND ABILITY ESTIMATIONS CALCULATED FROM THE PARAMETRIC AND NON-PARAMETRIC ITEM RESPONSE THEORY ACCORDING TO THE SEVERAL FACTORS

MOR DİRLİK, Ezgi

Ph.D., Department of Measurement and Evaluation

Supervisor: Prof. Dr. Nizamettin KOÇ

June, 2017, xvii + 209 pages

This study aimed to compare the item and ability parameters estimated from two different approaches of Item Response Theory, which are parametric and non-parametric, according to the factors of test length, sample size and the item psychometric qualities.

The study was prepared in a basic research method and the data of the study is obtained from the Trends of International Mathematic and Science (TIMSS) 2011 application. The mathematic booklet at eight grade level which had been determined as unidimensional at most was used and the data from the first 20 countries were gathered. The final data set of the research was composed of 7254 students from different countries. From the main data set, 16 different data sets were formed according to the research questions and the conditions of the factors. As for sample size factor, three different data sets were prepared consisted of 500, 1000 and 3000 students and sample size conditions were crossed with the test length conditions which are five, 15 ve 25 items. From these data sets, item and ability parameters were estimated by using the compatible parametric and non-parametric item response models and the results were compared by using the relevant correlation coefficients. According to the last factor of item psychometric quality factors, four different data sets were composed and ability parameters were estimated from them and compared. Lastly, the reliability of the data sets were investigated by using test information functions in parametric item response theory models and calculating Cronbach Alpha, Lambda2, LCRC and MS coefficients in non-parametric item response theory models.

The results showed that all of the item parameters estimated according to the parametric and non-parametric models from the data sets which are composed of different

number of items and sample sizes were correlated significantly and highly. However, as for ability parameters, it was determined that there is not consistency in the estimations according to the parametric models especially when the data sets are composed of five and 15 items. However, it was found that the abilities estimated from 25 item data set is correlated significantly and highly. Among of the abilities estimated by non-parametric models in all sample size and test length factors, it was concluded that ability parameters are significantly and highly correlated. Additionally, when the correlation of estimated ability parameters was examined between parametric and non-parametric models, parametric models produced consistent with non-parametric models only in the conditions of having 25 items or sample size is larger than 3000.

In the scope of this work, the another factor whose effects were examined on ability parameters were the levels of item discrimination and item difficulty parameters and according to this factor, it was observed that the abilities estimated from the four data sets that were grouped in terms of 'low and high' item difficulties and item discrimination, were correlated highly and significantly. Accordingly, it was found that both approaches had resulted in so highly and significantly correlated item and ability parameters. This finding showed that if the assumptions of parametric item response theory models are not met as a required level, the non-parametric alternative models can be used in the conditions that studied in this research.

Key Words: Parametric Item Response Theory, Non-Parametric Item Response Theory, Mokken Scaling, Sample Size, Test Length

ÖNSÖZ

Hızla gelişen bilgisayar teknolojisi, ölçme ve değerlendirme alanına katkıda bulunmuş ve karmaşık matematiksel modellemelerin gerekli olduğu Madde Tepki Kuramı yaklaşımı uygulamalarını artırmıştır. Klasik Test Kuramına göre değişmez madde parametreleri ve maddelerden bağımsız birey yetenek kestirimleri sağlayan, test eşitleme ve değişen madde fonksiyonu belirleme gibi alanlarda avantajlar sağlayan Madde Tepki Kuramı gerek eğitsel gerekse psikolojik ve klinik testlerde sıklıkla tercih edilmeye başlanmıştır. Kurama ilişkin söz konusu avantajların sağlanabilmesi ancak model veri uyumunun ve kurama ilişkin yaklaşımların karşılanmasına bağlıdır. Varsayımların yeterli düzey sağlanamadığı durumlarda kuram kapsamında yer alan alternatif yaklaşımlar kullanılmaktadır ve bunlardan biri de parametrik madde tepki kuramı modellerinin varsayımlarının gereken düzeyde karşılanmadığı durumlarda parametrik olmayan madde tepki kuramı modellerinin işe koşulmasıdır.

Bu çalışma kapsamında literatürde sıklıkla tercih edilen parametrik madde tepki kuramı modelleriyle, henüz parametrik yaklaşım kadar tanınmayan ve tercih edilmeyen parametrik olmayan madde tepki kuramı modellerinden çeşitli faktörlere göre elde edilen madde ve yetenek parametreleri karşılaştırılmıştır. Yapılan incelemelerden elde edilen bulguların ile parametrik ve parametrik olmayan madde tepki kuramı modellerinin uygulanacağı koşulların belirlenmesi konusunda araştırmacılara ve uygulayıcılara bilgi vereceği düşünülmektedir.

Bu çalışmanın yürütülmesinde her zaman destek veren, değerli görüş ve önerileriyle çalışmamı zenginleştiren ve akademik bir bakış açısı kazanmamı sağlayan değerli hocam Prof. Dr. Nizamettin KOÇ'a teşekkür eder, saygılarımı sunarım.

Ölçme ve değerlendirme alanına girdiğim andan itibaren, her türlü soruma yanıt olan, her durumda destek olan, bilgi ve deneyimleri ile önüme yeni kapılar açan Ankara Üniversitesi Eğitim Bilimleri Fakültesi Ölçme ve Değerlendirme Anabilim Dalının tüm değerli öğretim üyelerine çok teşekkür ederim. Ayrıca, tez izleme komitesinde yer alan ve görüşleriyle çalışmama yön veren, gösterdiği sabır ve anlayışla bu zor süreci benim için kolaylaştıran çok değerli hocalarım Prof. Dr. Ezel TAVŞANCIL ve Prof. Dr. Hülya KELECİOĞLU'na teşekkürlerimi ve saygılarımı sunarım. Son olarak gerek iş hayatımda gerekse özel hayatımda hep çok özel bir arkadaş ve dost olan, desteğini her zaman

hissettiğim ve sonsuz güvendiğim Arş. Gör. Seval KULA KARTAL'a da çok teşekkür ederim.

Lisansüstü eğitime başladığım andan itibaren, çabamı çabası olarak gören, beni her durumda destekleyen ve sabır ve anlayış gösteren bundan dolayı da bu çalışmanın en büyük paydaşlarından biri olan çok değerli annem ve babam –Ayşegül ve Doğan MOR ile canım kardeşim Zeki Can Mor'a teşekkürü bir borç bilirim.

Hayatıma kattıkları ile tüm sıkıntıları unutturan, en zor durumlarda bile desteğini esirgemeyen, tüm süreç boyunca her türlü yardımı sağlayıp, sorumluluklarımı azaltarak hayatımı kolaylaştıran ve minik bebeğimize bakan hayat arkadaşım Emrah DİRLİK'e sonsuz teşekkürlerimi ve sevgilerimi sunarım.



Can ođlum Denizime...

İÇİNDEKİLER

ONAY	ii
ETİK BİLDİRİM	iii
ÖZET	iv
SUMMARY	vi
ÖNSÖZ	viii
İÇİNDEKİLER	xi
ÇİZELGELER DİZİNİ	xiii
ŞEKİLLER DİZİNİ	xvii
BÖLÜM 1	1
GİRİŞ	1
Problem Durumu	1
Madde Tepki Kuramı	2
Parametrik ve Parametrik Olmayan Madde Tepki Kuramı.....	6
İlgili Araştırmalar	12
Amaç	21
Önem	22
Sınırlılıklar.....	23
Kısaltmalar	24
BÖLÜM 2	25
KURAMSAL ÇERÇEVE	25
Parametrik Olmayan Madde Tepki Kuramı	25
Parametrik Olmayan Madde Tepki Kuramı Modelleri	35
Parametrik Olmayan Madde Tepki Fonksiyonu Kestirimleri	39
Parametrik Olmayan Madde Tepki Kuramı Yaklaşımında Güvenirlik İncelemeleri... 40	
BÖLÜM 3	43
YÖNTEM	43
Araştırmanın Modeli	43
Çalışma Grubu.....	43
Veriler ve Elde Edilmesi	45
Çalışmada İncelemeye Alınan Faktörler	47
Parametrik Madde Tepki Kuramı Modellerine İlişkin Varsayımların Test Edilmesi .. 49	

Tek Boyutluluk.....	50
Yerel Bağımsızlık.....	59
Parametrik Olmayan Madde Tepki Kuramı Modellerine İlişkin Varsayımların Test Edilmesi.....	60
Tek Boyutluluk.....	60
Yerel Bağımsızlık.....	61
Madde Tepki Fonksiyonlarının Monotonluğu	61
Madde Tepki Fonksiyonlarının Kesişmemesi-Değişmez Madde Sıralaması.....	63
BÖLÜM 4	65
BULGULAR VE YORUMLAR	65
Farklı Örneklem Büyüklüğüne ve Test Uzunluğuna Sahip Veri Setlerinin PMTK ve POMTK Kapsamında Yer Alan Modellere Uyumunun Belirlenmesi	65
Farklı Örneklem Büyüklüğüne ve Test Uzunluğuna Sahip Veri Setlerinden PMTK ve POMTK Kapsamında Kestirilen Madde Parametrelerinin Karşılaştırılması	119
Farklı Test Uzunluklarına Göre Oluşturulan Veri Setlerinin PMTK ve POMTK'dan Kestirilen Madde Parametrelerinin Karşılaştırılması	130
PMTK ve POMTK Kapsamında Kestirilen Yetenek Parametrelerinin Farklı Örneklem Büyüklüğü ve Test Uzunluğu Koşullarına Göre ve Maddelerin Farklı Psikometrik Niteliklerine Göre Değişiminin İncelenmesi.....	145
Çalışmada Yer Alan Faktörlere Göre Oluşturulan Veri Setlerinin Güvenirliklerinin PMTK ve POMTK Yaklaşımlarına Göre İncelenmesi	170
BÖLÜM 5	188
SONUÇ VE ÖNERİLER	188
Sonuçlar.....	188
Öneriler.....	191
KAYNAKÇA	194
EKLER	201

ÇİZELGELER DİZİNİ

Çizelge 1.	Çalışma Grubunun Ükelere Göre Dağılımı ve Ortalama Ölçek Puanı.....	44
Çizelge 2.	AFA Sonucu Hesaplanan Bileşenler, Özdeğerleri ve Açıklanan Varyans Oranları ..	53
Çizelge 3 .	Maddelere İlişkin Faktör Yük Değerleri	55
Çizelge 4.	DFA'dan Elde Edilen Model Uyum İndeksleri	58
Çizelge 5.	500, 1000 ve 3000 Kişilik Veri Setlerinden Hesaplanan Test Puanlarına İlişkin Betimsel İstatistikler.....	66
Çizelge 6.	Seçkisiz Olarak Oluşturulan 500, 100 ve 3000 Kişilik Veri Setlerinin PMTK Modellerine Uyumu	67
Çizelge 7.	Seçkisiz Olarak Oluşturulan 500, 100 ve 3000 Kişilik Veri Setlerinin PMTK Modellerine Madde Düzeyinde Uyumu	68
Çizelge 8.	500, 1000 ve 3000 Kişilik Veri Setlerinin POMTK'ya Göre Hesaplanan Madde Güçlük Parametreleri.....	70
Çizelge 9.	Madde ve Ölçek Düzeyinde Hesaplanan Ölçeklenebilirlik Katsayıları	72
Çizelge 10.	Oluşturulan Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi.....	76
Çizelge 11.	500, 1000 ve 3000 Kişilik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler	78
Çizelge 12.	Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi.....	83
Çizelge 13.	Örneklemlere Göre Maddelerin Crit Değerlerinin İncelenmesi	84
Çizelge 14.	Kesişmeyen Madde Tepki Fonksiyonu Varsayımının P Matrisi Yöntemiyle Test Edilmesi	85
Çizelge 15.	Örneklemlere Göre Maddelerin Crit Değerlerinin İncelenmesi.....	86
Çizelge 16.	Örneklemler İçin Hesaplanan H^T Katsayıları.....	87
Çizelge 17.	Farklı Test Uzunlukları ve Örneklem Büyüklüğüne Göre Oluşturulan Veri Setlerinin PMTK Modellerine Uyumu.....	88
Çizelge 18.	Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Değerleri	91
Çizelge 19.	On Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Değerleri.....	92
Çizelge 20.	Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Değerleri.....	92

Çizelge 21.	Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları.....	94
Çizelge 22.	On Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları.....	95
Çizelge 23.	Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları.....	97
Çizelge 24.	Beş Maddelik Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi.....	101
Çizelge 25.	On Beş Maddelik Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi.....	102
Çizelge 26.	Yirmi Beş Maddelik Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi.....	103
Çizelge 27.	Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler.....	105
Çizelge 28.	On Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler.....	106
Çizelge 29.	Yirmi Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler.....	107
Çizelge 30.	Beş Maddelik Veri Setleri İçin Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları.....	110
Çizelge 31.	On Beş Maddelik Veri Setleri İçin Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları.....	111
Çizelge 32.	Yirmi Beş Maddelik Veri Setleri İçin Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları.....	112
Çizelge 33.	Beş Maddelik Veri Setleri İçin P Matrisi Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları	114
Çizelge 34.	On Beş Maddelik Veri Setleri İçin P Matrisi Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları.....	115
Çizelge 35.	Yirmi Beş Maddelik Veri Setleri İçin P Matrisi Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları.....	116
Çizelge 36.	Farklı Test Uzunlukları ve Örneklem Büyüklüğünden Oluşan Veri Setleri İçin Hesaplanan H^T Katsayıları.....	117
Çizelge 37.	500, 1000 ve 3000 Kişilik Veri Setleri İçin Uygun PMTK Modellerinden Kestirilen Madde Parametreleri.....	120

Çizelge 38.	Farklı Örneklem Büyüklüklerinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar.....	123
Çizelge 39.	Farklı Örneklem Büyüklüklerinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	124
Çizelge 40.	İki Parametrelili Lojistik Modelden ve Monoton Homojenlik Modelinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar.....	126
Çizelge 41.	İki Parametrelili Lojistik Modelden ve Monoton Homojenlik Modelinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	128
Çizelge 42.	Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük ve Ayırt Edicilik Parametreleri.....	131
Çizelge 43.	Beş Maddelik Veri Setlerinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar.....	132
Çizelge 44.	Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	132
Çizelge 45.	On Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük ve Ayırt Edicilik Parametreleri.....	133
Çizelge 46.	On Beş Maddelik Veri Setlerinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar.....	134
Çizelge 47.	On Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	135
Çizelge 48.	Yirmi Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük ve Ayırt Edicilik Parametreleri.....	136
Çizelge 49.	Yirmi Beş Maddeden Oluşan 500 ve 3000 Kişilik Veri Setlerinden Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyon Katsayısı.....	138
Çizelge 50.	Yirmi Beş Maddeden Oluşan 500, 1000 ve 3000 Kişilik Veri Setlerinden Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	139
Çizelge 51.	Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar.....	140
Çizelge 52.	On Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar.....	140

Çizelge 53.	Yirmi Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar.....	141
Çizelge 54.	Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	143
Çizelge 55.	On Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	144
Çizelge 56.	Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar.....	144
Çizelge 57.	POMTK ve PMTK Kapsamında 1000 Kişilik Veri Setlerinden Hesaplanan Madde Güçlük ve Ayırt Edicilik Parametreleri Sıralaması.....	146
Çizelge 58.	En Düşük ve En Yüksek Madde Ayırt Edicilik Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden Kestirilen Birey Yeteneklerine İlişkin Betimsel İstatistikler.....	150
Çizelge 59.	En Yüksek ve En Düşük Madde Ayırt Edicilik Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden POMTK ve PMTK'ya Göre Kestirilen Yetenekler Arasındaki İlişkiler.....	151
Çizelge 60.	En Düşük ve En Yüksek Madde Güçlük Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden Kestirilen Birey Yeteneklerine İlişkin Betimsel İstatistikleri.....	154
Çizelge 61.	En Yüksek ve En Düşük Madde Güçlük Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden POMTK ve PMTK'ya Göre Kestirilen Yetenekler Arasındaki İlişkiler.....	156
Çizelge 62.	Farklı Örneklem Büyüklüğü ve Madde Sayılarından Oluşan Veri Setlerinin PMTK Modellerine Uyumunun İncelenmesi.....	158
Çizelge 63.	Beş Maddeden ve Farklı Örneklem Büyüklüklerinden Oluşan Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler..	161
Çizelge 64.	On Beş Maddeden ve Farklı Örneklem Büyüklüklerinden Oluşan Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler.....	162

Çizelge 65.	Yirmi Beş Maddeden ve Farklı Örneklem Büyüklüklerinden Oluşan Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler.....	163
Çizelge 66.	Farklı Madde Sayılarının Oluşan Beş Yüz Kişilik Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler.....	165
Çizelge 67.	Farklı Madde Sayılarının Oluşan Bin Kişilik Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler.....	166
Çizelge 68.	Farklı Madde Sayılarının Oluşan Üç Bin Kişilik Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler.....	168
Çizelge 69.	500 Kişiden Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları.....	171
Çizelge 70.	1000 Kişiden Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları.....	173
Çizelge 71.	3000 Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları.....	174
Çizelge 72.	Beş Maddeden Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları	177
Çizelge 73.	On Beş Maddeden Oluşan Veri Setlerinden Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları	179
Çizelge 74.	Yirmi Beş Maddeden Oluşan Veri Setlerinden Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları.....	178
Çizelge 75.	Madde Güçlük ve Ayırt Edicilik Parametrelerine Göre Değişen Maddelerden Oluşan Veri Setlerinden Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları	181
Çizelge 76.	Farklı Faktörlere Göre Oluşturulan Veri Setlerinden POMTK Yaklaşımına Göre Kestirilen Güvenirlik Katsayıları.....	184

ŞEKİLLER DİZİNİ

Şekil 1.	Mokken Tarafından Önerilerin Güvenirlik Kestirim Formülü.....	41
Şekil 2.	Yamaç Birikinti Grafiği.....	54
Şekil 3.	DFA Sonucu Elde Edilen Yol Diyagramı ve T Değerleri.....	57

BÖLÜM 1

GİRİŞ

Bu bölümde araştırmanın problemi açıklanmış, ilgili araştırmalar, araştırmanın amacı, önemi, sınırlılıkları ve tanımlar belirtilmiştir.

Problem Durumu

Psikolojik özelliklerin ölçülmesi yüzyılı aşkın bir süredir sosyal bilimcilerin uğraş alanlarından biri olmuştur. Soyut ve çok boyutlu nitelik taşıyan psikolojik özellikler kendileriyle ilgili olduğu düşünülen ve gözlenebilen başka değişkenlerin aracılığı ile ölçülebilmekte ve bireylerin söz konusu değişkenlere verdikleri tepkilerden yola çıkılarak, gizil özelliklere yönelik çıkarımlar yapılmaktadır. Yapılan ölçme işlemlerinden elde edilen sonuçlardan yola çıkarak isabetli yargılara varılmasında ve ölçme işlemine karışan hata düzeyinin en aza indirgenmesinde önemli role sahip olan nitelikli ölçme araçlarını geliştirmek amacıyla çeşitli ölçme kuramları önerilmiştir (Baykul, 2000; Crocker ve Algina, 1986). Geliştirilen ölçme kuramlarından ilki Klasik Test Kuramı (KTK)'dir. Temelleri 1900'lü yılların başlarında Spearman tarafından atılan bu kuram, "Gerçek Puan Teorisi" olarak da bilinmektedir. Bu kuram ile ölçülen özelliğe ilişkin gerçek puan, ölçme sonuçlarından tahmin edilmekte ve bu kurama ilişkin temel denklem şu şekildedir:

$$\text{Gözlenen Puan (X)} = T (\text{gerçek puan}) + E (\text{Hata puanı}) \quad (1)$$

Eşitlik 1'de verilen denklem klasik test kuramına ilişkin kavramların çıkış noktasıdır ve KTK bu temel denklem üzerinden oluşturulan varsayımlara dayanmaktadır. Bu varsayımlar şu şekilde sıralanabilir (Crocker ve Algina, 1986; Lord, 1970):

- Hata puanlarının beklenen değeri sıfırdır.
- Gerçek puan ile hata puanları arasında ilişki yoktur.
- İki farklı ölçme işlemine ilişkin kestirilen hata puanları arasında ilişki yoktur.
- İki farklı ölçmeye ilişkin hata puanları ile gerçek puanlar arasında bir ilişki yoktur

KTK'nın belirtilen varsayımlarının birçok veri seti için kolayca karşılanabiliyor olması, kuramın uygulanabilirliğini genişletmiş olsa da, kuramın yapısından kaynaklanan bir takım sınırlılıklar bulunmaktadır (Baker ve Kim, 2004). Bu kuram kapsamında madde istatistiklerinin gruba bağlı olarak değişmesi, yetenek kestirimlerinin uygulanan teste bağlı olması ve test eşitleme ve madde/test yanlılığını belirleme gibi konularda yetersiz kalması nedenleriyle pratik kullanımı azalan kurama alternatif kuram arayışlarına girilmiştir. KTK'nın söz konusu sınırlılıklarını karşılamak üzere model dayanaklı bir kuram olan Madde Tepki Kuramı (MTK) ortaya atılmıştır (Embretson and Reise, 2000; Hambleton ve Swaminathan, 1985). Çalışma kapsamında temel alınan kuram olması nedeniyle bu aşamada Madde Tepki Kuramının temel özellikleri, varsayımları, modelleri, sınırlılık ve avantajları ile ilgili ayrıntılı bilgi verilmesi gerekli görülmüştür.

Madde Tepki Kuramı

1930'lu yılların sonlarında KTK'nın sınırlılıklarını ortadan kaldırdığı iddiasıyla ortaya atılan MTK, adından da anlaşılacağı üzere test maddelerine verilen yanıtların temel alınmasıyla oluşturulmuştur. Kuramın gelişimi ve kavramsal olarak temellendirilmesi, "Thurstone'un Psikolojik ve Eğitim Amaçlı Testlerin Ölçeklenmesi" makalesine dayanmaktadır ve kuram gelişimini 1943'te Lawley tarafından yeni bir model önerilmesiyle sürdürmüştür. Kuram kapsamında maddelerin özelliklerini belirlemede önemli olan "madde karakteristik eğrisi" kavramı ilk kez Tucker tarafından 1940'da kullanılmıştır (akt: Baker ve Kim, 2004; Hambleton ve Swaminathan, 1985). Kuramın gelişimine en büyük katkılardan birini Lord ve Novick 1968'de yayınladıkları 'Statistical Theories of Mental Test Scores' adlı çalışmalarında yapmış ve bu kitap içinde yer alan Birnbaum tarafından yazılan bölümler MTK'nın temellerini oluşturmuştur. Lord tarafından ogive modellerin geliştirilmesinin ardından, Birnbaum tarafından lojistik modeller kurama eklenmiştir. Ardından ise Rasch, kendi adıyla bilinen Rasch modelini kurama kazandırmıştır. Çok sayıda araştırmacı tarafından kullanılan MTK modelleri gelişimini sürdürmekte, geliştirildiği günden itibaren birçok psikolojik test uygulamasında kullanılmaktadır (van der Linden ve Hambleton, 1997).

Teknolojideki gelişmeler sayesinde kullanımı yaygınlaşan MTK modelleri, bireyin testle ölçülen gizil özellikleri ile test maddelerine verdiği tepkiler arasında bir örüntü kurmakta ve bireyin KTK'da gerçek puan olarak tanımlanan yeteneğini kestirmektedir. Madde Tepki Kuramı sayesinde test maddesi ile testi alan bireyin özellikleri ilişkilendirilmekte ve söz konusu ilişki matematiksel olarak ifade edilmektedir (Hambleton ve Swaminathan, 1985). Bu kurama göre geliştirilen testlerden elde edilen yetenek parametreleri, bireylere uygulanan testlerden bağımsız olarak kestirilebilmektedir. Bu özellik, test puanları eşitlendiğinde, bireylerin yeteneklerinin gruptan bağımsız olarak karşılaştırılmasını sağlamaktadır (Kelecioğlu, 2001). MTK'yı öncüsü olan KTK'dan ayırt eden bir başka özellik ise, her bir maddenin tüm kategorilerine ait yanıt verilme olasılığının ölçülen gizil değişkenin bir fonksiyonu olarak modellenmesidir. Madde tepki fonksiyonu (MTF) olarak isimlendirilen bu fonksiyon yetenek düzeylerine göre maddeye doğru yanıt verme olasılığını göstermektedir. Bireylerin maddelere verdiği yanıtlar ölçülen gizil değişkenin göstergesi olarak kabul edilmekte ve bireylerin yetenekleri maddelere verdikleri yanıtlara bağlı olarak kestirilmektedir (Embretson ve Reise, 2000; Hambleton, Swaminathan ve Rogers, 1991).

Madde tepki kuramının öncüsü KTK'dan ayrıldığı bir diğer nokta ise bireylere ilişkin yetenek kestirimidir. Klasik test kuramında bireylerin ölçülen gizil değişkene ilişkin yeteneği birey tarafından doğru yanıtlanan toplam madde sayısı olarak alınırken, MTK'da maddelere ilişkin ayırt edicilik, güçlük ve şans parametreleri de yetenek kestirimine dahil edilir. Bu durumda aynı sayıda maddeyi doğru yanıtlayan iki öğrenci, KTK'da olduğu gibi aynı yeterlik düzeyine sahip olmayabilir. Bu sayede ölçülen özellikle ilgili bireysel farklılıkların daha iyi ortaya konması olanaklı hale gelmektedir (Embretson ve Reise, 2000; Hambleton, Swaminathan ve Rogers, 1991). Bireysel farklılıkların daha iyi ortaya çıkarılması MTK'nın KTK'ya olan avantajlarından yalnızca biridir. Stocking (1997), MTK'nın diğer avantajlarını şu şekilde belirtmektedir: Madde özellikleri maddeyi yanıtlayan grubun özelliklerinden bağımsız olarak kestirilir, birey yeteneği de bireye uygulanan testin psikometrik özelliklerinden bağımsız olarak kestirilir. Kestirilen madde parametreleri ve birey yetenekleri kuramın varsayımları ve model veri uyumu sağlandığında değişmezlik özelliğine sahip olur. Bu avantajlara ek olarak MTK, KTK'da olduğu gibi tüm bireyler için tek bir standart hata değeri vermek yerine her yetenek düzeyi

için farklı standart hata değerleri vermekte ve KTK'dan farklı olarak kısa testlerde de yüksek güvenilirlik elde edilebilmektedir (Embretson ve Reise, 2000). Kurama ait bu avantajların ortaya çıkması için ise öncelikle kuramın varsayımlarının karşılanması gerekmektedir ve bu varsayımlardan ilki ölçülen psikolojik yapının tek boyutluluğudur. Bu varsayıma göre, madde tepki fonksiyonları yalnızca bir gizil değişkenin fonksiyonudur, bir diğer ifadeyle test maddeleri tek bir psikolojik özelliğe yönelik ölçme yapmalıdır. Test ile ölçülen bu gizil değişken eğitimsel testlerde yetenek ya da uzmanlık düzeyi olarak adlandırılmaktadır. Maddeler arasındaki istatistiksel bağlılık olarak da tanımlanan (Crocker ve Algina, 1986) bu özelliğin değerlendirilmesi için çok sayıda yöntem bulunmaktadır. Hattie (1984) çalışmasında 87 yöntemi kıyaslamış, Tate (2003) ise sık kullanılan dokuz yöntemi karşılaştırmıştır (akt. DeMars, 2010). Söz konusu varsayımın incelenmesinde sıklıkla kullanılan yöntemlerden biri ise faktör analizidir. Faktör analizi sonucunda baskın bir faktörün elde edilmesi, bir diğer ifadeyle, açıklanan varyansın büyük oranının bir faktörce açıklanması tek boyutluluğun göstergesi olarak kabul edilmektedir (Crocker ve Algina, 1986; Hambleton ve Swaminathan, 1985). Faktör analizi kadar sık kullanılan yöntemlerden biri de, Stout (1990) tarafından geliştirilen temel tek boyutluluk (essential unidimensionality) kavramını test eden DIMTEST yöntemidir. Aynı isimli bir program ile uygulanan yöntemde, yerel bağımsızlık ile tek boyutluluk kavramı birlikte ele alınmakta ve inceleme sonucunda elde edilen değerler bir takım sınır değerlere göre yorumlanarak veri setinin boyutluluğu hakkında karara varılmaktadır.

Madde tepki kuramı uygulamalarında tek boyutlulukla beraber incelenen bir diğer varsayım ise yerel bağımsızlıktır ve bu varsayım, testle ölçülen gizil değişken sabit tutulduğunda madde yanıtlarının birbirinden bağımsız olmasını ifade eder. Yerel bağımsızlık bireyin testte yer alan bir maddeye verdiği yanıtın aynı testte yer alan diğer maddelere verdiği yanıtın bağımsız olmasını gerektirir. Bu varsayım sayesinde her maddenin testin bütününe yaptığı katkıyı ayrı olarak değerlendirmek ve test düzeyindeki özellikleri madde özelliklerinin toplamıyla hesaplamak mümkün olur (Embretson ve Reise, 2000; Hambleton, Swaminathan ve Rogers, 1991). Yerel bağımsızlık varsayımı sıklıkla tek boyutluluk varsayımı içinde yer almakta ve maddelerin yerel bağımsızlığı ihlal ettiği durumlarda, bu duruma neden olabilecek teste ilişkin farklı bir boyut olduğu düşünülmektedir (DeMars, 2010). Bu nedenle yerel bağımsızlık varsayımının sağlanması

için tek boyutluluk ön koşul olarak gösterilmektedir. Hambleton ve Swaminathan (1985) yerel bağımsızlık varsayımının tek boyutlulukla olan ilişkisini belirterek, söz konusu varsayımın da tek boyutluluk gibi faktör analizi ile incelenebileceğini belirtmişlerdir. Ayrıca Yen (1984) tarafından önerilen Q_3 testi ile yerel bağımsızlık incelemesi madde çiftleri düzeyinde yapılabilmektedir. Yen istatistiğini, yerel bağımsızlıkla tek boyutluluğu eş zamanlı olarak inceleyen Stout'un analizinden ayıran nokta, bu testin madde çiftleri düzeyinde çalışıyor olmasıdır. Stout (1990) tarafından önerilen temel tek boyutluluk yaklaşımında toplam ya da ortalama kovaryanslar incelemeye alınırken, Yen istatistiğinde madde çiftleri için kalan kovaryanslar incelendiği için varsayım ihlalleri daha güçlü bir şekilde tespit edilebilir.

Kurama ilişkin üçüncü varsayım ise ölçeklenen veri seti ile model uyumunun test edilmesidir. Madde düzeyinde ve test düzeyinde uyumlar incelenmelidir. Bu incelemeler için modelin sahip olmadığı özelliklerin test edilmesi yöntem olarak uygulanabilir. Örneğin bir parametrelili lojistik model (1 PLM) ile ölçeklenen bir veri setinde yer alan maddelere ilişkin oluşturulan madde tepki fonksiyonlarının (MTF) eğimleri farklıysa ya da alt asimptot değerleri 0'dan başlamıyorsa, bu modelin veri seti için uygun olmadığı sonucuna ulaşılır. Madde düzeyindeki uyumlar, MTF'lerin incelenmesine ek olarak, gözlenen ve beklenen oranların farkına dayalı olarak incelenebilir ve aynı zamanda madde uyum indeksi olarak kullanılabilir birden çok indeks hesaplanabilir. Model düzeyindeki uyumlar için ise -2 loglikelihood değerleri incelenebilir, AIC ve BIC indeksleri hesaplanabilir (DeMars, 2010; Embretson ve Reise, 2000). Model düzeyindeki uyumlar için aynı zamanda değişmezlik incelemesi yapılabilir. Uyumlu bir model ile MTK'ya göre ölçeklenen veri setinde aranan özelliklerden biri değişmezlik özelliğidir ve yetenek ve madde parametrelerinin değişmezliği olmak üzere iki ayrı durum için değişmezlik incelemeleri yapılmaktadır. Aynı maddelerin farklı bireylere uygulanması ile madde parametrelerinin değişmezliği, aynı bireylere farklı maddelerin uygulanması ile ise yetenek değişmezliği incelenir. Yapılan bu incelemelerde her iki durum için de yüksek düzeyde değişmezlik sağlanıyorsa, model veri uyumunun sağlandığı yorumuna gidilebilir (Hambleton, Swaminathan ve Rogers 1991). Varsayımlar test edildikten sonra veri setine uyum sağlayan model belirlenerek madde ve yetenek parametreleri kestirimlerine geçilir.

Madde tepki kuramı test eşitleme, madde/test yanlılığını belirleme ve bireye uyarlanmış test geliştirme gibi karmaşık psikometri problemlerine, etkili çözüm bulmasına rağmen, bu çözümlerin anlamlılığı söz konusu varsayımların karşılanma düzeyine bağlıdır (Embretson ve Reise, 2000). Varsayımların ihlal edilmesi durumunda MTK uygulamaları ciddi düzeyde yanlılığa ve yanlış kararlar vermeye neden olabilir. Kurama ilişkin varsayımların veri setleri için karşılanmaması durumunda uygulanabilecek iki yaklaşım vardır. Bu yaklaşımlardan biri tek boyutluluk varsayımının sağlanmadığı durumda uygulanmaktadır ve test ile ölçülen özellik çok boyutlu bir nitelik taşıyorsa, tek boyutlu madde tepki kuramı modellerine karşılık çok boyutlu madde tepki kuramı modelleri işe koşulabilir. Bir diğer alternatif yaklaşım ise, Parametrik Madde Tepki Kuramı (PMTK) modellerinin uygulanamadığı durumlarda Parametrik Olmayan Madde Tepki Kuramı (POMTK) modellerinin kullanımınıdır. Parametrik olmayan madde tepki kuramını, PMTK'dan ayıran temel özellik maddelerle ölçülen gizil yetenek arasındaki ilişkinin belirli bir modele bağlı kalınarak tanımlanmamasıdır. Parametrik olmayan madde tepki kuramı, madde tepki fonksiyonlarını herhangi bir parametrik fonksiyona bağlı kalmadan kestirmeye olanak tanır ve böylelikle, parametrik madde tepki kuramına göre daha fazla veri setine uyum sağlayabilir ve deneysel verilere daha iyi uyum sağlar (Sijtsma ve Junker, 2006). Araştırmanın problem durumunun daha iyi anlaşılması amacıyla bu kısımda parametrik ve parametrik olmayan madde tepki kuramının ve bu yaklaşım kapsamında geliştirilen modellerinin ayrıntılı olarak tanıtılmasına gerek duyulmuştur.

Parametrik ve Parametrik Olmayan Madde Tepki Kuramı

Madde tepki kuramı, geliştirildiği tarihten itibaren psikometri alanında büyük ilgi gören ve çok sayıda araştırmaya konu olan bir ölçekleme yaklaşımıdır. Test performansı ile bireylerin yetenek düzeyleri arasında matematiksel bir fonksiyon tanımlayan bu kuram, adından da anlaşıldığı üzere, maddelerin yeteneklerle olan ilişkisini temel almaktadır (Stocking, 1997). Bu kuram çerçevesinde yer alan birden çok model bulunmaktadır ve bu modeller çeşitli ölçütlere göre sınıflandırılmaktadır. Söz konusu ölçütlerden biri, testle ölçülen psikolojik özelliğin boyutluluğudur. Ölçülen özelliğin niteliğine göre, tek boyutlu ve çok boyutlu madde tepki kuramı modelleri tercih edilebilir. Maddelerin puanlanmasına göre ikili ve çoklu puanlanan, doğrusallık yapısına göre doğrusal ve doğrusal olmayan

olarak ayrılan madde tepki kuramı modelleri, madde tepki fonksiyonlarının kestiriminde kullanılan yaklaşım temel alındığında ise, parametrik ve parametrik olmayan MTK modelleri olarak sınıflandırılmaktadır. Bu çalışma kapsamında parametrik MTK modelleri ile parametrik olmayan MTK modelleri arasında karşılaştırmalar yapılacağı için her iki yaklaşımın farklılıklarının aşağıda açıklanması gerekli görülmüştür.

Parametrik madde tepki kuramı modelleri ile parametrik olmayan madde tepki kuramı arasındaki temel fark her iki yaklaşımda benimsenen farklı MTF kestirim yöntemlerinden kaynaklanmaktadır. PMTK modellerinde madde tepki fonksiyonları lojistik ya da ogive fonksiyon şeklinde oluşturulmakta ve madde puanının yetenek ölçeği üzerindeki regresyonuna madde tepki fonksiyonu (MTF) ve bu fonksiyona ait eğriye ise Madde Karakteristik Eğrisi (MKE) adı verilir (Hambleton ve Swaminathan, 1985). Genellikle “S” şeklinde kestirilen bu fonksiyon, 0-1 arasında değerler alarak bireylerin yetenek düzeylerine göre maddeyi doğru cevaplamalarını olasılık olarak yorumlamaya olanak tanır. PMTK kapsamında, MTF’lerin şeklini maddeye ilişkin kestirilen parametreler belirlemektedir. Bu parametreler, madde güçlük, ayırt edicilik ve şans parametreleridir (Van der Linden ve Hambleton, 1997). Söz konusu parametrelerin PMTK ile kestirilen MTF’lere etkisini incelemek amacıyla PMTK modellerine açıklık kazandırılması gerekmektedir.

Bir parametrelili lojistik model.

PMTK yaklaşımı içinde yer alan bu model, tüm maddelerin madde tepki fonksiyonlarının eğimini eşit kabul etmektedir. Bir diğer ifadeyle, madde puanı ile ölçülen gizil yetenek arasındaki ilişkinin tüm maddeler için aynı olduğu, madde ayırt edicilik düzeylerinin testteki tüm maddeler için eşit olduğu kabul edilir ve bu çok kuvvetli bir varsayımdır. Ayırt edicilik parametreleri tek bir değere sabitlense de maddelerin güçlüklerine ilişkin herhangi bir kısıtlama yoktur ve madde düzeyinde madde güçlük parametreleri (b parametresi) hesaplanır. Bu modele dayalı olarak bir s bireyinin i maddesini doğru yanıtlama olasılığı Eşitlikte 2’de verildiği gibi hesaplanır:

$$P(X_{is} = 1 | \theta_s - \beta_i) = \frac{\exp(\alpha(\theta_s - \beta_i))}{1 + \exp(\alpha(\theta_s - \beta_i))} \quad (2)$$

X_{is} = s bireyinin i maddesine ilişkin yanıtı (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesine ilişkin hesaplanan güçlük parametresi

α = tüm maddeler için eşit olarak belirlenen madde ayırt edicilik indeksi

Eşitlik 2 incelendiğinde, 1 PLM'ye uyum gösteren bir veri setinde, bireyin yetenek düzeyi arttıkça maddeyi doğru yanıtlama olasılığının da arttığı görülmektedir. Ayrıca yetenek kestirimine dâhil edilen tek madde parametresinin madde güçlük parametresi olduğu görülmektedir. Bu modelde ayırt edicilik indeksine ek olarak, şans parametresinin ihmal edilebilir düzeyde olduğu kabulü vardır. Ayrıca bu model kapsamında, madde ayırt edicilik değerlerinin birbirine eşit olması nedeniyle, maddelerin MTF'leri kesişmemektedir (Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1985).

İki parametrelili lojistik model.

Birbaum (1968) tarafından önerilen iki parametrelili lojistik model (2 PLM), 1 PLM'den farklı olarak, madde tepki fonksiyonlarının farklı eğimlere sahip olmasına izin verir. Bir diğer ifadeyle, bu model kapsamında, madde güçlük parametresine ek olarak madde ayırt edicilik parametresi (α parametresi) hesaplanır. Hesaplanan bu parametrelerle maddelere ilişkin kestirilen madde tepki fonksiyonlarının şekli belirlenir. Bu modelde bir s bireyinin i maddesini doğru cevaplama olasılığı Eşitlik 3'teki gibi tanımlanır:

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i) = \frac{\exp(\alpha_i (\theta_s - \beta_i))}{1 + \exp(\alpha_i (\theta_s - \beta_i))} \quad (3)$$

X_{is} = s bireyinin i maddesine ilişkin yanıtı (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesine ilişkin hesaplanan güçlük parametresi

α_i = i maddesine ilişkin hesaplanan ayırt edicilik parametresi

Eşitlik 3 incelendiğinde 1 PLM'den farklı olarak madde ayırt edicilik değerlerinin madde düzeylerinde farklılık gösterdiği görülmektedir. Dolayısıyla bu durumda MTF'lerin kesişme durumu ortaya çıkabilir. Bireyin maddeyi doğru yanıtlama oranı hem madde güçlük hem de madde ayırt edicilik parametrelerine bağlıdır. Ayrıca 1 PLM'de olduğu gibi, yine şans parametresi kestirimlere dâhil edilmemektedir ve 0 olarak kabul edilmektedir (Crocker ve Algina, 1986; Hambleton ve Swaminathan, 1985).

Üç parametrelili lojistik model.

Üç parametrelili lojistik modelde, madde tepki fonksiyonunun şekli, madde güçlük ve ayırt edicilik parametresine ek olarak, maddeyi doğru yanıtlamak için gereken en düşük yetenek seviyesini belirten şans parametresi (c parametresi) kullanılarak belirlenir (Hambleton, Swaminathan ve Rogers, 1991). Dolayısıyla bireyin maddeyi doğru yanıtlama olasılığı üç parametreye bağlı olarak kestirilmektedir. Maddeyi yanıtlamak için gereken en düşük yetenek seviyesi olarak tanımlanan c parametresi, MTF'nin en düşük asimptotu olarak kestirilir ve 3 PLM'ye ilişkin olasılık fonksiyonun kestirimi Eşitlik 4'e göre yapılmaktadır.

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \frac{\exp(\alpha_i (\theta_s - \beta_i))}{1 + \exp(\alpha_i (\theta_s - \beta_i))} \quad (4)$$

X_{is} = s bireyinin i maddesine ilişkin yanıtı (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesine ilişkin hesaplanan güçlük parametresi

α_i = i maddesine ilişkin hesaplanan ayırt edicilik parametresi

γ_i = i maddesine ilişkin hesaplanan şans parametresi

Eşitlik incelendiğinde, γ_i olarak belirtilen şans parametresinin bu modelde yer aldığı görülmektedir. Bu parametrenin eklenmesi bu model kullanılarak kestirilen MTF'lerin 0'dan başlamamasına neden olmaktadır, bundan dolayı da madde güçlük parametresinin hesaplanmasında değişiklik olmaktadır. Bir ve iki parametrelili lojistik modelde, MTF'de 0.50 olasılığa denk gelen nokta maddenin güçlüğüne vermekte iken, bu modelde x ekseninde $1 + \gamma_i / 2$ 'ye denk gelen nokta madde güçlük parametresi olarak belirlenir (Hambleton ve Swaminathan, 1985).

PMTK kapsamında yer alan ve sıklıkla tercih edilen bu modeller incelendiğinde, maddelerle ölçülen gizil değişken arasındaki ilişkinin maddelerin güçlük, ayırt edicilik ve şans parametresine göre belirlendiği görülmektedir. Bu durumu modellere göre belirtmek gerekirse, 1 PLM'de maddenin doğru yanıtlanma olasılığını belirlemek için yalnızca bir parametre, madde güçlük parametresi, 2 PLM'de ise madde güçlük ve ayırt edicilik parametreleri yeterli kabul edilirken, 3 PLM'de güçlük ve ayırt edicilik parametrelerine ek olarak şans parametresi de kullanılmaktadır (Hambleton ve Swaminathan, 1985). PMTK modellerinin MTF'leri tamamen bu parametrelere bağlı olarak kestirmesi, parametre

kestiriminde görelî deęişmezlięi saęlayan bir avantaj olarak nitelendirilmektedir ve bu sayede psikolojik yapılarla maddelerin arasındaki ilişkiler uygun bir şekilde özetlenebilmektedir. Fakat bu kısıtlayıcı özellięin PMTK modellerinin yanıt fonksiyonunu belli bir formda üretmesi aynı zamanda bu modeller ile yapılan kestirimleri sınırlandıran bir durumdur (Molenaar, 2001). PMTK üzerine yapılan çalışmalarda, bu modellerin madde yanıtları ile gizil deęişken arasındaki karmaşık ilişkiyi keşfetmekte yetersiz kaldıęı belirtilmektedir (Ramsay ve Abrahamowicz, 1989). Thissen, Steinberg ve Mooney (1989) özellikle çok kategorili PMTK modellerinden elde edilen parametrelerin yorumlanmasının zor olduğunu ifade etmişler ve bu modellerden elde edilen eşik parametrelerinin madde kategori eğrilerinin şeklini ortaya çıkarmakta sadece bir araç olarak kullanılabileceğini belirtmişlerdir.

Santor ve Ramsay (1998) çalışmalarında PMTK modellerinin pratik açıdan iki temel sınırlılıęını tartışmışlar ve PMTK modelleri ile kararlı kestirimler yapılabilmesine rağmen, elde edilen sonuçların yanıltıcı olabileceğini belirtmişlerdir. PMTK modellerinin gizil deęişkenle madde yanıtları arasındaki ilişkinin tüm özelliklerini önceden tanımlanmış olarak açıklayan doğrulayıcı bir yaklaşımla kestirim yaptığı ancak kullanılan bu doğrulayıcı yaklaşımın veri setine her zaman uygun olmadığı belirtilmiştir. PMTK modellerinde kabul edilen doğrulayıcı yaklaşımın ancak ölçülen yapının belirtkelerinin çok dikkatli bir şekilde analiz edilerek oluşturulması ve hazırlanan maddelerin matematiksel modele uygunluęunun gözetildięi durumlarda geçerli olabileceęi belirtilmiştir. Böylesine sınırlayıcı ve doğrulayıcı nitelikleri karşılayan ölçme araçlarını geliştirmenin zor olduęu ve genellikle madde yanıtları ile ölçülen yapılar arasındaki ilişkinin PMTK modellerinde varsayıldıęı gibi açık bir şekilde tespit edilemedięi belirtilmiştir. PMTK modelleri ile kestirilen basit madde karakteristik eğrilerinin gizil deęişkenle madde yanıt olasılıkları arasındaki gerçek ilişkinin şeklini bozduęu, söz konusu psikolojik yapının keşfini ve bu yapıyı yorumlamayı zorlaştırdıęı da araştırmacılar tarafından ifade edilmiştir. Maddeler ve maddelerce ölçülen psikolojik yapı arasındaki gerçek ilişkinin ortaya çıkarılamaması, var olan ilişkilerin parametrik olarak kısıtlanması zayıf model veri uyumuna, modele uyum göstermeyen maddelerin ölçekten çıkarılması ile kapsam ve yapı geçerlięine ilişkin tehdit oluşturduęuna ve doğru olmayan parametre kestirimlerine yol açtıęı da ifade edilmiştir.

Araştırmacılar tarafından belirtilen PMTK'nın bir diğer sınırlılığı ise PMTK modellerinin örneklem verisinin özellikleriyle ilişkili olmasıdır. PMTK modelleri genellikle tüm örnekleme ulaşıldığı varsayımını temel alsa da, uygulamada ölçülen gizil yapıya ilişkin ranjin geniş olduğu örneklemlerde daha iyi model veri uyumu sağlandığı çeşitli çalışmalarda gösterilmiştir. Özellikle psikolojik araştırmalarda, küçük örneklemlerle çalışıldığı için, söz konusu psikolojik yapının tüm ranjına ilişkin ölçme yapılması mümkün olmamaktadır ve bu durum da PMTK modelleri ile iyi düzeyde uyum sağlayan ölçekler geliştirmeyi zorlaştırmaktadır (Santor ve Ramsay, 1998).

Parametrik madde tepki kuramı modellerinin geniş örneklem ve madde havuzu gibi sınırlılıklarının üstesinden gelmek için, daha esnek bir veri analizi yöntemi bulma amacıyla parametrik olmayan madde tepki kuramı araştırmaları ortaya çıkmıştır. POMTK modelleri PMTK modellerine göre daha esnek bir yaklaşımla modelleme yapmaktadır ve sağladığı esneklik sayesinde psikometri literatüründe gün geçtikçe daha sık tercih edilmektedir. PMTK'ya göre daha ılımlı varsayımlara sahip olan POMTK modellerinin psikometri alanında kullanım alanları Junker ve Sijtsma (2001) tarafından şu şekilde belirtilmektedir:

- POMTK modelleri PMTK modellerini daha iyi anlamaya olanak sağlar.
- PMTK modellerine zayıf uyum sağlayan ölçme uygulamaları için daha esnek modelleme çerçevesi sağlar.
- Geniş ölçekli test uygulamalarında kullanılan madde ve birey sayısından daha az sayıda madde ve bireyden oluşan veri setleri için analiz yapma olanağı sağlar.

Junker ve Sijtsma (2001) tarafından belirtilen POMTK'nın tercih edilme nedenleri bu alanda yapılan çalışmaların yönünü belirlemiştir. Özellikle Hollanda'da ve Amerika Birleşik Devletleri'nde sıklıkla araştırılan POMTK modelleri, Mokken (1971)'in çalışmalarına dayanmaktadır. Mokken tarafından yapılan POMTK çalışmaları tutum ve kişiliğin ölçülmesi üzerinde yoğunlaşmış ve Mokken'in öğrencileri olan Molenaar (1991) ve Sijtsma (1998) da tipik performansın ölçülmesinde POMTK modellerinin kullanımı üzerinde çalışmalar yapmışlardır. Araştırmacılar özellikle POMTK modellerinin küçük örneklemlerde ve kısa testlerdeki performansını incelemişlerdir. ABD'de ve Kanada'da yapılan çalışmalar ise Junker ve Sijtsma (2001)'nin belirttiği POMTK kullanım alanlarından 1. ve 2. maddenin araştırılmasına dayalı olarak gerçekleştirilmiş ve POMTK modelleri tipik performanstan çok yetenek ve başarı ölçülmesinde incelenmiştir (Douglas

ve Cohen, 2001; Junker, 2000; Ramsay, 1991; Stout, 1987, 1990). Hollanda ve ABD’de gerçekleştirilen bu öncü nitelikteki çalışmalar POMTK modellerinin alanda tanınmasını sağlamış ve bu modeller kısa sürede araştırmacılarda ilgi uyandırmıştır.

POMTK yaklaşımı genel olarak incelenecek olursa; PMTK’den farklı olarak veri yönelimli bir yaklaşım olduğu, doğrulama işlevinden çok açıklama işlevi taşıdığı, PMTK modellerine göre daha esnek bir yaklaşım olduğu ve PMTK’ya göre daha az sınırlayıcı varsayımlara sahip olduğu sonucuna ulaşılabilir. PMTK modellerinin aksine, POMTK modellerinde madde/ kategori tepki fonksiyonu önceden var olan modele bağlı kalınmadan, doğrudan ölçmelerden elde edilen veriler kullanılarak kestirilmektedir. POMTK uygulamalarında önceden tanımlanmış herhangi bir matematiksel model yoktur ve verinin var olan modele uyum sağlaması beklenmez (Ramsay, 1991; Van Schuur, 2011). Veri yönelimli bir yaklaşım olan POMTK, madde tepki kuramı çerçevesindeki temel ilişkileri inceler ve özellikle model veri uyumunun sağlanmadığı durumlarda PMTK modellerine alternatif olabilir (Junker ve Sijtsma, 2001; Stout, 2002). Ayrıca Douglas ve Cohen (2001), POMTK modellerinin maddeler ile yetenekler arasındaki ilişkiyi daha gerçekçi bir şekilde modellemeye olanak sağladığını belirtmektedirler. POMTK yaklaşımı ve bu yaklaşım içinde yer alan modeller ve modellere ilişkin incelenen varsayımlar ile verilerin bu yaklaşıma göre nasıl analiz edileceği raporun kuramsal çerçeve bölümünde ayrıntılı olarak tanıtılmış, bu kısımda ise POMTK yaklaşımını içeren ilgili araştırmalara yer verilmiştir.

İlgili Araştırmalar

POMTK alanı gelişmeye devam eden bir alandır ve özellikle Hollanda’da ve Amerika Birleşik Devletleri’nde, POMTK modelleri, varsayımları ve uygulamaları sıkça araştırılan konulardan olmuştur. Gelişmeye devam eden bir alan olduğu için literatürde, uygulamalı çalışmalardan çok, kuramsal çalışmalar yer almaktadır. Bu nedenle ilgili araştırmaların sunulması sırasında çalışmalar uygulamalı ve kuramsal olarak sınıflandırılmıştır.

Parametrik olmayan madde tepki kuramını temel alan ilk kuramsal çalışmalardan biri, Mokken modelleri için yapılan güvenilirlik ve iç tutarlılık analizlerini içeren çalışmadır. Molenaar ve Sijtsma (1984) tarafından yapılan bu çalışmada klasik test kuramındaki güvenilirlik kavramı POMTK çerçevesinde tartışılmış ve POMTK’da yapılabilecek

güvenirlilik analizleri incelenmiştir. Loevinger tarafından önerilen H katsayısı ile Guttman'ın homojenlik kavramı tartışılmış ve Guttman ölçekleme yaklaşımının ancak $H=1$ olduğu durumlarda mümkün olduğu belirtilmiştir. H katsayısının Mokken ölçekleme tekniğinde Guttman'dan farklı olarak esnetilerek kullanıldığı ve bu durumunda daha fazla veri setinin POMTK kapsamında ölçeklenmesine olanak tanınması anlamına geldiği belirtilmiştir. Ardından H katsayısı ile Cronbach alfa katsayısından çeşitli koşullara göre elde edilen değerler incelenmiş ve H katsayısının madde sayısının 10 ve daha üzerinde olduğu durumlarda Cronbach alfa katsayısından düşük çıktığını belirlemişlerdir. Ayrıca Rasch model ile ÇMM'den elde edilen parametreleri kıyasladıklarında ise, ÇMM'nin Rasch'ın sahip olduğu avantajlara sahip olmadığını belirtmişlerdir.

Güvenirlilik analizinin temele alındığı bir diğer kuramsal çalışma, Meijer, Sijtsma ve Molenaar (1995) tarafından yapılmıştır ve madde güvenirliklerinin temel alındığı bu çalışmada, madde düzeyinde güvenirlilik bilgisinin sağlanması için MTF'lerin monoton azalmayan ve kesişmeyen nitelikte olması gerektiği ifade edilmiştir. Madde düzeyinde yapılan güvenirlilik incelemelerinde, POMTK ile ölçeklenen veri setindeki madde güvenirliliğinin maddeye ilişkin ölçeklenebilirlik katsayısı ile yüksek düzeyde ilişkili olduğu, fakat aynı kavramlar olmadığı belirtilmiştir. Madde güvenirlilik katsayısının, güçlük sıralamasına göre maddenin altında veya üstünde yer alan madde ile değişimlenerek hesaplandığı ve bu nedenle H_i katsayısının tamamlayıcısı olarak nitelendirilebileceği bulunmuştur. Ayrıca MS katsayısının da POMTK kapsamında özellikle test düzeyinde güvenirlilik bilgisi verdiği belirtilmiştir. Son olarak PMTK kapsamında yer alan bir, iki ve üç parametrelili lojistik modellere göre test geliştirilirken, öncelikle teorik olarak konuya uyumlu maddelerin bir araya getirildiği ve ardından bazı maddelerin model uyumunu sağlamadığı için testten çıkarıldığı, bu durumun sonucu olarak da madde güvenirlilik indeksleri 0'dan farklı olmayan maddelerin testte yer aldığı belirtilmiştir.

Kuramın gelişimine katkıda bulunan bir diğer çalışma yine alanın öncü isimleri olan Sijtsma ve Junker tarafından 1996'da değişmez madde sıralaması varsayımının incelendiği çalışmadır. Bu çalışma kapsamında POMTK kapsamında değişmez madde sıralaması varsayımının veri setlerinde incelenebileceği üç yöntem olan P matrisi, madde kalan puanı yöntemi ve H^T katsayıları yöntemleri ile analizler yapmıştır. Madde kalan puanı gruplarına göre yapılan incelemelerin daha bilgi verici olduğu belirtilmiş ve bu yöntemlere ek olarak

Croon (1991) tarafından önerilen isotonik regresyon yönteminin MTF'lerin kesişme durumları hakkında bilgi verdiği ve kullanılabilceğini belirtilmiştir.

Kuramın kullanımlarını inceleyen bir diğer çalışma, Meijer, Egberink, Emons ve Sijtsma (2008) tarafından yapılan, ölçeklenemeyen maddelerin belirlenmesine ilişkin POMTK modellerinin kullanımını öneren bir çalışmadır. Bu çalışma kapsamında Harter's tarafından 1985 yılında geliştirilen Çocuklar İçin Öz-algılama Profili Ölçeği (Scale of Self-Perception Profile for Children) Hollanda'da yaşayan ve yaşları 7 ile 13 arasında değişen 702 çocuğa uygulanmıştır. Elde edilen veriler MSP 5 programı kullanılarak analiz edilmiştir. Yanıtlayıcılar arasında yer alan küçük çocukların yanıt örüntülerinin tutarsız olduğu fark edilmiş ve bazı gruplarda uygulamalar tekrarlanmıştır. Elde edilen bulgulara göre, ölçeğin MHM'ye uyum sağladığı belirlenmiştir. Özellikle kişi uyum indekslerinin hesaplanmasının ardından, ölçeklenemeyen maddeler için Mokken ölçekleme kapsamında incelenen H katsayılarının hesaplanması, Guttman hatalarının belirlenmesinin uygulamada kullanışlı olduğu sonucuna ulaşmışlardır.

Kuram kapsamında ölçeklenebilirlik katsayıları ve örneklem büyüklüğü arasındaki ilişkiyi inceleyerek Mokken ölçekleme yöntemi için gerekli olan en düşük örneklem büyüklüğünü belirlemeye çalışan bir araştırma ise Straat, van der Ark ve Sijtsma, tarafından 2014'de yapılmıştır. Mokken tarafından önerilen Otomatik madde seçim işlemi algoritmasının kullanıldığı bu çalışmada, örneklem büyüklüğü faktörünün maddeleri doğru ölçeğe yerleştirmeye etkisi incelenmiştir. Simülasyon yöntemi ile verilerin üretildiği çalışmada çok kategorili veriler incelemeye alınmıştır. Örneklem büyüklüğü faktörü içinde 16 düzey incelemeye alınmış, en küçük örneklem büyüklüğü 50 olarak belirlenmişken, en geniş örneklem büyüklüğü değeri için ise 3500 kişilik bir veri seti oluşturulmuştur. Test uzunluğu faktörü altında kısa testler için 10 maddelik, uzun testler için ise 20 maddelik testler oluşturulmuştur. Ayrıca yetenek düzeyleri arasındaki korelasyon düzeyleri ve madde düzeyi için ölçeklenebilirlik katsayılarının düzeyleri de çalışma kapsamında incelemeye alınmıştır. Elde edilen sonuçlar incelendiğinde, Mokken ölçekleme yöntemine göre maddelerin ait oldukları ölçeğe yerleştirilebilmeleri için, bir diğer ifadeyle OMSİ sürecinin doğru olarak işlemesi için özellikle madde düzeyindeki ölçekleme katsayısının önemli olduğu bulunmuştur. Ayrıca Mokken ölçeklemeye yaklaşımını kullanacak araştırmacılara, örneklem büyüklüklerinin en az 1500 ve üzerinde olması uyarısında bulunulmuştur. Bu

durumun gerekçesi olarak ise, madde düzeyindeki ölçeklenebilirlik katsayıları için belirlenen sınır değerlerin örneklem büyüklüğünün 1500'ün üzerinde olduğu durumlarda daha iyi incelenebildiği belirtilmiştir. Madde düzeyindeki ölçeklenebilirlik katsayıları için sınır değerinin düşük örneğin 0.3, seçildiği durumlarda, 250 kişiden oluşan bir örneklemden dahi etkili sınıflandırmaların yapılabildiği belirlenmiştir. Ayrıca madde ayırt edicilik indekslerinin yüksek olması örneklem büyüklüğünün 250 ve üstü olduğu durumlarda bile OMSİ'nin yanlış sınıflandırma oranının düşük olduğu bulunmuştur. Son olarak test uzunluğu ve gizil değişken sayısının OMSİ'nin doğru sınıflandırma performansı üzerindeki etkisinin ihmal edilebilir düzeyde olduğu çalışma kapsamında ulaşılan sonuçlardır.

Son zamanlarda yapılan ve POMTK'nın gelişime katkıda bulunan bir diğer çalışma ise, Kuijpers, van der Ark, Croon ve Sijtsma (2016) tarafından yapılan ölçeklenebilirlik katsayılarının standart hatalarına ilişkin bir simülasyon çalışmasıdır. Çok kategorili puanlanan maddeler için hesaplanan H katsayılarının örneklemden etkilendiği belirtilmiş ve bu nedenle çok kategorili puanlanana maddelerden oluşan veri setleri için hesaplanan H katsayılarına ilişkin standart hatalarının hesaplanmasının gerekli olduğu vurgulanmıştır. Yapılan simülasyon çalışmaları sonucunda, yalnızca oluşturulan madde adım fonksiyonlarındaki adımların birbirlerine çok benzer olması ve örneklem büyüklüğünün 200'den az olması durumlarında H katsayıların düşük oranda pozitif yanlılıkta kestirildiği belirlenmiş ve bu durum dışındaki durumlarda H katsayılarının ihmal edilebilir düzeyde yanlılığa sahip olduğu bulunmuştur. Ayrıca maddelerdeki seçenek sayısının azaltılması da H katsayılarının daha az yanlı kestirilmesini sağlamaktadır. Bu nedenle özellikle iki kategorili olarak puanlanan maddelerde madde ayırt edicilik parametrelerinin yüksek olması durumunda H katsayılarına ilişkin yapılan nokta kestirimlerin %95'inin güven aralığı içinde olduğu belirlenmiştir.

Parametrik olmayan madde tepki kuramının gelişimine katkıda bulunan kuramsal çalışmalara ek olarak, alanda toplanan veriler üzerinden gerçekleştirilen uygulamalı çalışmalar da alanyazın da mevcuttur ve sağlık bilimleri üzerine Amerika Birleşik Devletleri'nde yapılan çalışmalar uygulamalı çalışmaların çoğunluğunu oluşturmaktadır. Parametrik ve parametrik olmayan MTK modellerinden elde edilen kestirimleri karşılaştıran ve uygulama durumlarını inceleyen çalışmalardan biri, Meijer ve Baneke (2004) tarafından yapılmıştır. Bu çalışmada, kişilik ve psikopatoloji ölçeklerine

POMTK'nın uygulanabilirliđi incelenmiř ve parametrik modellerden elde edilen parametrelerle kıyaslamalar yapılmıřtır. Elde edilen sonuçlar incelendiđinde, PMTK'dan elde edilen bilgilere gre POMTK ile madde dzeyinde daha fazla bilgiye ulařıldıđı ve PMTK kapsamında st asimptotu 0'a ulařmadıđı iin incelenemeyen ve lek dıřında tutulan maddelerin aslında yksek ayırt edicilik deđerine sahip olduđu ve POMTK ile leklenebildiđi bulunmuřtur. Ayrıca alıřma kapsamında psikopatoloji maddelerini leklemede kullanılan POMTK modellerinin basit olması arařtırmacılar tarafından avantaj olarak nitelendirilmiř ve yorum kolaylıđı sađladıđı belirlenmiřtir. Sonu olarak da zellikle tipik performansın lldđ durumlarda, istatistiksel model uyum testlerini incelemek yerine ncelikli olarak POMTK'ya dayalı analizlerin yapılmasının hem lek hakkında hem de llen yapı hakkında daha fazla bilgi vereceđi sonucuna ulařılmıřtır.

Parametrik ve parametrik olmayan madde tepki kuramı modellerini kıyaslayan bir bařka alıřma Dyehouse (2009) tarafından gerekleřtirmiřtir. Arařtırmacı alıřmasında sıralama dzeyinde lme yapan bir alternatif deđerlendirme leđini hem POMTK hem de PMTK modellerine gre analiz etmiřtir. alıřmada PMTK modellerinden Genelleřtirilmiř Dereceli Aımlayıcı Model (Generalized Graded Unfolding Model) ile POMTK modellerinden Mokken leđi analizi kullanılmıřtır. 4449 đrenciye uygulanmıř olan 20 maddelik leđin POMTK varsayımlarına uyumu incelenmiř ve maddelerin ođunluđunun monoton homojenlik modeline uyum sađladıđı ancak ift monotonluk modelinin birok madde iin uyumlu olmadıđı belirlenmiřtir. Genel olarak model veri uyumunun Genelleřtirilmiř Dereceli Aımlayıcı Model iin daha yksek dzeyde olduđu belirlenmiřtir.

Uygulamalı alıřmalardan biri de Zhou (2011) tarafından gerekleřtirilen parametrik ve parametrik olmayan madde tepki kuramı modellerini gerek bir veri zerinde inceleyen arařtırmadır. alıřmanın amacı parametrik ve parametrik olmayan madde tepki kuramı modellerini kıyaslamaktır. Arařtırmacı alıřmanın amacına uygun olarak ABD'de de sıklıkla tercih edilen Positive and Negative Syndrome Scale (Pozitif ve Negatif Sendrom leđi) adlı řizofreni tanısında kullanılan ve 30 adet ok kategorili maddeden oluřan leđi veri toplama aracı olarak kullanmıřtır. Ticari bir lme merkezi tarafından uygulanan leđin, 9204 kiřiden oluřan veri seti kullanılarak, sz konusu lek PMTK modellerinden Modifiye Edilmiř Derecelendirilmiř Tepki Modeline POMTK

modellerinden ise Kernel düzgünleştirmesine göre ölçeklenmiştir. Farklı örneklem büyüklüklerine göre yapılan kestirimlerde elde edilen parametreler incelenmiş ve örneklem büyüklüğünün madde ve birey parametrelerine etkisi modellere göre analiz edilmiştir. Çalışmanın sonucunda örnekleme yer alan birey sayısı arttıkça PMTK modellerinin daha tutarlı sonuçlar verdiği, fakat ölçek ile ölçülen yapının ilişkisinin POMTK modelleri ile daha ayrıntılı olarak incelenebileceği bulgusuna ulaşılmıştır. Ayrıca POMTK modeli ile madde sayısı az olduğunda bile, PMTK modeline göre daha detaylı bilgilere ulaşıldığı belirtilmiştir.

Mokken ölçekleme tekniğini uygulamalı olarak inceleyen bir başka çalışmada ise, uluslararası düzeyde kullanılan Ulusal Sağlık Çıktıları Ölçeği (The Health of Nation Outcome Scale) kullanılmış ve daha önceki çalışmalarda PMTK modelleri ile ölçeklenen 80161 kişilik veri seti POMTK'ya göre ölçeklenmiştir (Muncer ve Speak, 2016). Mokken analizine ek olarak doğrulayıcı faktör analizinin de uygulandığı çalışmada, ölçeğin boyutluluğu her iki yönteme göre incelenmiş ve her iki yöntemden ortak olarak elde edilen sonuç söz konusu ölçeğin varsayıldığı gibi 12 ölçekten oluştuğu değil, iki faktörlü bir yapıda olduğudur. Belirlenen iki ölçek de Mokken modelleri içinde yer alan MHM'ye uyum sağlamış, madde değişmezlik varsayımı ise birkaç maddeden dolayı ihlal edilmiştir. Bu maddeler veri setinden çıkarıldığında ise, ölçeğe ilişkin hesaplanan H katsayısında azalma olduğu belirlenmiştir. Bulunan sonuçlara göre ölçeğin revize edilmesi önerilmiştir.

POMTK alanında yapılan çalışmaların odak noktalarından biri de model veri uyumu olmuştur. PMTK yaklaşımında yer alan modellere ilişkin model veri uyumunun beklenen düzeyde olmadığı durumlar için POMTK kullanımı önerilen uygulamalardan biridir ve araştırmacılar hangi durumlarda POMTK modellerinin kullanımının daha uygun olduğu üzerine çalışmalar yapmışlardır. Bu çalışmalardan biri Zhao (2008) tarafından yapılmıştır. Madde düzeyinde model veri uyumunun incelendiği çalışmada, model veri uyumunun yeterli düzeyde sağlanmamasının bireylerin yetenek düzeylerine göre gruplandırılmasının üzerindeki etkileri incelenmiştir. Model uyumunun grafiksel gösterimleri için parametrik olmayan madde tepki kuramından faydalanılmış ve gerçek bir veri seti üzerinde çalışılmıştır. Model veri uyumunun-uyumsuzluğunun yeterince incelenmediği konusunda eleştiri getiren araştırmacı, POMTK ile kestirilen madde tepki fonksiyonlarının model veri uyumunun geliştirilmesi için kullanılması gerektiğini belirtmiştir. Bu alandaki bir başka

çalışma ise Liang (2010) tarafından gerçekleştirilmiştir. Bu çalışmada MTK uygulamalarında model veri uyumunu bozan nedenlerden birinin de PMTK modellerine uymayan maddeler olduğunu belirtilmiş ve model veri uyumunu incelemek için POMTK kapsamında geliştirilen RISE istatistiği incelenmiştir. Model veri uyumunu belirlemeye yönelik geleneksel yöntemlerin örneklem büyüklüğüne duyarlı olması ve geniş örneklemelerde küçük sapmaları bile uyumsuzluk olarak nitelendirilmesi araştırmanın motivasyon kaynağı olmuştur. Bu çalışmada hem simülasyon hem de gerçek veri kullanılmış, simülasyon çalışmasında örneklem büyüklüğü, test uzunluğunu, yetenek dağılımı ve kullanılan model faktör olarak almıştır. POMTK kapsamında geliştirilen RISE istatistiğinden elde edilen sonuçlar kıyaslanmış ve RISE'in geleneksel model veri uyumu istatistiklerine göre çok büyük örneklemelerde bile daha az sayıda maddeyi uyumsuz olarak belirlediği bulgusuna ulaşmıştır.

POMTK kapsamında yer alan madde tepki fonksiyonu kestirim yöntemleri de literatürde araştırılan konulardan biri olmuştur. Hem gerçek hem de simülatif verilerle yapılan çalışmalarda farklı kestirim yöntemleri performansları bakımından kıyaslanmıştır (Lee, 2007; Ramsay, 1991; Ramsay ve Abrahamowicz, 1989; Ramsay ve Winsberg, 1991; Rossi, Wang, ve Ramsay, 2002). Birçok simülasyon çalışmasında, iki ve üç parametrelili lojistik modele uygun veriler üretilmiş ve kestirilen madde tepki fonksiyonları ile parametrik olmayan yöntemlerle kestirilen madde tepki fonksiyonları kıyaslanmıştır. Çalışmalar tutarlı sonuçlar vermiştir. Her ne kadar örneklemde yer alan birey sayısı arttıkça, madde tepki fonksiyonu kestirimleri daha hatasız olsa da, standart normal yetenek dağılımında 1000 kişilik veri seti verisinin uç yetenek düzeyleri olarak tanımlanabilecek +2 ve -2 yetenek düzeylerinde PMTK modellerinin gerçek madde tepki fonksiyonunu üretecek nitelikte olmadığı bulunmuştur.

POMTK'ya dayalı yurt dışında birçok çalışma yapılmış olmasına rağmen, yapılan alan yazın taraması sırasında, POMTK'yı temel alan yerli çalışmaların çok sınırlı olduğu görülmüştür. POMTK'yı değişen madde fonksiyonu analizinde kullanan çalışmalar bulunmaktadır ancak bu çalışmalarda POMTK modellerinden çok, değişen madde fonksiyonu belirleme yaklaşımları içinde yer alan parametrik olmayan yöntemler karşılaştırılmıştır. Değişen madde fonksiyonu belirleme yöntemleri bu çalışma kapsamı dışında tutulduğundan, söz konusu çalışmalara bu rapor kapsamında yer verilmemiştir.

POMTK modellerinin inceleyen ve Türkiye’de yapılan sınırlı çalışmalardan biri Koğar (2014) tarafından gerçekleştirilmiştir. Araştırmacı çalışmasında, basit ve iki boyutlu yapılarda, çeşitli örneklem büyüklükleri, test uzunlukları ve boyutlar arası korelasyon değerlerinde, tek boyutlu POMTK, tek boyutlu PMTK ve çok boyutlu MTK elde edilen madde parametreleri, maddelere ait model veri uyumları ve teste ait model veri uyumlarını belirlemeyi amaçlamıştır. Simülatif verilerle yapılan çalışmada, örneklem büyüklüğü, farklı test uzunlukları ve model veri uyum indeksleri kıyaslanmıştır. Çalışmanın bulguları incelendiğinde, tüm simülasyon düzeneklerinde tek boyutlu POMTK modellerinin düşük hatalar ürettiği sonucuna ulaşılmıştır. Ayrıca tek boyutlu PMTK’da a parametresinin, örneklem büyüklüğündeki ve test uzunluğundaki artış ile birlikte azalma eğilimi gösterdiği fakat hiçbir simülasyon koşulundaki değişimin, b parametresini etkilemediği belirlenmiştir.

POMTK ile ilgili Türkiye’de yapılan bir başka çalışma ise Şengül Avşar (2015) tarafından yapılan çalışmadır. Çalışma kapsamında çok kategorili puanlanan maddeler farklı test koşullarında POMTK yaklaşımı ile ölçeklenmiş ve maddelerin psikometrik nitelikleri incelenmiştir. Veriler simülasyon yöntemiyle üretilmiştir ve test uzunluğu, dağılım özelliği ve örneklem büyüklüğü faktörleri incelemeye alınmıştır. Yapılan incelemelerde madde sayısı, örneklemin dağılım şekli ve örneklem büyüklüğü faktörlerinin POMTK modellerinden biri olan Monoton Homojenlik Modeline uyumu etkilediği sonucuna ulaşılmıştır. Örneklem büyüklüğündeki artışın maddelerin ayırt edicilik düzeylerini artırdığı belirlenmiş, madde sayısındaki ve örneklem büyüklüğündeki artışın testlerin güvenilirliğini de yükselttiği bulunmuştur. Özetle farklı test koşullarında POMTK yaklaşıma göre çözümlenmelerin yüksek güvenilirlik değerlerine sahip olan testler ürettiği sonucuna ulaşılmıştır.

POMTK’ya ilişkin yapılan çalışmalar genel olarak incelendiğinde, model-veri uyumu çalışmalarının yoğunlukta olduğu, farklı madde tepki fonksiyonu kestirme yöntemlerine ve PMTK ile POMTK modellerine ilişkin kıyaslamalar üzerine çalışmalar yapıldığı belirlenmiştir. Yapılan çalışmalar incelendiğinde, POMTK modellerinin PMTK modellerine tamamen alternatif olması değil, her iki yaklaşımın da gereken durumlarda birlikte kullanılmasının uygun olduğu sonucuna ulaşıldığı görülmektedir. Bir diğer ifadeyle, yapılan çalışmaların tümünde PMTK yerine her durumda POMTK’nın kullanılmasını değil, her iki yaklaşımın da kendine özgü özelliklerinden faydalanılarak

yapılan çalışmaların geliştirileceği, ölçülen özelliğin daha iyi anlaşılacağı ve model veri uyumsuzluğu sorununa bir ölçüde çözüm bulunacağı savunulmaktadır.

PMTK modellerinin doğrulayıcı yaklaşımından önce POMTK modellerinin açılımlayıcı yaklaşımının kullanılması hem ölçülen özelliğin hem de veri yapısını daha iyi anlaşılmasını sağlayacağı belirtilmektedir. Ayrıca POMTK modelleri kapsamında veri yapısına dayalı olarak kestirilen madde tepki fonksiyonlarının ölçülen yapıya ilişkin daha fazla bilgi verdiği sonucuna ulaşılmış ve bu alanda özellikle deneysel çalışmaların yapılması gerekliliği alanda öncü araştırmacılar tarafından önerilmiştir (Christensen ve Kreiner, 2010; Junker, 2000; Van Schuur, 2011; Sijtsma ve Molenaar, 2002).

POMTK'nın model veri uyumunun sağlanamadığı durumlarda PMTK'ya göre daha iyi performans gösterdiği, gerçeğe daha yakın madde tepki fonksiyonları üretme olanağı sağlaması ve daha az sayıda madde ve örneklem ile PMTK'dan daha az hatalı kestirimler yapmaya olanak sağlaması, bu yaklaşımın araştırmacılar tarafından bilinmesi ve uygulanması gerekliliğini ortaya çıkarmıştır. Ayrıca POMTK'nın Junker ve Sijtsma (2001) tarafından belirtilen niteliklerinden biri olan PMTK modellerinin daha iyi anlaşılması amacıyla kullanımı da her iki yaklaşımı kıyaslayan çalışmalara ihtiyaç duyulduğunu belirtmektedir. Özellikle Türkiye'de bu alanda yapılmış olan çalışma sayısının çok sınırlı olması bu gerekliliği daha da arttırmaktadır. Hem POMTK'nın keşfedilmeye açık bir alan olması hem de Türkiye'de bu alanda oldukça az sayıda çalışma yapılmış olması ve dolayısıyla ülke çapındaki araştırmacılar tarafından bu alanın yeterince tanınmıyor olması göz önünde bulundurularak, POMTK ve PMTK modellerini kıyaslayan bir çalışmanın yapılmasının hem kurama hem de uygulamaya katkı sağlayacağı düşünülmektedir. Literatürde çalışılan çeşitli faktörler olan MTK modelleri, örneklem büyüklükleri, yetenek dağılımı ve maddelerin psikometrik nitelikleri gözetilerek PMTK ve POMTK modellerinin madde ve yetenek kestirimlerinin incelenmesi uygulamada ortaya çıkabilecek durumlarda araştırmacılara yol gösterici olabilecektir. Özellikle uygulamada ortaya çıkabilecek çeşitli durumların model kestirimleri üzerindeki etkilerinin incelenmesi, araştırmacılara kendi çalışma durumlarına göre model seçmeleri ve olası model veri uyumsuzluğu durumunda alternatif yollara başvurmaları konusunda yönlendirici olabilecektir. Tüm bu nedenlerle, bu çalışma kapsamında tek boyutlu ve iki kategorili puanlanan verilerde, çeşitli örneklem büyüklükleri, test uzunlukları ve maddelerin değişen psikometrik özelliklerinde, POMTK

ile PMTK yaklaşımlarından elde edilen madde parametrelerini ve birey yeteneklerini incelemek bu çalışmanın problemini oluşturmaktadır.

Amaç

Bu çalışmanın temel amacı, iki kategorili puanlanan verilerde PMTK modellerinden bir, iki ve üç parametrelili lojistik model ile POMTK kapsamında yer alan Mokken ölçekleme modellerinin, örneklem büyüklüğü, test uzunluğu ve maddelerin psikometrik özellikleri faktörlerine göre kestirilen madde ve birey yetenek parametrelerinin karşılaştırılmasıdır. Bu genel amaç kapsamında araştırmada yanıt aranan sorular şu şekildedir:

1. Araştırma kapsamında farklı örneklem büyüklüğü ve test uzunluğuna göre oluşturulan veri setlerinin,
 - a. PMTK modellerine,
 - b. POMTK modellerine uyumları nasıldır?
2. İncelenen veri setinde yer alan maddelerin, PMTK modellerinden bir-iki ve üç parametrelili lojistik modelden en iyi uyumu sağlayan model ile POMTK modellerinden uyumlu olduğu belirlenen model kullanılarak kestirilen madde ayırt edicilik ve güçlük parametreleri,
 - a. Örneklem büyüklüğüne,
 - b. Test uzunluğuna göre nasıl farklılaşmaktadır?
3. İncelenen veri setinde yer alan bireylere ilişkin PMTK modellerinden bir-iki ve üç parametrelili lojistik modelden en iyi uyumu sağlayan model ile POMTK modellerinden uyumlu olduğu belirlenen model kullanılarak belirlenen yetenek kestirimleri,
 - a. Maddelerin psikometrik niteliklerine,
 - b. Test uzunluğuna,
 - c. Örneklem büyüklüğüne göre nasıl farklılaşmaktadır?
4. Çalışma kapsamında incelenen faktörlere ilişkin oluşturulan veri setlerinin ölçüklendiği yaklaşıma göre güvenilirlik düzeyleri nasıldır?

Önem

Madde tepki kuramı modelleri ve uygulamaları psikolojik ölçmelerde birçok yeniliğe olanak sağlamış ve karmaşık psikometri problemlerine çözümler üretmiştir. Bireylerden bağımsız olarak kestirilen madde parametreleri ve maddelerden bağımsız olarak kestirilen yetenek parametreleri ile test eşitleme, bireye uyarlanmış test uygulamaları ve madde yanlılığı gibi önemli psikometrik konular bu kuram çerçevesinde araştırılmakta ve uygulamaya dönüştürülmektedir. Bu avantajlarına rağmen PMTK modellerinin parametrik doğasından kaynaklanan doğrulayıcı özelliği modellerin esnekliğini sınırlandırmakta ve gerçek uygulamalarda çeşitli nedenlerden dolayı model veri uyumunun sağlanmasını zorlaştırmaktadır. Yeterli düzeyde sağlanamayan model veri uyumu ise PMTK'nın söz konusu avantajlarını ortadan kaldırmakta, modele uyum sağlamayan maddelerin ölçekten çıkarılmasına neden olmakta ve bu nedenle aynı zamanda geçerliğe tehdit oluşturmaktadır. PMTK'nın sınırlayıcı özelliğine yönelik daha esnek modelleme yaklaşımlarının aranması sonucu ortaya çıkan POMTK ise daha zayıf varsayımlara sahip olması sayesinde daha fazla veri setine uyum sağlayabilmektedir. Doğrulayıcı bir özellikten çok açıklayıcı özellik taşıyan POMTK modelleri yanıt fonksiyonlarının belli bir forma sahip olmasını gerektirmediğinden, verilerin doğal yapısının değişmesine yol açmadan ölçülen gizil değişken hakkında daha gerçekçi bilgiler sağlamaktadır.

İlk ortaya çıkışı 1970'lere uzanan POMTK yaklaşımının teknolojik olanaklarla daha fazla kullanım alanı bulması ile PMTK'nın yetersiz kaldığı durumlarda kullanabilecek alternatif bir yaklaşım haline gelen POMTK uygulamalarının incelenmesi hem psikoloji hem de ölçme uygulamalarına katkı sağlayabilecektir. POMTK modellerinin incelenmesi ve uygulanması ile psikolojik yapılarla maddeler arasındaki ilişkiler daha ayrıntılı olarak incelenebilecek, böylelikle psikolojik ölçmenin en zor aşamalarından biri olan psikolojik yapıların tanımlanması konusunda yeni bilgiler edinilebilecektir. POMTK yaklaşımı ile PMTK'nın kıyaslanması her iki alanın da daha iyi anlaşılmasını sağlayarak, olası problemlere çözüm üretme olanağı sunabilecektir. Her iki yaklaşımında özelliklerinin ve farklı durumlardaki uygulamalarının incelenmesi araştırmacılara hem teorik hem de uygulama açısından yeni bakış açıları sunacaktır. Diğer yandan POMTK yaklaşımı ve uygulamaları dünya çapında incelenen bir alan olmasına rağmen, yapılan alan yazın çalışmasında PMTK ile POMTK'yı kıyaslayan çalışmaların sayısı yeterli düzeyde olmadığı

belirlenmiştir. Türkiye’de de dünya ile benzer bir durum söz konusudur ve POMTK konusunda yapılan çalışmalar çok sınırlıdır. MTK kapsamında yer alan bir diğer yaklaşım olan çok boyutlu madde tepki kuramı uygulamaları ülke çapında yer bulmaya başlamış olsa da, POMTK konusunda aynı durum geçerli değildir. Araştırmacılar gerek veri analizinin yapılacağı paket program bulma zorluğu gerekse PMTK’den oldukça farklı yorumlama gerektirmesinden dolayı, POMTK yaklaşımları konusunda çekimser kalmaktadır. Hem ülke çapında yapılan sınırlı araştırmalara bir yenisini katarak var olan bilgi birikimini geliştirmek hem de gerçek uygulamalara yön verici olmak amacıyla POMTK yaklaşımını geleneksel PMTK ile kıyaslayan bir çalışmanın yapılmasının alana yenilik getirici bir katkı sağlayacağı beklenmektedir.

Bu çalışma sonucunda elde edilen bulgularla araştırmacılar kendi veri setlerinin özelliklerini bu çalışma kapsamında oluşturulan koşullarla ilişkilendirebilecek, böylelikle çalışmalarında hangi yaklaşımın kullanılmasının daha uygun olduğu hakkında fikir sahibi olabileceklerdir. Bu durum da olası model veri uyumsuzluklarının önüne geçmeyi, madde yanıt fonksiyonlarının daha detaylı oluşturulması ile ölçülen yapı ile madde özelliklerinin daha etkili bir şekilde ilişkilendirilmesini ve farklı örneklem büyüklüklerinin kestirimler üzerindeki etkisi incelenmesini sağlayacaktır. Ayrıca çalışma kapsamında faktör olarak kabul edilen test uzunluğu, maddelerin farklı psikometrik özelliklerinin ve örneklem büyüklüğünün PMTK ve POMTK kestirimlerini nasıl etkilediği incelenmiş, böylece her iki yaklaşıma ilişkin daha detaylı bilgiler elde edilmiştir. Özetle bu çalışmanın hem araştırmacılara veri setlerine uygulayacakları modellere ilişkin yeni bakış açıları kazandırabilecek olması, hem de PMTK modellerinde ortaya çıkabilecek problemlere alternatif çözümler üretebilecek olması, söz konusu çalışmayı hem kuramsal düzeyde hem de uygulama düzeyinde önemli kılacaktır.

Sınırlılıklar

Bu çalışma;

1. İki kategorili puanlanan maddelerle sınırlıdır. Çok kategorili puanlanan maddeler farklı PMTK modelleri gerektirdiğinden dolayı çalışma kapsamı dışında tutulmuştur.

2. PMTK modellerinden bir-iki ve üç parametrelili modeller ile sınırlandırılmıştır. Dört parametrelili lojistik model ve ogive modellerin kullanım sıklıklarının görelili olarak az olması nedeniyle çalışma kapsamı dışından bırakılmıştır.
3. Çalışmada ele alınan faktörler örnekleme büyüklüğü, test uzunluğu ve maddelerin değişen psikometrik özellikleri ile sınırlandırılmıştır. Dağılımların çarpıklığı faktörü de çalışma kapsamında incelenmeye karar verilmiş, ancak gerçek verilerden çarpık dağılımlar üretilmediği için bu faktör çalışma dışında bırakılmıştır.
4. Değişen madde psikometrik nitelikleri faktörü içinde maddeler güçlük ve ayırt edicilik değerlerine göre gruplandırılmış, şans parametresi çalışma kapsamı dışında tutulmuştur.

Kısaltmalar

Bu çalışma kapsamında kullanılan kısaltmalar ve anlamları aşağıda verilmiştir:

AFA: Açıklayıcı Faktör Analizi

ÇMM: Çift Monotonluk Modeli

DFA: Doğrulayıcı Faktör Analizi

GA: Otomatik Madde Seçim İşlemi Yönteminde Kullanılan Generic Algoritması

MHM: Monoton Homojenlik Modeli

MKE: Madde Karakteristik Eğrisi

MTF: Madde Tepki Fonksiyonu

MTK: Madde Tepki Kuramı

NOR: Otomatik Madde Seçim İşlemi Yönteminde Kullanılan Normal Algoritması

OMSİ: Otomatik Madde Seçim İşlemi

PMTK: Parametrik Madde Tepki Kuramı

POMTK: Parametrik Olmayan Madde Tepki Kuramı

1 PLM: Bir Parametrelili Lojistik Model

2 PLM: İki Parametrelili Lojistik Model

3 PLM: Üç Parametrelili Lojistik Model

BÖLÜM 2

KURAMSAL ÇERÇEVE

Parametrik Olmayan Madde Tepki Kuramı

Parametrik olmayan madde tepki kuramına dayalı olarak gerçekleştirilen Mokken Ölçek Analizi (Mokken Scale Analysis), test verilerini ölçeklemede sıkça kullanılan tekniklerden biridir. Rasch ya da faktör analizi kadar popüler olmasa da, Google Akademik veri tabanında 5000'e yakın yayın kaydı bulunan bu teknik, maddelerin ölçeklenmesinde PMTK modellerine göre daha esnek bir şekilde işlemekte ve parametrik modellerin doğrulayıcı yaklaşımından farklı olarak açımlayıcı bir yaklaşıma göre ölçekleme yapmaktadır (Straat, van der Ark ve Sijtsma, 2014). İkili puanlanan maddeler için Mokken (1971) tarafından geliştirilen bu yaklaşım, Molenaar (1991, 1997) tarafından çok kategorili puanlanan maddelere uyarlanmıştır. Ardından Sijtsma ve Molenaar (2002), POMTK'nın genel çerçevesini oluşturmuş, Meijer ve Baneke (2004) ise bu yaklaşımın PMTK'ya alternatif kullanımını örneklendirmiştir.

Ortaya atıldığı günden itibaren PMTK'ya göre daha az kısıtlayıcı olduğu ve verinin yapısı hakkında açımlayıcı bilgiler sağladığı için araştırmacıların ilgisini çeken POMTK ölçekleme yaklaşımı olarak, bir çok farklı alanda kullanılmaktadır. Pazarlama (Pass ve Molenaar, 2003), sağlık ve tıp, siyasal bilimler (Van Schuur, 2003), psikoloji (alterman, Cacciola, Habing ve Lynch, 2011) ve sosyoloji (Gow, Watson, Whitmean ve Dreary, 2011) bu ölçekleme türünün kullanıldığı alanlardandır (akt. Straat, van der Ark ve Sijtsma, 2014). Gelişen teknolojiyle paralel olarak kullanımı artan Mokken ölçekleme yaklaşımı Gutmann ölçeklemesi ile PMTK yaklaşımı arasında yer almaktadır. Mokken ölçekleme tekniği, testi alanların yanıt örüntüsünü tek ve gizil bir değişkenin göstergesi olarak kabul etmekte ve Gutmann ölçekleme yaklaşımının parametrik olmayan olasılıklı modelini sunmaktadır. Belirleyici (deterministic) bir yaklaşımdan ziyade olasılıksal bir yaklaşıma sahip olması ve gizil değişkenlerin çok az sayıda madde ile temsil edilmesinin zorunlu olduğu durumlarda dahi uygulanabilmesi bu yaklaşımın diğer ölçekleme tekniklerine göre avantajlarından

(Van Schuur, 2011). Faktör analizi ve güvenilirlik analizleri maddeleri paralel olarak kabul edip ölçekleme yapmakta iken, Mokken ölçekleme tekniği, 1940'lı yıllarda Ferguson (1941) ve Carrol (1945) tarafından da belirttiği gibi, maddelerin frekans dağılımının farklı olduğunu kabul etmekte ve bu sayede faktör analizinde zor olan ikili puanlanan verilerin analizlerini olanaklı hale getirmektedir (akt. Sijtsma ve Molenaar, 2002).

Parametrik olmayan madde tepki kuramı yaklaşımını, PMTK'dan ayıran iki özelliği vardır. Bunlardan ilki PMTK'ya göre daha geniş bir teorik çerçeve sahip olması ve bu sayede MTK'nın genel olarak daha iyi anlaşılmasını sağlamasıdır. Diğer özellik ise, PMTK ile model veri uyumuna dayalı olarak doğrulayıcı bir yaklaşımla veri setleri analiz edilirken, POMTK kapsamında ise test verilerini daha iyi anlamak, niteliğini sorgulamak üzere açıklayıcı bir analizin gerçekleştiriliyor olmasıdır (Sijtsma ve Molenaar, 1984). Ayrıca bir yaklaşımın parametrik, diğerinin parametrik olmayan olarak nitelendirilmesinin bir başka nedeni ise, PMTK'da madde tepki fonksiyonlarının lojistik ya da normal ogive olarak belli bir parametrik forma göre oluşturulurken, POMTK'da madde tepki fonksiyonlarına yalnızca sıralama düzeyinde sınırlılık getirilmesidir. Örneğin POMTK modellerinden bazıları, iki kategorili madde puanı ile gizil değişken arasında pozitif ve monoton ilişki olduğunu ve bunun tüm yetenek düzeyleri için geçerli olduğunu varsaymaktadır (Sijtsma ve Meijer, 2007). Bu yaklaşıma ilişkin varsayımların incelenmesi ile PMTK ile aralarındaki farklılıkların ve ilişkilerin daha iyi belirlenebileceği düşüncesi ile varsayımların aşağıda açıklanması uygun bulunmuştur.

Parametrik Olmayan Madde Tepki Kuramı Varsayımları

POMTK modelleri PMTK modellerine göre daha zayıf varsayımlara sahip olsa da, POMTK kapsamında incelenecek verilerin de bazı varsayımları karşılaması beklenmektedir. Bu varsayımlar tek boyutluluk, yerel bağımsızlık, madde tepki fonksiyonlarının monotonluğu ve kesişmeyen madde tepki fonksiyonlarıdır.

Tek boyutluluk

Parametrik olmayan madde tepki kuramı yaklaşımını içinde var olan tek boyutluluk varsayımı genel olarak PMTK yaklaşımında incelenen tek boyutluluk varsayımından farklı değildir. Bu varsayım gereği, testte yer alan tüm maddeler aynı gizil değişkeni ölçmeye yönelik olarak hazırlanmış olmalıdır. Tek boyutluluğun MTK modellerindeki matematiksel

karşılığı ise, veri yapısını, birey yanıt örüntüsünü, açıklayacak tek bir gizil değişkenin bulunmasıdır. Bu varsayım özellikle eğitimde kullanılan testlerde ve psikometride sıklıkla kabul edilir çünkü tek boyutlu olarak gerçekleştirilen ölçme sonuçlarını yorumlamak daha kolaydır. Birden fazla psikolojik yapının aynı anda ölçülmesi hem bireysel puanların yorumlanmasını zorlaştırır hem de bireyler arasında kıyaslama yapmayı karmaşık hale getirir. Söz konusu varsayım her ne kadar ölçme sonuçlarını yorumlama kolaylığı sağlasa da, psikolojik yapıların çoğu tek bir özellik ile nitelendirilemeyecek kadar karmaşıktır. Bu durumda göz önünde bulundurularak tek boyutluluk varsayımı altında, psikolojik yapıların karmaşık nitelikleri göz ardı edilmeden, testlerin başat tek bir yeteneği-özelliği temsil etmesi beklenir (Baker ve Kim, 2004; Sijtsma ve Molenaar, 2002; Stout, 1990).

Parametrik olmayan madde tepki kuramı yaklaşımı kapsamında tek boyutluluk varsayımı maddeler arası ikili kovaryansın pozitif olmasını gerektirir, ancak bu durum varsayımın sağlanması için yeterli değildir. Madde çiftleri arasındaki tek bir negatif kovaryans dahi POMTK uyumunu zedelemekteyken, tüm maddeler arasındaki kovaryansların pozitif olması tek boyutluluğu garanti etmez. Bundan dolayı söz konusu varsayım çeşitli tekniklerle incelenmesi gerektiği alanyazında belirtilmektedir (Sijtsma ve Meijer, 2007; Straat, van der Ark ve Sijtsma, 2014)

POMTK kapsamında tek boyutluluk varsayımının incelenebileceği yöntemlerden biri, DETECT yöntemidir. Boyutluluğu ölçmek için maddeler arasındaki koşullu kovaryansları temel alan bu yöntem Stout'un (1987, 1990) çalışmalarına dayanmaktadır. Stout çalışmalarında temel yerel bağımsızlık, temel tek boyutluluk ve zayıf monotonluk kavramlarını tanımlamış ve temel tek boyutluluk için bir baskın gizil değişkenin olmasının yeterli olduğunu belirtmiştir. Tek boyutluluk varsayımının test edilmesi için de benzer madde kalan puanına sahip bireylerden gruplar oluşturarak, bu gruplar için hesaplanan maddeler arası ikili kovaryansın negatif olmaması gerektiğini belirtmiştir. DETECT yönteminde maddeler arası kovaryanslar bazı sınır değerler göre incelenmekte ve veri setinin boyutluluğu hakkında karar verilmektedir. Stout, Habing, Douglas, Kim ve Zhang (1996) tarafından, DETECT değerinin 0.1'in altında çıkması temel boyutluluğunun, 1'in üzerinde çıkması ise, çok boyutluluğunun göstergesi olarak kabul edilmiştir. Ancak Roussos ve Özbek (2006), DETECT değerinin 0.2'nin altında çıkmasını zayıf çok boyutluluk, 0.2 ile 0.4 arasında çıkmasını orta düzeyde çok boyutluluk, 0.4 ile 1 arasında

çıkmasını ise, yüksek çok boyutluluk göstergesi olarak belirtmiştir. Bonifay, Reise, Scheines ve Meijer (2014) ise, bu değerlerin verinin faktör yapısına bağlı olarak değişebileceğini, genel ve grup faktörleri arasındaki ilişkiden etkilendiğini belirtmiş ve bu değerlerin dikkatli kullanılmasını tavsiye etmişlerdir (akt. Meijer, Tenderio ve Wanders, 2015). DETECT yöntemi aynı isimli ücretsiz bir program ile uygulanmaktadır.

Parametrik olmayan madde tepki kuramı kapsamında, PMTK’da tek boyutluluğun analizinde uygulanan faktör analizine alternatif olarak önerilen bir başka teknik de Otomatik Madde Seçim İşlemi (OMMS-Automated Item Selection Procedure-AISP) olarak adlandırılan yöntemdir. Bu yöntem, maddeler arası kovaryansa dayalı bir algoritma ile çalışmakta ve maddelere ilişkin hesaplanan ölçeklenebilirlik katsayılarını (H_i) kullanmaktadır. Gözlenen kovaryansın olası maksimum kovaryansa olan oranından hesaplanan ölçeklenebilirlik katsayıları farklı araştırmacılar tarafından önerilen ve örneklem büyüklüğüne göre değişen farklı sınır değerlere göre yorumlanır ve bu değerler OMSİ sürecinde ölçüt olarak alınır (Kuijpers, van der Ark, Croon ve Sijtsma, 2016).

Temel olarak maddeler arasındaki ikili kovaryansa ve maddeler arasındaki ilişkinin gücüne dayanan OMSİ yönteminin amacı, madde setini veri setinin izin verdiği ölçüde, bir ya da daha fazla alt ölçeğe, Mokken ölçeğine ayırmaktır. Maddeler arasındaki kovaryans düzeyleri gözetilerek birbirleri ile ilişkili olan maddeler kümelendirilir. Bu kümelendirme yapılırken ise, öncelikli olarak H_{ij} değerleri 0’dan manidar olarak farklı olan madde çiftleri belirlenir. Bu işlemin ardından tüm olası madde çiftleri için en yüksek H_{ij} değerine sahip olan madde çifti belirlenir ve bu madde setiyle, ilk olarak pozitif korelasyona sahip, ikinci olarak ise H_i katsayısı sıfırdan büyük, son ölçüt olarak da H_i değeri araştırmacı tarafından belirtilen sınır değerden-c değeri- büyük olan bir madde seçilerek ölçeklemeye devam edilir. Bu süreç belirtilen koşulları sağlayan madde kalmayınca kadar devam eder. Sonuç olarak ise, bir ya da birden fazla aynı niteliği ölçen madde seti oluşturulmuş olur. Bu madde setlerine ilişkin yorumlar yapılırken maddelerin ölçtükleri içerikler ve veri yapısı hakkında bilenenler kullanılır (Meijer, Tenderio, ve Wanders, 2015; Straat, van der Ark ve Sijtsma, 2014).

OMSİ algoritması Mokken ölçekleme için kritik düzeyde öneme sahiptir ve Sijtsma ve Molenaar (2002), Mokken ölçeklemeyi “ortak bir özelliği, araştırmacı tarafından

belirlenen c değerinden (H_i katsayıları için temel alınan sınır değer) yüksek bir ayırt edicilik düzeyinde ölçen maddelerden oluşturulan ölçekleme yaklaşımı” olarak belirtmişlerdir.

Otomatik madde seçim işlemi algoritması aşağıdan yukarıya doğru işler ve bir kez seçilen bir maddenin geri dönüp seçilme, tekrar işleme alınma gibi bir durumu olamaz. Bundan dolayı bu yöntem, tüm madde gruplamalarını dikkate almaz. Straat, Van der Ark, ve Sijtsma (2009), OMSİ'nin bu dezavantajını ortadan kaldırmak için “generic search algorithm” (GA) algoritmasını önermişlerdir. OMSİ'nin tersine, GA yöntemi olasılıksaldır ve tüm olası alt bölümleri değerlendirmeye almaktadır. Bu nedenle OMSİ'den farklı sonuçlar üretebilmektedir ve araştırmacıların bu yöntemi de kullanarak elde ettikleri sonuçları değerlendirmeleri önerilmektedir. Bu çalışma kapsamında OMSİ ve GA algoritmaları birlikte kullanılmış ve elde edilen bulgular yorumlanmıştır.

Otomatik madde seçim işlemi ve GA algoritması ortak olarak araştırmacı tarafından belirlenen ve ölçeklenebilirlik katsayısının sınır değeri olan c değerlerine dayanmaktadır. Araştırmacı c değerini algoritmaya ne kadar yüksek olarak tanımlarsa, oluşturulan ölçeğin ayırt edicilik düzeyi de o kadar yüksek olur. Ölçeklenebilirlik katsayılarının sınırını belirten c değeri, H_i değerlerinin sınır değeridir ve ölçeğe alınacak maddelerin bireylerin doğru sıralanmasına yapacağı en az katkı düzeyini belirtir. Bu nedenle c değeri, bireylerin sıralanmasına düşük katkısı olacak maddelerin ölçeğe dahil edilmesinin önüne geçen bir teminat olarak görülebilir. Bu yöntem kullanılarak oluşturulan bir ölçekte tüm maddelerin ölçeklenebilirlik katsayılarının sıfırdan yüksek, c değerine eşit ya da bu değerden daha yüksek olması ve madde çiftleri için hesaplanan H_{ij} katsayılarının tümünün ise pozitif olması garantilenmiş olur. Dolayısıyla c değerinin yüksek seçilmesi, ölçeğin daha yüksek ayırt ediciliğe sahip maddelerden oluşması anlamına gelmektedir. Alan yazın incelendiğinde c değerine ilişkin farklı değerlerin denenmesi gerektiği belirtilmektedir. OMSİ analizinin ilk aşamasında, c değeri en alt sınır olan sıfır olarak seçilebilir ve bu durumda ölçeğe alınacak maddelerde aranan tek koşul, H_i katsayılarını pozitif olması olur. Bu durum ise ölçeğin tümü için hesaplanan H katsayısının çok düşük olmasına neden olabilir, çünkü H_i değerleri ile H katsayısı arasında şu şekilde bir bağlantı bulunmaktadır: $\min(H_i) \leq H \leq \max(H_i)$. Bu bağlantıya göre ölçeğe alınacak maddelerin H_i değerleri, belirlenecek c değeri ile araştırmacı tarafından kontrol edildiğinde, ölçeğin ölçeklenebilirlik katsayısı da kontrol edilmiş olur. Alan yazında c değerine ilişkin öneriler, H_i katsayısına

ilişkin önerilen değerlerle aynıdır. Buna göre c değerinin 0.3'ün altında bir değer olarak belirlenmesi, kuramsal olarak POMTK modellerine uyum sağlayan bir ölçekleme yapılsa dahi, pratikte bireylerin sıralanmasında bu ölçeğin düşük kesinlikte olması anlamına gelmektedir. Aynı zamanda H değerlerine ilişkin yöneltilen tüm eleştiriler c değeri için de geçerlidir ve bundan dolayı 0.3'ün yüksek bir değer olduğu düşünülebilir. Bu nedenle c değerleri için de H_i için uygulanan kurallar göz önünde bulundurulmalı ve çeşitli değerlere göre analizler tekrarlanarak veri setinin yapısı incelemelidir (Wismeijer, Sijtsma, Van Assen ve Vingerhoets, 2008).

Yerel Bağımsızlık

Bu varsayım PMTK modellerinde aranan yerel bağımsızlık varsayımıyla benzer niteliktedir. Testi alan bireyin testte yer alan i maddesine verdiği yanıtın aynı testte yer alan ve bireyin yanıtladığı diğer maddelerden etkilenmemesi anlamına gelir (Hambleton ve Swaminathan, 1985). Güçlü fakat ihlal edilmesi kolay bir varsayımdır. Örneğin testteki diğer maddelerle oluşan pratik etkisi ile gerçekleşen öğrenme ile yerel bağımsızlık ihlal edilebilir. Bu durum test devam ederken gerçekleşir ve testi alan birey testteki maddeleri çözerek gizil değişkendeki becerisini geliştirebilir. Bu durum yalnızca maksimum performans ölçülürken değil tipik performans ölçülürken de ortaya çıkabilir. Tutum ya da kişilik ölçeğini yanıtlayan bireyler, ölçekte yer alan maddelerle kendi davranışlarını kıyaslayıp değerlendirebilir ve ardından gelen maddelere farklı tepkiler verebilir. Bu durumlarda, test boyunca bireyin yeteneği değişiklik gösterir, fakat bu varsayım test boyunca birey yeteneğinin sabit kalmasını gerektirir (Baker ve Kim, 2004). Bu nedenle yerel bağımsızlık varsayımı sabit tutulan yetenek düzeyleri için maddeler arasındaki kovaryansların hesaplanması ile incelenir. Belli yetenek düzeyleri için maddeler arasındaki koşullu kovaryansın 0 olması zayıf yerel bağımsızlık varsayımının karşılandığını gösterir (Emberston ve Reise, 2000; Stout, 1990).

Chen ve Thissen (1997) maddelerin yerel bağımsızlığı ihlal ettiği durumlarda, PMTK yaklaşımında, madde ayırt edicilik parametrelerinin gerçek değerinden daha yüksek değerler aldığını göstermiştir. POMTK kapsamında yapılan bir simülasyon çalışmasında ise, yerel bağımsızlık varsayımını ihlal eden madde sayısının az olması durumunda ilişkili maddeler dışında kalan madde parametrelerinin etkilenmediği belirlenmiştir, ancak

buradaki az kavramı geniş bir kavram olarak kalmış ve yazarlar tarafından daha fazla çalışmaya ihtiyaç olduğu belirtilmiştir (akt. Meijer, Tendeiro ve Wanders, 2015).

Tek boyutluluk ve yerel bağımsızlık hem parametrik hem de parametrik olmayan madde tepki kuramı modellerinin ortak varsayımdır. POMTK modellerini PMTK modellerinden ayıran nitelik madde tepki fonksiyonlarının daha esnek kestirilmesi olsa da, parametrik olmayan madde tepki kuramı modelleri madde karakteristik eğrisine ilişkin varsayımlardan tamamen arınık değildir. Aksi durumda, koşullu bağımsızlık ve tek boyutluluğunun sağlandığı tüm durumlarda POMTK modellerinin veriye uyum sağlanması beklenir ve bu durum Suppes ve Zonotti (1981)'nin çalışmasında kanıtlanmıştır. Dolayısıyla POMTK ile PMTK arasındaki temel fark; POMTK'nın madde tepki fonksiyonlarına ilişkin herhangi bir varsayımı olmaması değil, PMTK kadar sınırlayıcı varsayımlarının olmamasıdır. POMTK'da yer alan madde karakteristik eğrisinin şekline ilişkin varsayımlar (düzgünleştirme gibi) PMTK'daki varsayımlar kadar sınırlayıcı değildir. Bu durum da verinin asıl yapısının korunmasını ve ölçülen psikolojik değişkenin daha iyi anlaşılmasını sağlamaktadır (Molenaar, 1997; Sijtsma ve Molenaar, 2002). POMTK modellerinin madde tepki fonksiyonlarına ilişkin varsayımlarından ilki madde tepki fonksiyonlarının monotonluğudur.

Madde Tepki Fonksiyonlarının Monotonluğu

Bu varsayım bireylerin maddeyi doğru yanıtlama olasılıklarının ölçülen gizil yeteneğin monoton azalmayan fonksiyonu olduğu belirtir. Yani bireyin ölçülen özelliğe sahip oluş derecesi arttıkça maddeye doğru cevap verme olasılığının aynı kalması ve/veya artması anlamına gelir. Eğitsel ve psikolojik testler gözetildiğinde, bu özellik şu şekilde yorumlanabilir: yeteneği, tutumu ya da becerisi daha üst düzeyde olan bireylerin söz konusu ölçülen özelliğe daha düşük düzeyde sahip olan bireylere göre maddelere doğru yanıt verme olasılıkları daha yüksek olmasıdır (Meijer, Sijtsma ve Smidt 1990; Sijtsma ve Molenaar, 2002).

Monoton azalmayan ifadesi madde tepki fonksiyonunu tanımlayan önemli bir ifadedir çünkü bazı durumlarda madde tepki fonksiyonu monoton artış göstermez. Bazı madde tepki fonksiyonları çan eğrisi şeklinde olabileceği gibi bazıları ise adimsal bir şekilde olabilir. POMTK modelleri ile bu gibi değişken formdaki maddeleri ve bu

maddelerle ölçülen psikolojik özellikler arasındaki ilişkiyi incelemek olanaklı hale gelir. Ayrıca özellikle madde ayırt edicilik değerlerinin çok yüksek ve madde güçlük parametrelerinin ise uçlarda olduğu ölçmelerde, söz konusu maddelere ilişkin kestirimler PMTK modelleri ile gerçekleştirilemeyebilir. Söz konusu maddeler için PMTK modelleri ile madde model uyumu sağlanamayabilir ve maddenin test dışında tutulmasına karar verilebilir. Bu gibi durumlarda, maddeyi test dışında tutmak yerine, madde tepki fonksiyonlarının şeklinin katı bir şekilde sınırlandırılmadığı POMTK modelleri ile söz konusu maddelere ilişkin MTF'ler kestirilebilir ve böylece ölçülen gizil değişkenle maddeler arasındaki ilişki daha gerçekçi ve ayrıntılı olarak incelenebilir (Junker, 2000).

Sijtsma ve Meijer (2007), POMTK kapsamında yer alan Monoton Homojenlik Modeli (MHM) ile kestirilen MTF'ler ile bu modelin PMTK'daki karşılığı olan 2 PLM ile kestirilen MTF'ler arasındaki farkları incelemiş ve POMTK ile kestirilen MTF'lere ilişkin şu sonuçlara ulaşmıştır:

- MTF'ler S şeklinde olmak zorunda değildir. Hatta düzgün bile olmayabilir.
- Birden çok büküm noktasına sahip olabilir ve birkaç yetenek aralığı boyunca sabit kalabilir.
- En düşük asimptotu 0'ın üzerinde, en yüksek asimptotu ise 1'in üzerinde olabilir.

POMTK ile kestirilen MTF'lerin belirtilen özellikleri incelendiğinde, PMTK ile POMTK'yı ayıran özellikler de ortaya çıkmaktadır. POMTK ile kestirilen MTF'ler, PMTK ile kestirilenlere göre oldukça esnektir ve madde puanı ile gizil değişken arasındaki ilişkiyi daha iyi ortaya çıkarmakta ve bu yaklaşıma açılmalı bir nitelik kazandırmaktadır. Bu varsayım R programında yer alan “*mokken*” paketinin bir fonksiyonu olan basit bir istatistiksel manidarlık testi ile sınanabilir. Söz konusu test ile bu varsayım, benzer toplam puana sahip olan bireylerin maddeye verdikleri yanıtlar gözetilerek madde kalan puanı gruplarının oluşturulması ve her grup için doğru yanıtlama oranlarının hesaplanması ile incelenir. Bu yöntemin ayrıntılı olarak açıklaması verilerin analizi kısmında verilmiştir. Ayrıca bu varsayım, veriye uygun parametrik olmayan regresyon yöntemi ile maddelere ilişkin MTF'lerin oluşturulmasıyla da incelenebilir (Meijer, Tenderio ve Wanders, 2015).

Tek boyutluluk, yerel bağımsızlık ve monotonluk varsayımı POMTK kapsamında ölçekleme yapmak için yeterlidir ancak MTF'lere getirilen ek bir sınırlama ile birey yeteneklerini eşit aralıklı ölçek düzeyinde kestirmek olanaklı hale gelmektedir. Bu ek

sınırlama ise, kesişmeyen madde tepki fonksiyonları ya da değişmez madde sıralaması adıyla bilinen varsayımdır.

Kesişmeyen Madde Tepki Fonksiyonları

Birçok test uygulaması yalnızca bireylerin yetenek ya da başarıları bakımından sıralanmasını gerektirmektedir. İşe eleman alımı, üst düzey eğitimler için en başarılı 10 öğrencinin belirlenmesi gibi durumlar bireylerin yetenek ya da başarıları bakımından sıralanmalarının gerekli olduğu durumlara örnektir. Bireyleri sıralamanın yanında maddeleri de güçlük düzeylerine göre sıralamak birçok durum için kullanışlıdır. Maddeler testi alan her birey için aynı güçlük sıralamasına sahip ise, elde edilen sonuçların yorumlanması daha kolay olacaktır ve bu özellik “değişmez madde sıralaması” olarak adlandırılır. Özellikle zekâ testlerinde maddeler bireylere artan güçlük sıralamasına göre uygulanır ve belli yaş grupları için bazı maddeler cevaplanmadan geçilebilir. Böylece bireyler kendi düzeylerine göre çok kolay olan maddelerle karşılaşmaz ve motivasyon düşüklüğü yaşamaz. Madde sıralamasının önemli olduğu bir başka durum ise, test uygulamalarında Değişen Madde Fonksiyonunun (DMF) ortaya çıktığı durumlardır. Özellikle tek biçimli DMF belirlendiğinde, maddelerin farklı gruplar için güçlük sıralamasının aynı olmadığı ve farklı gruplara göre farklı fonksiyon gösterdiği sonucuna ulaşılır. Değişmez madde sıralaması varsayımının ölçme uygulamalarındaki önemini belirten başka örnek ise maddelerin gelişimsel kuramlara göre oluşturulduğu ölçme uygulamalarıdır. Çocukların nesnelere ilişkin özellikleri kavramları bu duruma örnek olarak verilebilir. Çocuk içinde bulunduğu gelişim dönemi itibari ile yalnızca belli özellikleri belli bir düzeyde kazanmış olabilir. Örneğin, sadece nesnelere boyları arasındaki farklar büyükse, bu farkları belirleyebiliyor ya da sadece renklere göre nesnelere gruplandırabiliyor, boy farklılıklarını aynı anda algılayamıyor olabilir. Bu gibi durumlarda maddelerin gelişimsel kurama uygun bir şekilde güçlük düzeylerine göre sıralanması ve bu sıralamanın tüm yetenek düzeylerinde korunması gerekmektedir (Sijtsma ve Junker, 1996).

Psikolojik ölçmelerde grup düzeyinde maddelerin ortalama puanlarına göre yapılan sıralamanın genel olarak bireysel düzeyde de aynı olduğu düşünülür. Ancak bu varsayım sadece madde sıralamasının gizil değişkenin farklı düzeylerinde aynı olduğu durumlarda ortaya çıkar ve bir diğer ifadeyle değişmez madde sıralamasının karşılandığı durumlarda

gerçekleşir ve bu varsayımın veri seti tarafından karşılanma durumu ölçeklemenin yapıldığı yaklaşıma göre incelenmelidir (Watson, Deary & Shipley, 2008). Söz konusu varsayımının veri setleri için karşılanmasının ön koşulu, tüm maddelere ilişkin kestirilen madde tepki fonksiyonlarının Eşitlik 5'te verildiği gibi sıralanmasıdır (Sijtsma ve Meijer, 2007; Sijtsma ve Junker, 2006).

$$P_1(\theta) \leq P_2(\theta) \leq \dots \leq P_K(\theta) \quad (5)$$

Eşitlik 5 incelendiğinde, güçlük düzeyi en yüksek olan madde 1, güçlük düzeyi ikinci olarak en yüksek olan madde 2 tarafından takip edilir ve tüm maddeler için bu ilişkiler bu şekilde devam eder. Bu varsayımın karşılanması için maddeler arasındaki bu ilişkinin belli yetenek düzeyi aralıkları (θ) için sağlanması gerekir. Burada vurgulanması gereken nokta, varsayımla maddeyi doğru yanıtlama olasılıklarının değil madde sıralamasının aynı olması gerekliliğidir. Olasılıklar alt gruplara göre değişiklik gösterebilir çünkü olasılıklar yetenek dağılımından etkilenir. Örneğin bir maddeyi doğru yanıtlama olasılığı erkekler için 0.15; kızlar için 0.27 olabilir ve bu maddeler için değişmez madde sıralaması varsayımı sağlanıyor ise, her iki grup için de söz konusu maddenin diğer maddelere göre sıralamadaki yerinin aynı olması beklenir. Bu özellik daha önce de belirtildiği gibi, teorilerin test edilmesi, teste ilişkin başlama ve bitirme kurallarının belirlenmesi, değişen madde fonksiyonlarını ortaya çıkarılması ve özellikle sağlık alanlarında kullanılan tanı ve teşhis amaçlı testlerle yapılan ölçmeler gibi birçok durum için oldukça önemlidir (Sijtsma ve Molenaar, 2002).

Ölçme uygulamalarında önemli yeri olan bu varsayımın, iki kategorili verilerde incelenmesi birden çok yöntem ile gerçekleştirilebilir. Meijer ve Egberink (2012), olası ihlallere ilişkin istatistiksel testlere geçilmeden önce, madde kalan puanı regresyonlarının oluşturulmasını ve incelenmesini önermişlerdir. Daha ayrıntılı inceleme için ise, kalan puan yöntemi ve P-matrisi yöntemlerinin kullanılmasının gerekli olduğunu belirtmişlerdir. Ayrıca MTF'lerin birbirlerine olan uzaklığının ölçüsünü veren H^T katsayısının hesaplanması ihlallerin manidarlığı hakkında bilgi vermekte, böylece varsayımın veri setinde ne düzeyde karşılandığını göstermektedir. Bu çalışma kapsamına alınan veri setleri için yapılan değişmez madde sıralaması varsayımı incelemesinde, hem madde kalan puanı

gruplarına göre hem de P matrisi yöntemine göre analizler yapılmış ve elde edilen sonuçlarla hesaplanan H^T katsayıları karşılaştırılarak, söz konusu varsayımın veri setleri için karşılanma durumları belirlenmiştir. Bu yöntemlerin detayları verilerin analizi kısmında yer almaktadır.

Tek boyutluluk, yerel bağımsızlık, monoton azalmayan MTF'ler varsayımlarını karşılayan veri setleri POMTK kapsamında yer alan esnek bir model olan Monoton Homojenlik Modeli'ne uyum sağlarken, bu varsayımlara ek olarak, değişmez madde sıralaması varsayımını karşılayan veri setleri ise daha kısıtlayıcı bir model olan Çift Monotonluk Modeline uyum sağlar. Söz konusu modeller araştırma kapsamında kullanıldığı için çalışmanın bu kısmında bu modellerin tanıtılması uygun görülmüştür.

Parametrik Olmayan Madde Tepki Kuramı Modelleri

Mokken, Guttman ölçeğinin uzantısı olarak POMTK kapsamında yer alan iki model üretmiştir. Bu modeller; Monoton Homojenlik Modeli (Monotone Homogeneity Model-MHM)ve Çift Monotonluk Modeli (Double Monotonicity Model –ÇMM)'dir. Mokken tarafından yalnızca iki kategorili veriler için önerilen bu modeller, Molenaar (1982, 1997) tarafından çok kategorili veriler için genişletilmiş ve bu sayede sıralanmış ve kategorik puanlar için de POMTK modellerinin kullanılması sağlanmıştır. Çalışmanın bu kısmında, bu modeller ayrıntılı olarak tanıtılmıştır.

Monoton Homojenlik Modeli

Hem çok kategorili hem de iki kategorili puanlanan maddeler için uygulanan Monoton Homojenlik Modelinde her iki veri seti için de aynı varsayımlar incelemeye alınır. Bu varsayımlar tek boyutluluk, madde tepki fonksiyonlarının monotonluğu ve yerel bağımsızlıktır. Çok kategorili puanlanan maddeler için monotonluk varsayımı madde adım fonksiyonlarına uygulanır (Molenaar, 2001; Van Schuur, 2011).

Monoton homojenlik modelinde yer alan varsayımların veri seti için karşılandığı durumda, bireylerin toplam puanları gözetilerek yetenek düzeylerine göre stokastik (raslantısal) olarak sıralanabileceği Grayson (1988) tarafından belirtilmiştir (akt. Straat, van der Ark ve Sijtsma, 2014). Bu varsayım matematiksel olarak şu şekilde ifade edilir:

$$P(X_+ \geq X_+ | \theta = \theta_a) \leq P(X_+ \geq X_+ | \theta = \theta_b), \theta_a < \theta_b \quad (6)$$

Eşitlik 6'da X_+ toplam puanı, θ ise yetenek düzeyini ifade etmektedir. Eşitlikten de görüldüğü üzere, monoton homojenlik modeline uyum sağlayan veri setlerinde toplam puan sıralaması olasılıklı olarak bireylerin yetenek sıralamasını verir (Mokken, 1971).

Model uyumunun belirlenmesi için kurama ilişkin varsayımlar incelendikten sonra, POMTK analizlerinde bir diğer koşul olan madde çiftleri arasındaki kovaryansların incelenmesi gerekmektedir. Model uyumunun yeterli düzeyde olması için testte yer alan maddelerin tümünün arasındaki kovaryans değerleri pozitif olmalıdır. Madde çiftleri arasındaki negatif olmayan kovaryanslar model uyumu için gereklidir fakat madde tepki fonksiyonlarının monoton azalmayan nitelikte olduğunu garantilemez. Bu nedenle bu varsayımın da karşılanıp karşılanmadığı belirlenmelidir (Sijtsma ve Molenaar, 2002). Mokken(1971), MHM modeline ait varsayımların karşılanmasının bireylerin yetenek ölçeğinde sıralanması için gerekli olduğunu belirtmektedir.

Çift Monotonluk Modeli

Çift Monotonluk Modeli (ÇMM), Mokken (1971) tarafından önerilen Monoton Homojenlik Modelinden daha kısıtlayıcı bir modeldir. Monoton homojenlik modelinde yer alan varsayımlara ek olarak, kesişmeyen madde tepki fonksiyonları varsayımının bu model için karşılanması gerekmektedir. Bu varsayım altında maddeler de yetenek düzeyleri gibi güçlüklerine göre sıralanabilir ve bu sayede MHM'den farklı olarak yalnızca maddeler değil, bireyler de aynı yetenek sürekliliğinde sıralanabilir. Bu özellik PMTK modelleri içinde yer alan Rasch modelde sağlanan bir özelliktir (Sijtsma ve Molenaar, 2002).

Çift Monotonluk Modeli, MHM'de karşılanması gereken tüm varsayımları taşıdığı için bu modelin özel hali olarak nitelendirilir ve ÇMM'ye uyum sağlayan tüm veri setleri aynı zamanda MHM'ye de uyum gösterir. Fakat bu durumun tersi geçerli değildir. Dolayısıyla MHM, ÇMM'nin daha zayıf varsayımlara sahip özel bir hali olarak nitelendirilir (Molenaar, 1997).

POMTK kapsamındaki modeller sadece ölçek düzeyinde değil madde düzeyinde de analizler yapar. Her iki model için de geçerli olarak ölçme yapan maddelerin belirlenmesi için Mokken (1971, 1997), ilk olarak Loevinger'in (1947, 1948) kullandığı H

ölçeklenebilirlik katsayısının kullanılmasını önermiştir. H katsayısı Guttman ölçeğindeki hata kavramıyla ilişkilidir ve bireyin zor maddeye doğru yanıt verirken kolay maddeye yanlış yanıt verme düzeyine ilişkin bilgi verir. Bir diğer ifadeyle veri setinde yer alan uygun olmayan/beklenmeyen yanıt örüntülerini hakkında bilgi veren bir ölçüttür. Çok kategorili maddelerde ise, maddeye ilişkin madde güçlük adımlarıyla ilişkilendirilir ve hatalar madde adım çiftleri gözetilerek hesaplanır. H katsayısı bireyleri ölçülen özelliğe göre sıralamada maddenin ne kadar uyumlu olduğunu gösterir ve POMTK kapsamında “ölçeklenebilirlik katsayısı” olarak da adlandırılır (Sijtsma ve Molenaar, 2002). Ölçeklenebilirlik katsayısının üç türü vardır: maddeler için H katsayısı (H_i), madde çiftleri için H katsayısı (H_{ij}) ve ölçek için H katsayısı (H) olmak üzere üç tür ölçeklenebilirlik katsayısı hesaplanır. Teorik olarak H değerinin 0 ile 1 aralığında değer alması beklenir ve pozitif olan tüm H değerleri kabul edilebilir düzeydedir. H değerinin 1 olması ise Guttman hatasının 0'a yaklaştığını belirtir fakat uygulamada pek gerçekleşmeyen bir durumdur. Pratikte 0.0-0.3 aralığında H_i değerine sahip olan maddeler pozitif fakat düşük ayırt ediciliğe sahip maddeler olarak kabul edilir ve bireyleri toplam puana dayanarak güvenilir olarak sıralamaya katkısı düşük olduğundan test için kullanışsız olarak nitelendirilir (Mokken ve Lewis, 1982). Madde düzeyi için hesaplanan H_i ile ölçeğin tümü için hesaplanan H arasında şu şekilde bir ilişki bulunmaktadır: $H \geq \min (H_i)$. Bu eşitliğe göre ölçeğin tümü için hesaplanan H katsayısı en düşük H_i katsayısından yüksek olacaktır. Mokken (1971), ölçeğin tümü için hesaplanan H katsayısının $0.3 \leq H < 0.4$ olduğu durumlarda ölçeğin “zayıf”, $0.4 \leq H < 0.5$ olduğu durumlarda “orta” ve $H \geq 0.5$ olduğu durumlarda ise ölçeğin “güçlü” olarak nitelenebileceğini belirtmiştir. Son zamanlarda yapılan bir çalışma da ise, Mokken tarafından önerilen bu ölçek analizinin yalnızca POMTK modelleri için değil, PMTK modelleri için de analize başlamadan önce yapılmasının oldukça faydalı olduğunu ve bu incelemenin model veri uyumunu arttıracığı öne sürülmüştür (Christensen ve Kreiner, 2010). Mokken tarafından önerilen bu ölçek analizinin özellikle küçük örneklerde etkili bir ölçek geliştirme stratejisi sunduğu da alan yazında yer almaktadır (Junker, 2000).

Mokken ile başlayan parametrik olmayan ölçekleme yöntemini, Madde Tepki Kuramı çatısı altında inceleyen öncü isimlerden Sijtsma, Meijer ve Van der Ark (2011), çalışmalarında POMTK kapsamında yapılacak incelemelerin Çift Monotonluk Modeli

temel alınarak yapılmasını önermektedirler. Bu duruma neden olarak da Çift Monotonluk Modelinin Monoton Homojenlik Modelini kapsadığı ve Çift Monotonluk Modeline uyumlu olan veri setinin Monoton Homojenlik Modeline zaten uyumlu olmak zorunda olduğunu belirtmektedirler. İkili puanlanan maddeler için yapılacak POMTK analizlerinde izlenmesi gereken adımlar araştırmacılar tarafından şu şekilde belirtilmektedir:

1. Veri setinde yer alan tüm maddelerin ortalama puanları (Item popularities), bir diğer ifadeyle Klasik Test Kuramındaki madde güçlük parametreleri hesaplanır.
2. Veri setinde yer alan maddelere, madde çiftlerine ve tüm ölçeğe ilişkin ölçeklenebilirlik katsayıları, sırasıyla, H_i , H_{ij} ve H katsayıları hesaplanır. Söz konusu katsayıların sınır değerlerden yüksek olup olmadığı incelenir.
3. Veri setinin tek boyutluluğu Otomatik madde seçim işlemi (OMSİ- Automated Item Selection Procedure- AISP) kullanılarak belirlenir.
4. Maddelerin monotonluk varsayımına olan uyumu madde kalan puanı regresyonu hesaplanarak incelenir.
5. Değişmez madde sıralaması (Invariant Item Ordering), bir diğer ifadeyle kesişmeyen madde tepki fonksiyonları varsayımını incelemesi yapılır.
6. Son olarak madde güçlük sıralamasının doğruluğu H^T katsayıları kullanılarak incelenir.

Belirtilen aşamalardan geçilerek veri setine uygun modelin bulunmasının ardından, söz konusu modele göre ölçeklenen maddelere ilişkin MTF'ler kestirilip incelenebilir. Bu sayede özellikle tipik performansın ölçüldüğü durumlarda ölçülen gizil değişkene ilişkin önceden kestirilemeyen farklı özellikler ortaya çıkarılabilir (Meijer, Tenderio ve Wanders, 2015). Madde tepki fonksiyonlarının POMTK kapsamında yer alan modellere göre kestirilmesi sırasında parametrik olmayan MTF kestirim yöntemleri kullanılmaktadır.

Mokken ölçekleme tekniği alan yazında POMTK'dan daha önce tanıtılmış olsa da parametrik olmayan MTF kestirimi yöntemleri Mokken ölçekleme tekniklerinin öncesine dayanmaktadır. İlk parametrik olmayan MTF, Lord tarafından 1970'de geliştirilmiştir, bundan dolayı parametrik olmayan madde tepki fonksiyonu kestiriminin Mokken'ın parametrik olmayan madde tepki kuramı modelleri içinde geliştirildiği ileri sürülemez, fakat parametrik olmayan MTF kestirimleri, Mokken tarafından önerilen monoton homojenlik modeli varsayımları ile aynı varsayımları gerektirmektedir. Bu nedenle

monoton homojenlik modeline ilişkin varsayımların karşılandığı durumlarda parametrik olmayan MTF kestirim yöntemleri uygulanır (Junker, 2000; Sijtsma ve Molenaar, 2002; Van Schuur, 2011). Parametrik olmayan MTF kestirimi ve Kernel Düzgünleştirilmesi (Kernel Smoothing) aşağıda genel olarak tanıtılmıştır.

Parametrik Olmayan Madde Tepki Fonksiyonu Kestirimleri

Parametrik olmayan madde tepki kuramı kapsamında özellikle parametrik olmayan regresyon teknikleri sıklıkla tercih edilmektedir. Parametrik olmayan madde tepki fonksiyonu kestirim yöntemleri tamamen parametrik olmayan ve yarı parametrik olan yaklaşımlar olarak iki temel kategoriye ayrılmaktadır. Tamamen parametrik olmayan yaklaşımlar her yetenek düzeyinde doğru yanıt verme olasılığını kestirir. Kestirilen madde tepki fonksiyonları parametrik bir forma sahip olmaz ve bu fonksiyon oluşturulurken, tüm madde yanıt örüntüleri kullanılır. Diğer yandan yarı parametrik yaklaşımlar, madde tepki fonksiyonlarını temel fonksiyonların doğrusal kombinasyonu olarak kestirir. Kestirilen fonksiyonlar birkaç parametre ile tanımlanan sürekli fonksiyonel bir formda olur. Yakın tarihli çalışmalarda, yarı parametrik yaklaşımlar geleneksel parametrik yaklaşımlardan daha esnek oldukları için parametrik olmayan yaklaşım olarak nitelendirilmektedir (Douglas ve Cohen, 2001; Lee, 2007; Meijer, 2004).

POMTK modellerine uyum sağlayan maddelerin madde ve kategori karakteristik eğrilerini kestirmek için Kernel düzgünleştirilmesi, B-Splines ve Monoton Splines (Ramsay ve Abrahamowicz, 1989) gibi farklı yöntemler bulunmaktadır. Ramsay (1991)'in Kernel Düzgünleştirilmesi yaklaşımı tamamen parametrik olmayan tekniklerin içinde yer almakta ve diğer yöntemlere göre daha çok tercih edilmektedir. Ramsay (1991, 1997, 2000) tarafından geliştirilen bu yaklaşım hem iki kategorili hem de çok kategorili maddelerin analizinde kullanılır. Esnek bir modelleme yaklaşımı olarak kabul edilen bu yaklaşım çoğu durumda gerçek verilere parametrik modellerden daha iyi uyum sağlar. Ayrıca özellikle asimetric madde tepki fonksiyonlarında bu yaklaşımın kullanışlı olduğu ve gizil yetenekteki değişimlere duyarlı olduğu belirtilmiştir (Lee, 2007; Santor ve Ramsay, 1998).

Bu yaklaşımda madde tepki fonksiyonları PMTK modellerinde olduğu gibi önceden belirlenen ya da modelde belirtilen şekilde kestirilmez. Bunun yerine var olan veri yapısı temele alınarak kestirimler yapılır. Bu nedenle söz konusu yaklaşım PMTK modellerinde

olduđu gibi dođrulamayı nitelikte deđil, aımlayıcı nitelikte kestirim yapar. Bu yntemde madde tepki fonksiyonları eřit olarak ađırlıklandırılmıř ve seilmiř deđerlendirme noktalarına gre kestirilir ve yerel ortalamalar tekniđi kullanılır. Yerel ortalamalar tekniđi, gzlenen verilerin belirlenmiř noktalarına ađırlıklandırma yaparak, verilerin kestirim srecine dâhil edilmesini sađlar (Santor ve Ramsay, 1998).

Kernel dzgnleřtirmesi diđer parametrik olmayan madde tepki fonksiyonu kestirim yntemlerine gre bir takım avantajlara sahiptir. Bu yntemin ilk avantajı hesaplama kolaylıđıdır. Madde tepki fonksiyonu gzlenen madde yanıtlarının ađırlıklandırılmıř ortalamaları temel alınarak hesaplandıđı iin, iteratif (tekrarlamalı) bir kestirim srecine ihtiya duyulmaz ve daha kısa srede kestirimler yapılabilir. Bu yntemin bir diđer avantajı da hem ok kategorili hem de iki kategorili verilere uygulanabilir olmasıdır. Bu yntemin nc ve en byk avantajı ise sıralı yetenek kestirimi ile Kernel dzgnleřtirmesinden elde edilen madde tepki fonksiyonlarının tutarlılıđıdır. Carroll, Maca ve Ruppert (1999) ise geleneksel parametrik olmayan regresyon yntemlerinin tutarsız kestirimler yaptıđını, Fan ve Troung (1993) ise Kernel dzgnleřtirmesinin farklı lme hatası dađılımlarında asimptotik sonular verdiđini gstermiřtir (akt. Ramsay, 2000). Douglas (1997) alıřmasında, madde ve birey sayısı yeterince ok olduđunda, Kernel dzgnleřtirmesinin PMTK modellerine gre daha dođru kestirimler yaptıđı sonucuna ulařmıř ve PMTK varsayımları karřılanmadıđında bu yntemin kullanılabileceđini belirtmiřtir. POMTK modelleri iinde PMTK modelleri ile dođrudan kıyaslanabilir nitelikler tařıyan bu yaklařım Ramsay (2000) tarafından geliřtirilen TESTGRAF programında incelenebilmekte, ayrıca R programında da “KernSmoothIRT “adıyla yer alan bir paket ile analizler yapılabilir.

Parametrik Olmayan Madde Tepki Kuramı Yaklařımında Gvenirlik İncelemeleri

Parametrik MTK yaklařımında madde yanıt rntleri kullanılarak bireylere iliřkin gizil yetenek kestirilir ve aynı testin hipotetik olarak tekrarlı uygulamalarından elde edilen varyans kestirilen yetenek puanına iliřkin varyans olarak kabul edilir. Elde edilen bu varyans deđerinin kk olması kestirimin netliđini gsterirken, byk olması ise kestirilen yetenek puanının gvenilir olmadıđı anlamına gelir, bu durumun nedeni ise bu varyansın bireysel dzeyde hesaplanıyor olmasıdır. Genellikle her yetenek dzeyine iliřkin birey ii

varyansı veren test bilgi fonksiyonları PMTK kapsamında güvenilirlik incelemelerinde kullanılır.

Parametrik olmayan madde tepki kuramı yaklaşımında PMTK'daki gibi bir yetenek kestirimi olmadığı için güvenilirlik Çift Monotonluk Modelinde karşılanan kesişmeyen madde tepki fonksiyonları varsayımı ile ilişkilendirilmiştir. Söz konusu varsayımın incelenmesi için oluşturulan ve simetrik bir nitelik taşıyan P matrisinde, kolaydan zora doğru sıralanmış maddeler yer almakta, hücrelerde ise madde çiftlerini doğru yanıtlayan bireylerin oranları bulunmaktadır. Eğer veri seti tarafından ÇMM'nin nitelikleri karşılanıyorsa, hücrelerde yer alan oranların soldan sağa ve yukarıdan aşağıya doğru artış göstermesi beklenmektedir. Oluşturulan matrisin köşegeninde yer alan hücreler boş bırakılmıştır ancak bu hücreler ilk ve son hücre değerleri ile ara değerlendirme (interpolation) yapılarak kestirilir ve üretilen bu değerler de maddenin bireye ikinci defa uygulanma durumunda pozitif yanıt verme olasılığının kestirimini vermektedir. Bu olasılık Mokken (1971) tarafından POMTK uygulamalarında test-tekrar test güvenilirliği olarak belirtilmekte ve bu değerlerin Cronbach alfa ile karşılaştırmalı olarak kullanılması önerilmektedir. Mokken tarafından önerilen bu istatistik “ p_{ii} ” olarak gösterilmekte ve istatistiğin hesaplanmasına ilişkin formül Şekil 2 'de verilmektedir (Van Schuur, 2003.)

$$\rho = \frac{\sum_{i=1}^k [p_{ii} - p_i^2] + 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^k [p_{ij} - p_i \cdot p_j]}{\sum_{i=1}^k [p_i(1 - p_i)] + 2 \sum_{i=1}^{k-1} \sum_{j=i+1}^k [p_{ij} - p_i \cdot p_j]}$$

Şekil 1. Mokken Tarafından Önerilerin Güvenirlik Kestirim Formülü

Mokken (1971)'in ardından, Molenaar ve Sijtsma (1984) POMTK kapsamında güvenilirliğin incelendiği ilk çalışmalardan birini yapmış ve POMTK kapsamında iç tutarlılık kavramını Loevinger tarafından önerilen H katsayısı ve Cronbach alfa katsayısı ile karşılaştırarak incelemişlerdir. Her iki katsayı için yapılan birçok inceleme sonucunda, bu katsayılar arasında yüksek pozitif korelasyon olduğunu belirlemişlerdir. Aynı zamanda test uzunluğunun 10 madde ve üstü olduğu durumlarda H katsayısının neredeyse her zaman Cronbach alfa katsayısından daha düşük olduğunu belirlemişlerdir.

Sijtsma ve Molenaar (1987) tarafından önerilen POMTK kapsamında güvenilirlik incelemeleri için kullanılması önerilen katsayı ise MS (Molenaar-Sijtsma) katsayısıdır. Bu

katsayı madde yanıt matrisinde yer alan gözlenemeyen oranlara dayalı olarak hesaplanmaktadır. Araştırmacılar çalışmalarında Mokken tarafından önerilen yöntemle kendilerinin önerdikleri MS katsayısını simülatif veriler kullanarak karşılaştırmışlar ve klasik güvenilirlik kestirimleri olan Cronbach Alfa ve Guttman tarafından önerilen Lambda 2 katsayılarından daha az yanlı kestirimler yaptığını belirlemişlerdir. Buna ek olarak, klasik güvenilirlik kestirimlerinin daha az örnekleme hatasına sahip olduğu çalışma kapsamında belirlenmiş ancak Sijtsma ve Molenaar (1987), söz konusu farkların küçük ve pratikte manidar olma olasılığının düşük olduğunu belirtmişlerdir.

Bu çalışma kapsamında oluşturulan veri setlerinin POMTK'ya dayalı olarak güvenilirliklerinin incelenmesinde klasik güvenilirlik indeksleri olan Cronbach Alfa ve Lambda 2'ye ek olarak, Sijtsma ve Molenaar (1987) tarafından önerilen MS katsayısı ile Van der Ark, Van der Palm ve Sijtsma (2011) tarafından gizil sınıf analizinde kullanılmak üzere geliştirilmiş fakat POMTK yaklaşımında da kullanılabilceği belirtilen LCRC (Latent Class Reliability Coefficient) hesaplanmıştır.

Cronbach Alfa katsayısı da POMTK kapsamında incelenen veri setleri için hesaplanabilecek bir diğer güvenilirlik belirleme katsayısıdır. Cronbach Alfa katsayısı, Kuder Richardson katsayısının genel hali olarak tanımlanmakta ve testin olası tüm iki yarılarından kestirilen iki yarı korelasyonlarının ortalaması olarak hesaplanmaktadır (Cronbach, 1951). Guttman (1945) bu katsayının güvenilirliğin sınır değeri olarak kullanılabilceğini belirtmekte ancak testlerin tek boyutluluk varsayımını karşıladığı durumlarda gerçek güvenilirliğe eş değerler üretilebileceği belirtilmiştir (Feldt ve Qualls, 1996). Söz konusu katsayının POMTK uygulamaları için de güvenilirliğin alt sınır değerine ilişkin bilgi verdiği, dolayısıyla kullanılabilceği belirtilmektedir (Molenaar ve Sijtsma, 1984). Bu nedenle oluşturulan veri setleri için Cronbach Alfa katsayıları da hesaplanmış ve yorumlanmıştır.

BÖLÜM 3

YÖNTEM

Bu bölümde araştırmanın modeli, evren ve örneklem, veriler ve elde edilmesi ile verilerin analizinde kullanılan yöntemler tanıtılmıştır.

Araştırmanın Modeli

Bu çalışma kapsamında çeşitli faktörlere göre düzenlenen verilerle parametrik ve parametrik olmayan madde tepki kuramı modelleri kullanılarak elde edilen madde ve yetenek parametrelerinin kıyaslanması amaçlanmaktadır. Örneklem büyüklüğü, test uzunluğu ve maddelerin psikometrik niteliklerine göre düzenlenen veri setlerinden PMTK ve POMTK'ya uygun olan modeller belirlenerek madde ve yetenek parametreleri kestirilmiş ve elde edilen sonuçlar arasındaki korelasyonlar incelenmiştir. Araştırma farklı MTK yaklaşım ve modellerinin çeşitli faktörlere göre karşılaştırılmasını ve böylelikle söz konusu yaklaşımların özelliklerini ortaya çıkarmayı amaçladığı için temel araştırma niteliği taşımaktadır. Temel araştırmalar, kuramlara dayalı olarak çeşitli varsayımları geliştirerek ve inceleyerek, elde edilen sonuçları bilimsel olarak yorumlayan araştırmalardır ve bu çalışma da belirtilen yönleriyle temel araştırmalar içinde yer almaktadır (Karasar, 2010).

Çalışma Grubu

Çalışmanın evreni TIMSS 2011 uygulamasına katılan tüm ülkeler ve tüm katılımcılardır. Söz konusu evrenin genişliği göz önüne alındığında, örnekleme yöntemleri kullanılarak bir örneklem belirlenmesi gerekliliği ortaya çıkmaktadır. Araştırmanın genel amacına ve alt amaçlarına uygun olarak seçkisiz olmayan örnekleme belirleme yaklaşımından, uygun örnekleme yöntemi kullanılarak birden çok örneklem belirlenmiştir.

Çalışma TIMSS 2011 uygulamasından elde edilen veriler ile gerçekleştirilmiş ve öncelikli olarak TIMSS 2011 uygulamasına ilişkin tüm veri seti söz konusu uygulamanın resmi internet sitesinden alınmıştır. Uygulama kapsamındaki başarı testlerine göre katılımcı ülkeler arasında ilk 20'de yer alan ülkeler çalışma grubuna dâhil edilmiştir. Başarı sıralarına göre ilk 20 ülkeye ilişkin verilerin çalışmaya dâhil edilme nedeni, aynı kitapçığı

yanıtlayan birey sayısını artırmaktır. Ayrıca alan yazında testin maddelerine ilişkin verilen yanıtların ranjının ve doğru yanıtlanma oranlarının yüksek olduğu uygulamalarda, MTK kapsamında model veri uyumunu etkilediği belirtilmektedir (Baker ve Kim, 2004). Bu nedenle çalışmanın temel veri seti oluşturulurken, TIMSS 2011 uygulamasında başarı sıralaması bakımından yüksek olan ülkeler tercih edilmiştir. Toplam 7242 kişiden oluşturulan çalışma grubunun ülkelere göre dağılımı ve ülke düzeyinde TIMSS 2011 uygulamasından elde edilen ortalama ölçek puanı dağılımı Çizelge 1’de verilmektedir.

Çizelge 1.

Çalışma Grubunun Ülkelere Göre Dağılımı ve Ortalama Ölçek Puanları

Çalışma Grubuna Alınan Ülkeler	Ortalama Ölçek Puanı	13 nolu Kitapçığı Yanıtlayan Birey Sayısı
Kore	613	370
Singapur	611	424
Çin-Tayvan	609	357
Hong Kong	586	290
Japonya	570	307
Rusya	539	357
İsrail	516	319
Finlandiya	514	312
Amerika	509	745
İngiltere	507	264
Macaristan	505	370
Avustralya	505	537
Slovenya	505	310
Litvanya	502	341
TIMSS Ölçek Orta Noktası	500	
İtalya	498	278
Yeni Zelanda	488	379
Kazakistan	487	316
İsveç	484	400
Ukrayna	479	240
Norveç	475	286
Toplam /Ortalama		7242

Çizelge 1’de özellikleri görülen bu veri seti, çalışmanın amaçları doğrultusunda farklı veri setlerine ayrılmıştır. Çalışma kapsamında madde ve yetenek parametrelerine etkisi incelenen olan ilk faktör örneklem büyüklüğü faktörüdür ve bu faktör altında incelemeye alınacak üç koşul bulunmaktadır. Bunlar örneklem büyüklüğünün 500, 1000 ve 3000 olduğu koşullardır. Söz konusu koşullar gözetilerek oluşturulan temel veri setinden seçkisiz olarak 500, 1000 ve 3000 kişilik veri setleri oluşturulmuştur.

Çalışma kapsamında madde ve yetenek parametrelerine etkisi incelenen bir diğer faktör ise, test uzunluğu koşuludur ve bu faktör altında da üç koşul incelemeye alınmıştır. Bu koşullar test uzunluğunun beş, 15 ve 25 olduğu koşullardır. Bu koşullar örneklem büyüklüğü faktöründe yer alan koşullar ile çaprazlanarak dokuz farklı veri seti oluşturulmuştur. Son olarak ise maddelerin psikometrik nitelikleri faktörü incelemeye alınmış ve bu faktör kapsamında incelemeye alınan dört koşula göre veri setleri oluşturulmuştur. Madde güçlük parametrelerinin yüksek ve düşük, ve madde ayırt edicilik parametrelerinin yüksek ve düşük olduğu dört ayrı veri seti oluşturulmuş ve tüm bu veri setlerinden birey yetenekleri ve madde parametreleri kestirilmiştir.

Veriler ve Elde Edilmesi

Bu çalışmada kapsamında incelenmeye alınan veriler TIMSS 2011 uygulaması kapsamında elde edilen verilerdir. Simülatif verilerle yapılan çalışmalarda ölçülen psikolojik değişkenlere ilişkin yorum yapmanın zorluğu ve gerçek verilerle çalışmanın uygulamaya yönelik katkılarının daha fazla olacağı göz önünde bulundurularak bu çalışma kapsamında TIMSS 2011 uygulamasında elde edilen verilerin incelenmesine karar verilmiştir.

Çalışma kapsamında incelemeye alınan verilerin belirlenmesi için, öncelikli olarak TIMSS 2011 uygulamasına ilişkin sunulan tüm veri seti söz konusu geniş ölçekli test uygulamasının internet sayfasından temin edilmiştir. TIMSS 2011 veri seti incelendiğinde, 4. ve 8. sınıf düzeyinde uygulanan Matematik ve Fen Bilimleri testlerine ek olarak, katılımcı ülkelerin eğitim sistemlerini, okul organizasyon yaklaşımlarını ve öğretim tekniklerini de incelemeyi amaçlayan geniş bir çerçevede bilgi sunduğu görülmekte ve toplanan tüm veriler uygulamaya ilişkin internet üzerinde oluşturulan veri tabanında yer almaktadır. Bu çalışma kapsamında incelemeye alınan veriler de söz konusu uygulamaya ait uluslararası veri tabanı kullanılarak elde edilmiştir (<http://timss.bc.edu/timss2011/international-database.html>).

TIMSS veri setinin çok geniş olması nedeniyle, çalışma kapsamında incelemeye alınan veri seti bazı kriterler gözetilerek belirlenmiştir. Bu kriterlerden ilki, kitapçıklarda yer alan madde sayısıdır. Sınıf düzeyi gözetildiğinde, her kitapçıkta 8. sınıf düzeyinde 4. sınıf düzeyine göre daha fazla madde yer aldığı belirlenmiş ve bu nedenle çalışma

kapsamında 8. sınıf başarı testlerinin kullanılmasına karar verilmiştir. İncelemeye alınan kitapçığın Matematik veya Fen Bilimleri testlerinden hangisine yönelik olacağına belirlenmesi için ise 8. sınıf düzeyinde uygulanan tüm kitapçıklar incelenmiştir.

TIMSS uygulamalarında yer alan Matematik ve Fen Bilimleri testleri 14 kitapçıktan oluşmakta ve her kitapçıkta iki Fen Bilimleri ve iki Matematik olmak üzere dört blok yer almaktadır. Her iki bloktan biri farklı kitapçıklarda kullanılarak, test eşitleme çalışmaları yapılmaktadır. Söz konusu kitapçıklar incelendiğinde, 8. sınıf düzeyinde Matematik başarı testi kapsamında dört içerik alanına yönelik maddelerin bulunduğu belirlenmiştir. Bunlar; Sayılar, Cebir, Geometri ve Veri ve Olasılıktır. TIMSS kapsamında uygulanan bir diğer başarı testi olan Fen Bilimleri testinde de 8. sınıf düzeyinde dört içerik alanı olduğu belirlenmiştir. Bunlar; Biyoloji, Kimya, Fizik ve Yer Bilimleridir. Bu çalışma kapsamına alınan kitapçığın belirlenmesinde, madde sayısına ek olarak gözetilen diğer kriter ise, söz konusu testin tek boyutluluk özelliğini mümkün olduğunca yüksek düzeyde karşılıyor olmasıdır. Bu kriterin incelemeye alınma nedeni hem PMTK hem POMTK modellerinin tek boyutluluğu varsayım olarak kabul etmesidir. Bu varsayımının hangi alanda daha fazla karşılandığını belirlemek için, Matematik ve Fen Bilimleri alanında uygulanan tüm kitapçıklardan elde edilen verilere tetrakorik korelasyon matrisi kullanılarak Açıklayıcı Faktör Analizi (AFA) uygulanmış ve Matematik testlerinin Fen Bilimleri testlerine göre, tek faktörde daha yüksek öz değere sahip olduğu ve Matematik testlerinde, daha fazla maddenin tek boyut altında toplandığı belirlenmiştir. Bu nedenle çalışma kapsamına incelenen verilerin 8. sınıf düzeyinde Matematik başarı testine yönelik olmasına karar verilmiştir.

Çalışma kapsamında incelemeye alınan kitapçığın belirlenmesi için, 8. Sınıf düzeyinde uygulanan 14 Matematik testi kitapçığı TIMSS 2011 veri tabanından elde edilen Madde Bilgi Tabloları (Item Information Tables) kullanılarak incelenmiştir. Matematik testi kapsamında yer alan 14 blok öncelikle madde sayıları bakımında incelenmiş ve 9. ve 13. kitapçığın sırasıyla 35 ve 33 maddeden oluşarak en fazla maddeye sahip olduğu belirlenmiştir. Çalışma kapsamında incelemeye alınan kitapçığın belirlenmesi için ise yine tek boyutluluk varsayımını göz önünde bulundurulmuş ve belirlenen iki kitapçığa ait verilere tetrakorik korelasyon matrisine dayalı olarak Açıklayıcı Faktör Analizi uygulanmıştır. Her iki kitapçık için gerçekleştirilen analizde, 13 nolu kitapçıkta tek faktör

için açıklanan varyans, %30.737 iken, 9 nolu kitapçıkta tek faktör için açıklanan varyans %27,572 bulunmuştur. Veri setinin belirlenmesinde tek boyutluluk varsayımının en yüksek düzeyde karşılanması istendiği için, 13 nolu kitapçığın çalışma kapsamında analize alınmasına karar verilmiştir. İncelemeye alınan kitapçıkta yer alan 33 maddenin 3'ü 0-1-2 şeklinde çok kategorili olarak puanlanırken, kalan 30 maddenin 1-0 şeklinde puanlanmakta olduğu belirlenmiştir. Bu çalışma iki kategorili maddelere sınırlandırıldığı için, çok kategorili puanlanan üç madde çalışma kapsamı dışında tutulmuş ve 30 madde ile analizlere devam edilmiştir.

Çalışmada İncelemeye Alınan Faktörler

Bir önceki başlıkta belirtilen işlemlere göre elde edilen veri seti çalışma kapsamında incelemeye alınan faktörlere göre gruplara ayrılmıştır. Çalışma kapsamında incelemeye alınan faktörler; veri setlerinde yer alan birey sayısı, test uzunluğu ve maddelerin psikometrik nitelikleridir. Uygulamada bu faktörlerin PMTK ve POMTK modelleri ile yapılan madde parametre kestirimlerini, birey yeteneklerini ve madde tepki fonksiyonlarını etkilediği belirtilmektedir (Molenaar, 1997; Van Schuur, 2011; Sijtsma ve Molenaar, 2002).

Araştırmada ele alınan ilk faktör gruplarda yer alan birey sayısıdır. Bu faktör için 500, 1000 ve 3000 kişilik veri setleri oluşturulmuştur. Bu örneklem büyüklüklerinin belirlenmesinde PMTK uygulamaları temel alınmıştır. İncelemeye alınan örneklem büyüklüğü değerlerinden 500 kişilik veri seti PMTK uygulamaları için küçük örneklem olarak nitelendirilirken, 1000 orta büyüklükte, 2000 ve 3000 ise büyük örneklem olarak nitelendirilmektedir. Büyük örneklem için 2000 yerine 3000 değerinin alınmasının nedeni, PMTK modellerinin 2000 kişilik bir örneklemde dahi, ölçülen psikolojik özelliğe ilişkin ranjının geniş olmadığı durumlarda, uç yetenek düzeylerinde yer alan bireylere ilişkin yetenek kestirimlerinin yüksek standart hata içermesinden dolayı alan yazında eleştiri almış olmasıdır. En küçük örneklem büyüklüğü için 500 kişilik veri setinin oluşturulma nedeni ise, POMTK modellerinin küçük örneklemelerde de etkili kestirimler yaptığı iddiasının test edilmesidir (Sijtsma ve Molenaar, 2002; Van Schuur, 2011). Alan yazında PMTK modelleri için yapılan analizlerde hatasız madde parametreleri kestirimi için en az 1000 kişilik veri setinin gerekli görüldüğü göz önünde bulundurularak (Mislevy ve Stocking,

1989), 500 kişilik veri setinin küçük ve 1000 kişilik veri setinin ise orta büyüklükte olarak nitelendirilebileceği düşünülmüştür.

Çalışma kapsamında incelenen bir diğer faktör ise test uzunluğudur. Bu faktör kapsamında, testte yer alan madde sayısının madde ve birey yetenek kestirimlerine etkisinin incelenmesi amaçlanmıştır. Bu durumun nedeni ise POMTK modellerin testteki madde sayısının az olduğu durumlarda dahi birey yeteneklerine ve madde parametrelerine ilişkin PMTK modellerinden daha gerçekçi ve etkili kestirimler yapıyor olmasıdır (Van Schuur, 2011; Sijtsma ve Molenaar, 2002). Ayrıca özellikle madde sayısının az olduğu durumlara psikolojik testlerde sıkça rastlanmaktadır ve bu koşulun incelemeye alınması çalışma sonuçlarının uygulamada ortaya çıkabilecek durumlara yansıtılmasını sağlayabilecektir. Bu faktör altında yer alan koşulların incelenmesi için, TIMSS 2011 Matematik 8. Sınıf başarı testinden belirlenen kitapçık beş, 15 ve 25 maddeden oluşacak şekilde seçkisiz olarak üç ayrı teste ayrılmış ve bu testlerden madde ve birey parametreleri kestirilmiştir.

Çalışma kapsamında incelemeye alınan son faktör ise maddelerin özellikleridir. Maddelerin güçlük ve ayırt edicilik özelliklerinin POMTK ve PMTK modelleri ile kestirilen madde ve yetenek parametrelerine olan etkisinin incelenmesi için madde güçlük ve ayırt edicilik değerlerinin yüksek değerler aldığı maddeler belirlenmiş ve bu maddelere ilişkin madde parametreleri ile birey yetenekleri her iki yaklaşıma göre kestirilmiştir. Bu faktörün çalışmaya dâhil edilme nedeni, özellikle ayırt edicilik ve güçlük değerleri normalden düşük ya da yüksek olan maddeler için PMTK modellerine göre POMTK modellerinin daha ayrıntılı kestirimler yapıyor olması, daha gerçekçi MTF'ler ürettiyor olması ve bu maddeler için madde model uyumunun PMTK'ya göre POMTK'da daha yüksek düzeyde sağlanıyor olmasıdır (Junker, 2000). Bu faktör kapsamında dört ayrı koşul incelenmiştir. Bunlar madde ayırt edicilik indekslerinin yüksek ve düşük olduğu koşullar ile madde güçlük parametresinin yüksek ve düşük olduğu koşullardır. Bu faktöre ilişkin tüm koşullar ayrı olarak incelenirken, örneklem büyüklüğü ve test uzunluğu faktörlerine ilişkin koşullar çaprazlanarak incelemeye alınmış ve toplam 16 veri seti ile çalışma yürütülmüştür.

Verilerin Analizi

Bu kısımda öncelikli olarak araştırma kapsamında incelenmesine karar verilen 33 maddeden oluşan TIMSS 2011 uygulaması 8. sınıf Matematik alt testinin 13 nolu kitapçığının MTK varsayımlarını karşılayıp karşılamadığı incelenmiştir. MTK'nın varsayımlarından tek boyutluluğun sınanmasında Açımlayıcı Faktör Analizi (AFA) ve Doğrulayıcı Faktör Analizi (DFA) kullanılmıştır. Açımlayıcı faktör analizinin 1-0 şeklinde puanlanan kategorik verilerde uygulanabilmesi için öncelikli olarak STATISTICA programında tetrakorik korelasyon matrisi üretilmiştir. Oluşturulan tetrakorik korelasyon matrisi SPSS 22.0 programına aktarılmış ve bu matris üzerinden açımlayıcı faktör analizi yapılarak, PMTK ve POMTK modellerinin gereği olan tek boyutluluk varsayımının karşılanma durumu incelenmiştir. Her iki kuramında bir diğer varsayımı olan yerel bağımsızlık incelemesi için AFA kapsamında artık korelasyon matrisi oluşturulmuş ve bu matriste yer alan değerler incelenmiştir.

PMTK kapsamında verinin en iyi uyum sağladığı modeli belirlemek için BILOG-MG programı ve R programında yer alan “irtoys” ve “lrm” paketleri kullanılarak madde parametreleri ve birey yetenek parametreleri kestirilmiştir. PMTK yaklaşımı kapsamında bir, iki ve üç parametrelili lojistik modele ilişkin model veri uyumları ayrı ayrı incelenmiş ve model veri uyumu test edilirken -2 loglikelihood değerleri ile maddelere ilişkin hesaplanan ki-kare değerleri kullanılmıştır. Çalışma kapsamında PMTK ve POMTK kestirimlerinde kullanılan teste ilişkin güvenilirliğin belirlenmesi için ise PMTK kapsamında test bilgi fonksiyonları oluşturulmuş ve marjinal güvenilirlik katsayıları hesaplanmıştır. POMTK ile incelenen veri setleri için yapılan güvenilirlik analizinde ise, Lambda 2 ve Cronbach Alfa katsayılarına ek olarak alana özgü LRSCR ve MS katsayıları da hesaplanmıştır. Oluşturulan temel veri setinin PMTK ve POMTK modellerine uygunluğunun test edilmesine ilişkin gerçekleştirilen analizler bu bölümde sırasıyla verilmiştir.

Parametrik Madde Tepki Kuramı Modellerine İlişkin Varsayımların Test Edilmesi

Bu başlık altında oluşturulan temel veri setinin PMTK varsayımlarına uyumu incelenmiş ve söz konusu varsayımların veri seti için karşılanma durumları belirlenmiştir. İncelenen varsayımlar, tek boyutluluk ve yerel bağımsızlıktır.

Tek Boyutluluk

Parametrik Madde Tepki Kuramı modellerinin varsayımlarından biri olan tek boyutluluk aynı zamanda parametrik olmayan modeller için de geçerlidir. Bu varsayım kapsamında belirli bir maddeye doğru yanıt verme olasılığının bireyin tek bir özelliği ya da yetenek düzeyiyle belirlendiği kabul edilmektedir (Hambleton ve Swaminathan ve Rogers, 1991). Bu araştırmada PMTK için tek boyutluluk varsayımı “Lumsden Yöntemi” olarak da bilinen açımlayıcı faktör analizi ile incelenmiştir.

Açımlayıcı faktör analizi.

Veri setinin boyutluluk özelliğini incelemek adına yapılan AFA'ya geçilmeden önce söz konusu analizin varsayımlarının test edilmesi ve verilerin bu analize uygun olup olmadığının belirlenmesi gerekmektedir. Bu nedenle AFA'nın varsayımları olan; uç değerler (birey ve madde düzeyinde), örneklem büyüklüğü, kayıp veri (değişken ve birey düzeyinde), normallik (tek ve çok değişkenli), doğrusallık, çoklu bağlantılılık ve teklik durumu ile faktörleştirilebilirlik incelemeleri yapılmıştır (Tabachnick ve Fidell, 2007). Tüm varsayımlar SPSS 22 programında incelenmiştir. Oluşturulan temel veri seti için ilk olarak değişken ve birey düzeyinde uç değer incelemesi yapılmıştır. Uç değerler tek değişkenli ve çok değişkenli olarak incelenmiş, tek değişkenli uç değerler için bireylerin Z puanları hesaplanmış, çalışma örneklemini geniş örneklem olarak nitelendirilebilecek değerlerde olduğu için sınır değer olarak 4, -4 değerleri kabul edilmiştir (Field, 2009). Bireylerin toplam puanları temel alınarak hesaplanan Z puanları incelendiğinde, +4 ve -4 aralığı dışında yer alan herhangi bir değer olmadığı belirlenmiş ve çok değişkenli uç değer analizine geçilmiştir. Çok değişkenli uç değerlerin incelenmesi için ise Mahalanobis uzaklıkları hesaplanmış, her katılımcı için hesaplanan Mahalanobis değerleri, bu çalışma için 0.001 manidarlık derecesi ve 29 serbestlik derecesi temel alınarak belirlenen 50.892 ile kıyaslanmıştır. Mahalanobis uzaklığı belirlenen sınır değerden yüksek olan 80 katılımcıya ait veriler veri setinden çıkarılmıştır ve yeni veri seti 7162 kişiden oluşturulmuştur.

Uç değer analizinden sonra, AFA'nın bir diğer gerekliliği olan örneklem büyüklüğünün yeterliliği incelemesi yapılmıştır. Örneklem büyüklüğünün yeterliği için literatürde farklı görüşler bulunmaktadır. Hair, Black, Anderson ve Tatham (1997), değişken sayısının en az beş katı büyüklüğünde örnekleme sahip olunmasının yeterli

olduğunu belirtirken, bunun yanında bire on ya da bire 20 oranlarının da literatürde yer aldığını ifade etmektedir. Bu çalışma kapsamında 30 madde incelemekte ve 7162 kişilik bir çalışma grubu kullanılmaktadır. Dolayısıyla literatürde verilen bire 10 ve bire 20 oranlarının oldukça üstünde geniş bir katılımcı grubuyla çalışılmaktadır. Bu duruma ek olarak, bu çalışma için örneklem büyüklüğünün yeterliği açılımlayıcı faktör analizi kapsamında yer alan Kaiser-Meyer Olkin testi ile de incelenmiş ve test sonucu ,970 olarak bulunmuştur. Bu değer örneklem büyüklüğünün analiz için mükemmel düzeyde yeterli olduğu anlamına gelmektedir (Tabachnick ve Fidel, 2007) ve bu gerekliliğin sağlandığı kabul edilerek analizin bir diğer varsayımı olan kayıp değer incelemesine geçilmiştir.

AFA yapılmadan önce incelenmesi gereken bir diğer varsayımı ise veri setinin kayıp verilerden arınık olmasıdır. Bu varsayım hem değişken hem de katılımcıların yanıt örüntüleri düzeyinde incelenmiştir. Değişken düzeyinde kayıp veriler incelendiğinde, incelemeye alınan 30 maddelik değişken setinde değişkenlere göre kayıp veri oranlarının %1,7 ile %9,7 arasında değiştiği belirlenmiştir. Değişkenlere göre yapılan kayıp veri analizinde, kayıp veri örüntüleri incelenmiş, veri setinde düzeyinde %6 oranında kayıp veri bulunduğu belirlenmiştir. Tabachnick ve Fidell (2007) geniş veri setlerinde seçkisiz bir örüntü sergileyen az sayıda kayıp değer var ise veri dosyasındaki sorunun çok ciddi olmadığını savunmaktadır. Geniş bir veri seti olarak nitelendirilebilecek çalışmanın veri setinde yer alan kayıp verilerinin seçkisizliği Little M'CAR testi ile incelenmiş ve kayıp değerlerin kısmen seçkisiz olarak dağıldığı belirlenmiştir. Gerek düşük düzeyde kayıp verinin olması gerekse seçkisizliğin kısmen sağlanmış olması göz önünde bulundurularak, kayıp veriler sıfır olarak kodlanmıştır. Kayıp değerlerin yerine çoklu atama yoluyla değer atamak yerine doğrudan sıfır atanmasının nedeni, çalışma kapsamında yapılacak PMTK ve POMTK incelemelerinde kullanılacak BILOG ve R programı kapsamındaki Mokken paketinin çoklu atama yöntemiyle oluşturulan veri setlerinin analiz edilememesidir. Kayıp verilerin sıfır olarak tekrar kodlanmasından sonra, açılımlayıcı faktör analizinin gereklerinin incelenmesine devam edilmiştir.

Kayıp verilerin incelenmesinden sonra, açılımlayıcı faktör analizinin bir diğer gerekliliği olan tek ve çok değişkenli normallik veri seti tarafından sağlanma durumu incelemeye alınmıştır. Tek değişkenli normallik incelemesi için, veri setinde yer alan tüm değişkenlerin çarpıklık ve basıklık değerleri incelenmiştir. Tüm değişkenler için bu

değerlerin -1, +1 arasında olduğu bulunmuş ve bu durum da söz konusu değişkenler için tek değişkenli normalliğin sağlandığı anlamına geldiğinden, çok değişkenli normallik incelemesine geçilmiştir. Çok değişkenli normallik incelemesi için ise Bartlett Küresellik Testi uygulanmış ve bu test sonucunun manidar olduğu bulunmuştur. Bu teste ek olarak, çok değişkenli normallik incelemesi için grafiksel yöntem de kullanılmış ve veri setinde yer alan tüm değişkenlerin ikili kombinasyonları saçılma diyagramları oluşturularak incelenmiştir. Tüm ikili kombinasyonların elips şeklinde dağılım gösterdiği belirlenmiş ve bu sonucun Bartlett Küresellik Testi sonucu ile uyumlu olduğu dolayısıyla incelenen veri setinin tek ve çok değişkenli normallik varsayımını karşıladığı sonucuna ulaşılmıştır. Çok değişkenli normallik incelemesi için oluşturulan saçılma diyagramları aynı zamanda analizin bir diğer varsayımı olan doğrusallık analizi için de kullanılmış ve elde edilen elips şeklindeki dağılımlar, veri setinin doğrusallık varsayımını karşıladığını göstermiştir.

Değişkenler arasındaki çoklu bağlantı problemi de faktör analizi için önemli ve incelenmesi gereken bir varsayımdır. Analize alınan değişkenler arasında çoklu bağlantı problemi olduğu saptanırsa, bu sorunu oluşturan değişkenlerin belirlenip analizden çıkarılması gerekmektedir. Değişkenler arasındaki çoklu bağlantılılık durumunun incelenmesi için çeşitli istatistikler dikkate alınmakta ve bu çalışmadaki verilerin çoklu bağlantı probleminin olup olmadığının incelenmesi için VIF, CI ve tolerans değerleri incelenmiştir. Yapılan analiz sonucunda veri setinde yer alan değişkenler için hesaplanan varyans artış faktörlerinin 1.088 ile 1,680 arasında değiştiği, tolerans değerlerinin 0.85 ile 0.60 arasında değiştiği ve CI indeksinin de 1.00 ile 18.055 arasında değiştiği bulunmuştur. Tüm bu istatistiklerin ilgili sınır değerlerin altında olduğu bulunmuş, dolayısıyla veri seti için çoklu bağlantı durumunun olmadığı sonucuna ulaşılmıştır.

Açımlayıcı faktör analizi için incelenen son varsayım ise, değişkenlerin faktörleştirilebilirliğinin incelenmesidir. Faktörleştirilebilirlik, faktör analizi için oluşturulan matrisin, belli düzeyde korelasyon içermesini gerektirmekte ve 0.30'dan daha düşük korelasyonların çoğunlukta olduğu bir veri seti için faktör analizinin yapılmasının uygunluğunun tekrar gözden geçirilmesi gerekmektedir. Bartlett'in Küresellik Testi bu konuda bilgi verebilen bir istatistik olmasına rağmen, örneklem büyüklüğü söz konusu testin manidar çıkma olasılığını artırmaktadır, dolayısıyla testin güvenilirliği örneklem büyüklüğüne bağlı olarak azalmaktadır (Tabachnick ve Fidell, 2007). Bu çalışmadaki

değişkenlerin faktörleştirmeye uygun olma durumu anti-imaj korelasyon matrisi ile incelenmiştir. Söz konusu matrisin köşegen değerleri dışında yer alan değerlerin büyük çoğunluğu 0'a yakındır, dolayısıyla değişkenlerin faktörleştirmeye uygun olduğu sonucuna ulaşılmıştır.

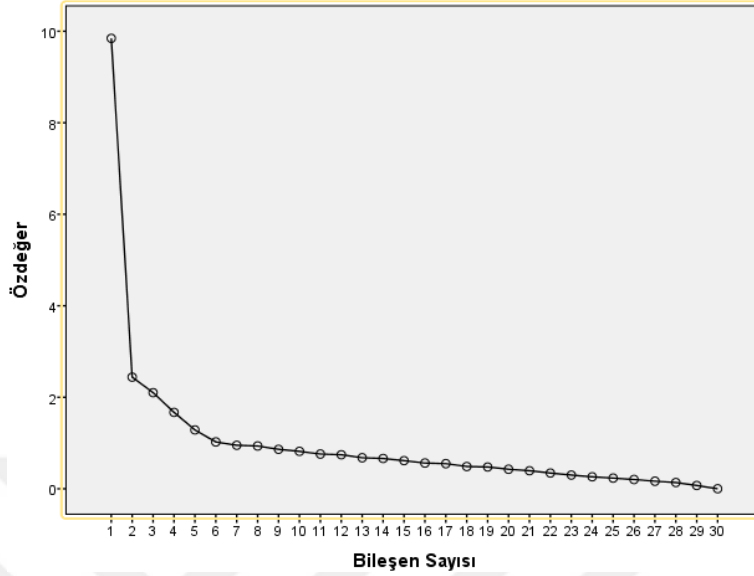
Açımlayıcı faktör analizine ilişkin varsayımların ve gereklerin veri seti tarafından karşılandığı sonucuna ulaşıldıktan sonra, Temel Bileşenler Analizi kullanılarak analiz gerçekleştirilmiştir. Analizin ilk aşamasında veri setinin öz değeri birden büyük olan altı bileşene sahip olduğu ve bu bileşenlerin toplam varyansın %61,223'ünü açıkladığı belirlenmiştir. Hesaplanan altı bileşen ve bu bileşenlerin özdeğerleri ile açıkladıkları varyans oranları Çizelge 2'de verilmiştir.

Çizelge 2.

AFA Sonucu Hesaplanan Bileşenler, Özdeğerleri ve Açıklanan Varyans Oranları

Bileşen	Öz değer	Açıklanan Varyans %	Açıklanan Toplam Varyans %
1	9,846	32,820	32,820
2	2,440	8,135	40,995
3	2,101	7,004	47,958
4	1,670	5,567	53,525
5	1,286	4,286	57,811
6	1,024	3,412	61,223

Çizelge 2'de yer alan değerler incelendiğinde, AFA sonucu elde edilen ilk bileşence açıklanan varyans oranının %32,820 ve bu bileşene ilişkin hesaplanan öz değerinin 9,846 olduğu görülmektedir. İkinci bileşence açıklanan varyans oranının ise %8,135 ve bu bileşene ilişkin öz değerinin 2,440 olduğu çizelgede görülmektedir. Bu değerlere göre ilk bileşenin açıkladığı varyans oranının ikinci bileşenin açıkladığı varyans oranından yaklaşık olarak dört kat daha yüksek olduğu ve öz değerler arasında da benzer bir oranın olduğu görülmektedir. Bu durum veri setinin tek boyutlu olarak kabul edilebileceğine ilişkin kanıt oluşturmaktadır. Ayrıca ilk bileşenden sonraki bileşenler açıklanan varyans oranlarında büyük bir düşüş olduğu ve diğer bileşenlerin açıkladıkları varyans oranlarının ve dolayısıyla öz değerlerinin benzer olduğu göz önüne alındığında, veri setinin başat bir boyutla özetlenebileceği sonucuna ulaşılmıştır. Benzer bir sonucu oluşturulan yamaç birikinti grafiğinden gözlemlemek de mümkündür. Bu çalışmada kullanılan veri seti için oluşturulan yamaç birikinti grafiği Şekil 2.'de görülmektedir.



Şekil 2. Yamaç Birikinti Grafiği

Alan yazında tek boyutluluğa ilişkin yer alan bilgiler incelendiğinde, AFA sonucunda başat bir faktörün elde edilmesi tek boyutluluk göstergesi olarak kabul edilmektedir (Crocker ve Algina, 1986; Hambleton ve Swaminathan ve Rogers, 1991; Stout, 1997). AFA’da, başat bir faktöre ait açıklanan varyans değerinin %30’dan büyük olması, yamaç birikinti grafiğinde bileşenlerin ivmelerine göre farkların anlamsız hale gelmesi, özdeğerler arasındaki farkın 1/3 oranından büyük olması gibi kriterler dikkate alınarak veri setinin boyutluluğuna karar verilmektedir (Field, 2009). Bu analiz sonucu elde edilen bulguların alan yazındaki bilgilerle paralellik gösterdiği göz önüne alınarak, veri setinin tek boyutlu bir yapıda olduğuna karar verilmiş ve analiz tek bileşen için tekrarlanmıştır.

Tek bileşen için tekrarlanan AFA sonuçlarına göre ilk incelenen bulgu, maddelere ilişkin ortak varyansın verildiği ortak varyans (communalities) değerleridir. Bu değerler maddelerin ortak varyansa olan katkılarını belirtmektedir ve bu analizdeki maddelerin ortak varyans değerlerinin 0.01 ile 0.71 arasında değiştiği belirlenmiştir. Veri setinde yer alan maddeler içinde ortak varyans değeri 0.10’un altında olan üç madde olduğu belirlenmiştir ve bu maddelerin faktör yüklerinin de 0.32’nin altında olduğu bulunmuştur. Faktör yükleri, 0.32’nin altında olan üç madde faktör yükü en düşük olandan en yükseğe göre 25, 24 ve 12

nolu maddeler olarak sıralanmaktadır. Testte yer alan tüm maddelerin faktör yükleri Çizelge 3'te verilmiştir.

Çizelge 3.

Maddelere İlişkin Faktör Yük Değerleri

Madde	Faktör Yük Değeri	Madde	Faktör Yük Değeri
M1	0.495	M16	0.630
M2	0.619	M17	0.728
M3	0.490	M18	0.624
M4	0.852	M19	0.706
M5	0.479	M20	0.451
M6	0.631	M21	0.586
M7	0.549	M22	0.554
M8	0.748	M23	0.834
M9	0.843	M24	0.293
M10	0.359	M25	0.270
M12	0.309	M26	0.390
M11	0.446	M27	0.517
M13	0.647	M28	0.438
M14	0.637	M29	0.591
M15	0.576	M30	0.349
		Açıklanan Varyans	34,021

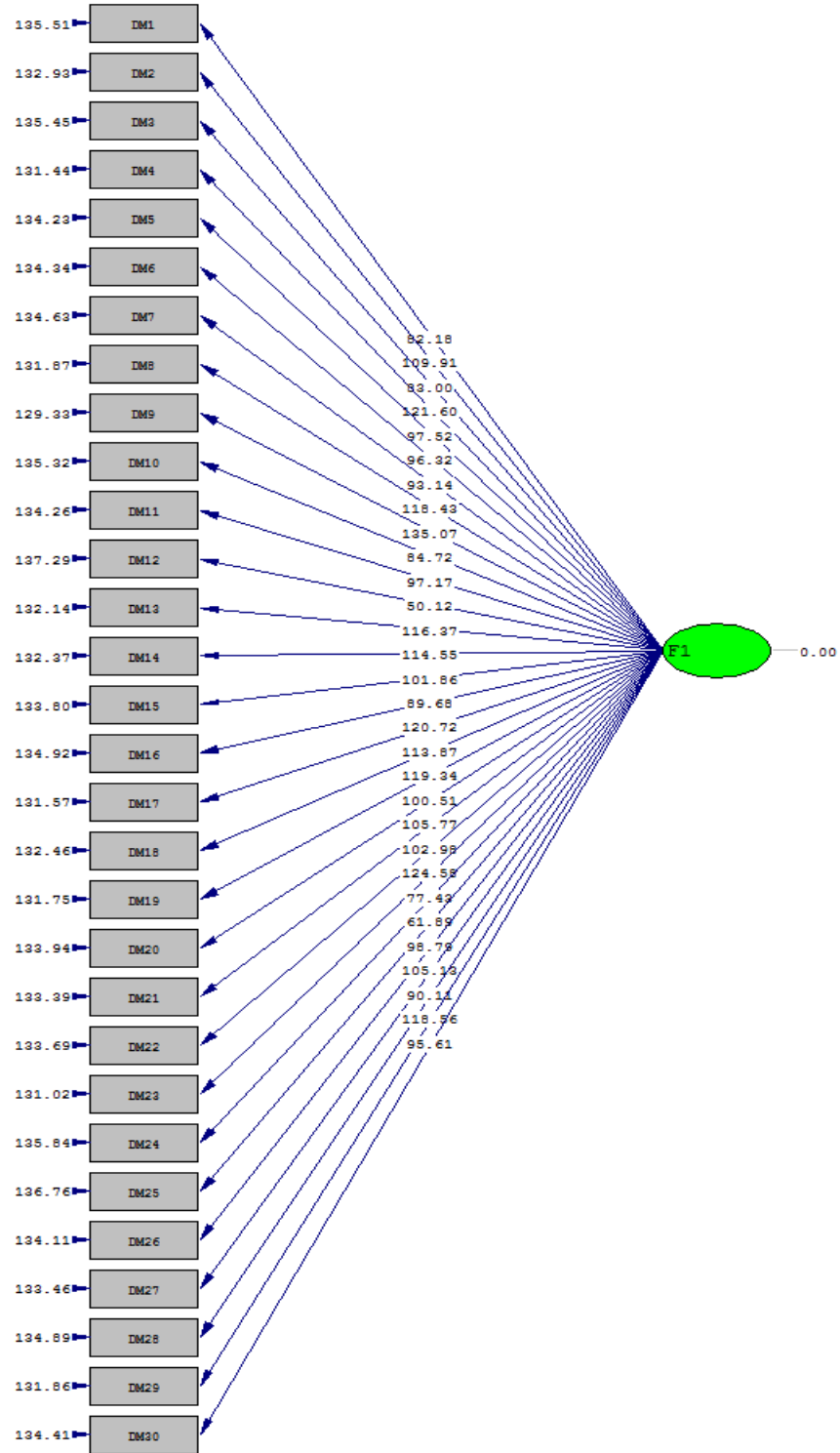
Çizelge 3 incelendiğinde, daha önce belirtilen üç madde hariç tüm maddelerin faktör yüklerinin sınır değer olan 0.32'nin üstünde olduğu görülmektedir. Ayrıca tek bileşen tarafından açıklanan toplam varyansın da 30'un üstünde olduğu görülmektedir. Faktör yük değerleri kabul değerlerinin altında olan maddeler veri setinden çıkarılmadan önce, belirlenen tek boyutlu yapının doğrulayıcı faktör analizi ile incelenmesine karar verilmiştir. Bu duruma karar verilirken de, incelemeye alınan testin daha önce faktör yapısının belirlenmiş olduğu göz önünde bulundurulmuş ve doğrulayıcı faktör analizine başvurmanın daha kesin bilgi vereceği düşünülmüştür. Bu nedenle analizlere doğrulayıcı faktör analizi ile devam edilmiştir.

Doğrulayıcı faktör analizi.

Bu çalışma kapsamında yapılan açımlayıcı faktör analizinde başat bir boyutun olduğu sonucuna varılan veri seti için Doğrulayıcı Faktör Analizi (DFA) yapılarak yapının tek boyutluluğunun test edilmesine karar verilmiştir. Oluşturulan tek boyutlu model için DFA yapılmış ve analiz sonucunda ilk olarak maddelere ilişkin t değerleri incelenmiştir. Gizil değişkenlerin gözlenen değişkenleri açıklama durumlarına ilişkin bilgi veren t

değerleri 1.96'nın üzerinde olması 0.05 düzeyinde, 2.56'nın üzerinde olması ise 0.01 düzeyinde manidar olarak kabul edilmektedir (Kline, 2000). Bu bilgilere göre elde edilen t değerleri incelenmiş ve maddelere ait t değerlerinin 61.89 ile 135.07 arasında değiştiği gözlenmiştir. Dolayısıyla modele ilişkin tüm t değerlerinin 0.01 düzeyinde manidar olduğu sonucuna ulaşılmıştır. Hesaplanan t değerleri ve önerilen modele ilişkin oluşturulan yol diyagramı Şekil 3'te verilmiştir.





Şekil 3. DFA Sonucu Elde Edilen Yol Diyagramı ve T Değerleri

Doğrulayıcı faktör analizinde maddelere ilişkin incelenmesi gereken bir başka değer de standartlaştırılmış yüklerdir. Standartlaştırılmış yükler, gözlenen değişkenlerle ilgili olduğu gizil değişkenler arasındaki ilişkiye yönelik bilgi vermektedir ve bu yükler gizil değişkenlerin maddelerde neden olduğu etki göstergeleri olarak da adlandırılabilir (Kline, 2000). Yapılan analiz sonucunda testte yer alan maddelerin standartlaştırılmış yüklerinin 0.32 ile 0.64 arasında değiştiği ve tüm katsayıların sınır değer olarak kabul edilen 0.30'un üzerinde olduğu görülmüştür. Doğrulayıcı faktör analizinde madde düzeyindeki uyum istatistiklerine ek olarak, modelin genel olarak uyumunun değerlendirilmesi için farklı göstergeler de işe koşulur ve bunlardan biri de modele ilişkin ki-kare değeridir. Bu değer, 20728.38 olarak hesaplanmıştır ve 0.01 düzeyinde manidardır. Söz konusu değer model uyumu genel olarak serbestlik derecesine oranlanarak incelenmektedir. Kelloway (1996; akt. Kline, 2000), χ^2/sd (ki kare/serbestlik derecesi) oranının beşten küçük olmasını iyi uyum göstergesi olarak belirtmektedir. Analiz sonucunda χ^2/sd oranının 51.1 bulunduğu belirlenmiştir ve bu değer iyi uyum göstergelerinin çok üzerinde olduğu sonucuna ulaşılmıştır, fakat söz konusu değer yüksek olma nedeni olarak örneklem büyüklüğünün çok yüksek olması düşünülmüş ve modelin uyumuna ilişkin bilgi veren diğer istatistikleri Çizelge 4'te incelenmiştir.

Çizelge 4.

DFA'dan Elde Edilen Model Uyum İndeksleri

Uyum İndeksi	RMSEA	GFI	AGFI	SRMR	IFI	NNFI	CFI
Elde Edilen Değer	0.039	0.96	0.95	0.027	0.99	0.98	0.98
Uyum Düzeyi	İyi	Çok iyi	Çok iyi	Çok iyi	Çok iyi	Çok iyi	Çok iyi

Çizelge 4'te yer alan indeksler var olan model ile önerilen modelin genel uyumunu yansıtmaktadır. Ortalama Hataların Karekökü (RMSEA) Çizelge 4'te yer alan ilk indeks değeridir ve bu değer sifıra yakın olması iyi uyum göstergesidir. Elde edilen RMSEA değerinin 0'a yakınlığı gözetildiğinde, modelin iyi düzeyde uyum sağladığını göstermektedir. Çizelgede yer alan bir diğer indeks olan SRMR istatistiği de RMSEA'ya benzer şekilde yorumlanır ve küçük değerler iyi uyum göstergesi olarak kabul edilir. Bu analiz için elde edilen SRMR değerinin de 0'a çok yakın olduğu göz önünde bulundurulduğunda, model uyumun iyi olduğu sonucuna ulaşılabilir. Ki-kare uyum testine

alternatif olarak gösterilebilen Uyum İyiliği İndeksi, GFI, 0.96 olarak hesaplanmış ve bu değere göre de model veri uyumun oldukça iyi düzeyde sağlandığı sonucuna ulaşılmıştır. Örneklem büyüklüğü dikkate alınarak hesaplanan Düzeltilmiş İyilik Uyum İndeksi olan AGFI indeksi de 0.95 olarak bulunmuş ve çok iyi uyumun göstergesi olarak kabul edilmiştir. Bir diğer uyum iyiliği indeksi olan, Artmalı Uyum İndeksi (IFI), örneklem büyüklüğünü ve modelin serbestlik derecesini dikkate alan bir indekstir ve bire yakın indeks değerleri uyumun oldukça iyi olduğunu göstermektedir. Bu modeldeki IFI değeri 0.99 olarak bulunmuş, dolayısıyla model veri uyumunun çok iyi düzeyde olduğu sonucuna ulaşılmıştır. Oluşturulan modelin serbestlik derecesini dikkate alarak hesaplanan Normlaştırılmış Uyum İyiliği İndeksi (NNFI), 0.98 olarak hesaplanmış ve bu değer Schermelleh-Engel ve Moosbrugge (2003) tarafından belirlenen ölçütlere göre mükemmellik düzeyinde uyumu göstermektedir. Çizelge 8’de yer alan son model uyum iyiliği indeksi olan Karşılaştırmalı Uyum İndeksi (CFI), öngörülen modelin ürettiği kovaryans matrisi ile önerilen modelin kovaryans matrisini kıyaslamakta ve bire yakın değerler mükemmel uyuma karşılık gelmektedir. Bu analizde CFI değeri, 0.96 olarak bulunmuş, dolayısıyla model veri uyumunun çok iyi düzeyde olduğunu sonucuna ulaşılmıştır (Jöroskog ve Sörbom, 2001; Schermelleh-Engel ve Moosbrugger, 2003).

Çizelge 4’te verilen model veri uyumu indekslerine göre, tek boyutlu olarak önerilen modelin veriyle yüksek düzeyde uyum sağladığı sonucuna ulaşılmıştır. Ayrıca madde düzeyindeki istatistiklerin de uyumlu olduğu sonucu gözetilerek, testte yer alan 30 maddenin tek boyutlu bir yapıya sahip olduğu sonucuna ulaşılmış ve PMTK ve POMTK’nın ortak varsayımı olan tek boyutluluğun bu test için sağlandığı kabul edilerek, bir diğer varsayım olan yerel bağımsızlığın test edilmesine geçilmiştir.

Yerel Bağımsızlık

Parametrik ve parametrik olmayan madde tepki kuramı modelleri için tek boyutluluktan sonra sorgulanması gereken bir diğer varsayım yerel bağımsızlıktır. Yerel bağımsızlık, tek boyutluluk varsayımı ile paralel olarak kabul edilmekte ve tek boyutluluk varsayımının karşılandığı durumlarda yerel bağımsızlık varsayımının da karşılanacağı Lord (1970) tarafından belirtilmektedir. Ayrıca yerel bağımsızlık varsayımının ayrı bir varsayım olarak incelenmesine gerek olmadığı da aynı kaynakta ifade edilmektedir. Farklı

kaynaklarda ise bu varsayımın da test edilmesi gerektiği belirtilmekte ve çeşitli yöntemler önerilmektedir. Değişkenlere ilişkin artık korelasyon matrisinin incelenmesi ve tüm ikili çiftlere ait korelasyonların katsayılarının düşük bulunması bu varsayımın karşılandığına ilişkin kanıt olarak kabul edilmektedir (Embretson ve Reise, 2000). Çalışma kapsamında bu varsayımın test edilmesi için açılımlayıcı faktör analizinden sonra, artık korelasyon matrisi oluşturulmuş ve değişkenler arasındaki ikili korelasyonlar incelenmiştir. İkili çiftlere ait korelasyon katsayılarının -0.00 ile 0.13 arasında olduğu bulunmuştur (İkili çiftlerin 107'inde (%26) korelasyon katsayısı mutlak olarak 0.05'ten büyük çıkmıştır). Artık değerlere ilişkin korelasyon katsayılarının sıfıra yakın çıkması ve tek boyutluluk varsayımının sağlanmasından dolayı yerel bağımsızlık varsayımının da karşılandığı kabul edilmiştir.

Parametrik Olmayan Madde Tepki Kuramı Modellerine İlişkin Varsayımların Test Edilmesi

Parametrik olmayan madde tepki kuramı için var olan varsayımlar PMTK ile benzer de olsa hem kullanılan yöntemlerde hem de model veri uyumun belirlenmesinde farklar bulunmaktadır. Bu yaklaşım için tek boyutluluk ve yerel bağımsızlık varsayımlarına ek olarak, madde tepki fonksiyonların monoton azalmayan nitelikte olduğu ve madde tepki fonksiyonlarının kesişmemesi, bir diğer adıyla değişmez madde sıralaması varsayımı da test edilmektedir. Bu yaklaşım çerçevesinde yapılan varsayımların analizi ile modelin belirlenmesi ve parametrelerin elde edilmesi aynı anda gerçekleşmektedir. Daha açık bir ifadeyle, varsayımlar verinin hangi POMTK modeline uygun olduğunu belirlemekte ve varsayımlar test edilirken madde parametreleri de aynı anda kestirilmektedir. Bundan dolayı bu kısımda söz konusu varsayımların test edilmesinde kullanılan yöntemler tanıtılmış ancak her veri seti için varsayımın analiz edilmesi ve elde edilen sonuçlar çalışmanın bulgular kısmında verilmiştir.

Tek Boyutluluk

Parametrik olmayan madde tepki kuramı yaklaşımında, tek boyutluluk yetenek parametresinin yalnızca maddelere verilen yanıtlara göre kestirildiği anlamına gelmektedir (van der Ark, Croon ve Sijtsma, 2008). Tek boyutluluk maddelere verilen yanıtları

yorumlamayı kolaylaştırdığı için ölçmelerde istenen bir özelliktir (Wismeijer, Sijtsma, van Assen ve Vingerhoets, 2008). Tanım olarak PMTK yaklaşımı ile aynı nitelikte olsa da, POMTK kapsamında tek boyutluluk Otomatik Madde Seçim İşlemi (OMSİ) ve bu analiz içinde yer alan Generic Algorithm (GA) ile test edilmektedir. Parametrik olmayan madde tepki kuramı modelleri kapsamında incelenen tek boyutluluk varsayımı ve bu kapsamda hesaplanan OMSİ ile GA algoritmaları, R programında “mokken” paketi aracılığı ile yapılmaktadır. Bu çalışma kapsamında da POMTK modelleri için yapılan tek boyutluluk analizinde R programı ve “mokken” paketinden yararlanılmıştır.

Yerel Bağımsızlık

Mokken ölçekleme tekniği içinde yer alan yerel bağımsızlık varsayımı, PMTK'daki yerel bağımsızlık varsayımı ile aynıdır. Bireyin maddelere verdiği yanıtların yalnızca bireyin yeteneği tarafından yönlendiriliyor olmasını ve bir maddeye verilen yanıtın başka bir maddeye verilecek yanıtı etkilememesi anlamına gelmektedir (Hambleton ve Swaminathan, 1985). POMTK kapsamında yerel bağımsızlık varsayımı, PMTK'ya göre daha esnek bir şekilde yer almaktadır ve Stout (1990) tarafından tanımlanan temel tek boyutlu modeller ve bu modellerin varsayımı olan temel yerel bağımsızlık varsayımı incelenmektedir. Temel yerel bağımsızlık, sonsuz sayıda madde için hesaplanan maddeler arası kovaryansların ortalamalarının belli bir yetenek düzeyi için sifıra eşit olması olarak tanımlanmaktadır. Koşullu yerel bağımsızlık olarak da adlandırılan bu yaklaşım, bireylerin yetenek ölçeği boyunca sıralanması için yeterli olmaktadır. POMTK kapsamında bu varsayım için maddeler arasındaki koşullu kovaryanslar incelenir ve varsayımın veri seti için karşılanması için bu değerlerin pozitif olması gerekmektedir (Sijtsma ve Meijer, 2007). Bu çalışmada yapılan POMTK analizlerinde yerel bağımsızlık varsayımı için koşullu kovaryanslar incelenmesi yapılmış ve elde edilen sonuçlar çalışmanın bulgular kısmında sunulmuştur.

Madde Tepki Fonksiyonlarının Monotonluğu

Mokken ölçekleme tekniği içinde madde tepki fonksiyonlarının şekline ilişkin incelenen ilk varsayım madde tepki fonksiyonlarının monoton azalmayan nitelikte olmasıdır. Varsayımlardan biri olan tek boyutluluğun incelenmesi sırasında ölçeklenebilirlik katsayıları için sınır değer olan c , 0.3'ten büyük olarak alındığında, seçilen

maddeler büyük olasılıkla monotonluk varsayımını karşılama eğiliminde olacaktır ancak c değerini 0.3'ten daha düşük belirleyerek, ölçüğe maddeler seçildiğinde, monotonluktan ciddi düzeyde ihlaller gösteren maddeler de ölçekte yer alabilir. Ayrıca bu varsayımı karşılamayan maddelerin çoğu c değerine göre ölçek dışında tutulsa da, varsayımı ihlal eden tüm maddeler c değerine göre belirlenemeyebilir. Bu nedenle, MTF'lerin monotonluğu varsayımı, veri setlerinin POMTK modellerine uyumu incelenirken araştırılması gereken bir varsayımdır (Wismeijer, Sijtsma, Van Assen ve Vingerhoets, 2008).

Madde tepki fonksiyonlarının monoton azalmayan nitelikte olması varsayımı POMTK kapsamında çeşitli yöntemlerle incelenebilir. Bu yöntemlerden ilki MTF'lerin kestirilmesi ve incelenmesidir. Parametrik olmayan regresyon yöntemlerinden olan Kernel düzgünleştirme tekniği ile maddelere ilişkin MTF'ler kestirilir ve monoton azalmayan nitelikte olup olmadığı incelenebilir. Bu yöntemi ücretsiz bir program olan TESTGRAF'da uygulamak mümkündür. Ayrıca R'da yer alan "KernSmoothIrt" paketi de Kernel düzgünleştirmesini kullanarak MTF kestirmeyi sağlamaktadır (Meijer, Tenderio ve Wanders, 2015).

Kestirilen MTF'lerin görsel olarak incelenmesi monotonluk hakkında bilgi verse de, bu inceleme ile var olan ihlallerin ne düzeyde olduğunu belirlemek her zaman mümkün olmayabilir. Bu durumda R'da yer alan "mokken" paketinde bulunan madde kalan puanı yöntemi uygulanarak monotonluk varsayımından olan ihlaller ve bu ihlallerin ciddiyetlerinin derecesi belirlenebilir. Madde kalan puanı yönteminde benzer toplam puana ya da kalan puana sahip yanıtlayıcılardan gruplar oluşturulur. Kalan puanlar hesaplanırken, araştırılan maddeye ilişkin bireyin puanı bireyin toplam puanından çıkarılır. Her grup için maddenin doğru cevaplanma oranı belirlenir ve bu süreç tüm maddeler için tekrarlanır. Elde edilen oranlar toplam puanın bir fonksiyonu olarak belirlenir ve MTF'ler oluşturulur. Toplam puanın yerine kalan puanın kullanılma nedeni ise, toplam puanın, özellikle çok kategorili puanlanan maddelere ilişkin oluşturulan madde adım fonksiyonlarının, yetenek ölçüğünde monotonluk göstermesinin gerekli olmamasıdır. Bu nedenle incelemeler kalan puan üzerinden yapılır (Sijtsma ve Meijer, 2007). Bu çalışma kapsamında yapılan monotonluk incelemelerinde R programı ve bu programda yer alan "mokken" paketi

kullanmıştır. Yöntem olarak ise, kalan puanı grupları oluşturulmuş ve monotonluk varsayımını ihlal eden maddeler ve söz konusu ihlallerin düzeyleri belirlenmiştir.

Madde Tepki Fonksiyonlarının Kesişmemesi-Değişmez Madde Sıralaması

Mokken ölçekleme tekniği içinde yer alan Monoton Homojenlik Modeline dayalı ölçekleme yapmak için bu aşamaya kadar incelenen varsayımlar yeterlidir ve bu varsayımların sağlandığı veri setlerinde bireyleri toplam puanlarına göre sıralamak olanaklıdır. Ancak MHM'ye göre daha kısıtlayıcı bir model olan Çift Monotonluk Modeline (ÇMM) göre ölçekleme yapabilmek için MTF'lerin monotonluğa ek olarak, kesişmeyen nitelikte olması gerekmektedir. PMTK kapsamında Rasch modelde sağlanan bu varsayım, POMTK'da, ÇMM'de sağlanmakta ve madde güçlük düzeylerinin tüm yetenek düzeylerinde aynı sıralamaya sahip olması anlamına gelmektedir. Madde tepki fonksiyonlarının kesişmemesi tek boyutlu bir testten elde edilen verilerin yorumlanmasını ve analizini birçok durumda kolaylaştırmaktadır. Bu varsayımın test edilmesinde çeşitli yöntemler bulunmaktadır ve bu yöntemler üç kategoride incelenmektedir. İlk yöntemde maddelerin marjinal dağılımlarının özellikleri artan ve azalan yönde çaprazlanarak incelenir. Maddelerin MTF'leri test kalan eğrilerine göre incelenir. Bu yöntemle yapılan incelemelerde, değişmez madde sıralamasının karşılandığı durumlarda, farklı kalan puanı grupları için hesaplanan madde yanıtlanma oranları sıralamasının tüm grup için hesaplanan madde yanıtlanma oranları ile aynı olması beklenir. Oluşturulan kalan puanı fonksiyonları tüm madde çiftleri için karşılaştırılır ve doğru yanıtlanma oranları tüm grup için yapılan sıralamadan farklılık gösterdiğinde ihlaller ortaya çıkar (Sijtsma ve Junker, 1996). Bazı durumlarda manidar ihlallerin uygulamada düşük etkisi olduğu göz önünde bulundurularak, Molenaar ve Sijtsma (2000) tarafından ihlallerin etki büyüklüğünün belirlenmesini sağlayan bir ölçüt geliştirilmiştir. Crit olarak adlandırılan bu ölçüte göre manidar olan ihlaller şu sınır değerlere göre belirlenmektedir: 40'tan düşük olan değerler önemsiz ihlalleri, 40 ile 80 arasındaki değerler düşük ihlalleri ve 80'in üzerindeki değerler ise ciddi ihlalleri belirtmektedir.

Değişmez madde sıralaması varsayımının test edildiği bir diğer yöntem ise P matrisi yöntemidir. Bu yöntemle yapılan değişmez madde sıralaması incelemesinde, maddelerin zordan kolaya doğru sıralandığı iki simetrik matris kullanılır. İlk matristeki hücreler, P(+

+) , her iki maddeyi de doğru yanıtlayan bireylerin yanıt örüntülerinden oluşurken, ikinci matristeki hücreler, P(- -) her iki maddeyi de yanlış yanıtlayan bireylerin yanıt örüntülerinden oluşur. Madde tepki fonksiyonlarının kesişmemesi durumu bu matriste şu şekilde ifade edilmektedir: Madde tepki fonksiyonların kesişmediği durumda P(+ +) matrisindeki satır ve sütunlar azalmayan, P(- -) matrisindeki satır ve sütunlar ise artmayan özellik göstermektedir. Kesişmeyen madde tepki fonksiyonları varsayımından olan ihlaller ise, P(+ +) matrisinin hücrelerinde azalma, P(- -) matrisinin hücrelerinde artma olduğu durumlarda ortaya çıkar. Madde kalan puanı yönteminde olduğu gibi, varsayımdan olan ihlallerin manidarlığı test edilir ve manidarlık testi McNemar analizi ile yapılır (Meijer, Tenderio ve Wanders, 2015).

Değişmez madde sıralaması varsayımının test edildiği son yöntem, söz konusu varsayımdan olan ihlallerin düzeyini belirten H^T katsayının hesaplanmasıdır. Sijtsma ve Meijer (1992) tarafından önerilen bu katsayı, madde sıralamasının doğruluğu hakkında bilgi vermektedir. Maksimum değeri 1 bu katsayı arttıkça, MTF'lerin birbirinden uzak olduğu, azaldıkça MTF'ler arasındaki mesafelerin az olduğu anlamına gelmektedir. Teorik olan H^T katsayının 0 ile 1 arasında olması, değişmez madde sıralaması varsayımının karşılandığı anlamına gelse de, pratikte bu katsayının en az 0.3 ve üzeri olması beklenmektedir. Ayrıca bu katsayı veri setinde yer alan tüm maddelere ilişkin olarak hesaplanan ortak bir değerdir ve madde düzeyinde bilgi vermez. Bu nedenle Sijtsma ve Meijer (1992), bu katsayının farklı değişmez madde sıralaması inceleme yöntemlerinden elde edilen bilgiler ile birlikte yorumlanmasını önermektedirler. Bu çalışma kapsamında yapılan POMTK analizlerinde, değişmez madde sıralaması varsayımı için, madde kalan puanı grubu ve P matrisi yöntemleri kullanılmış ve her veri seti için H^T katsayısı hesaplanarak her iki yöntemden elde edilen bilgiler birlikte yorumlanmıştır.

BÖLÜM 4

BULGULAR VE YORUMLAR

Bu bölümde, araştırmanın amaçları doğrultusunda yapılan istatistiksel analizlerden elde edilen bulgular ve bulgulara ilişkin yorumlar yer almaktadır. Elde edilen bulgular araştırma sorularının sırasına göre aşağıda sunulmuştur.

Farklı Örneklem Büyüklüğüne ve Test Uzunluğuna Sahip Veri Setlerinin PMTK ve POMTK Kapsamında Yer Alan Modellere Uyumunun Belirlenmesi

Çalışma kapsamında yanıt aranan ilk araştırma sorusu incelenen örneklem büyüklüğü ve test uzunluğu koşullarına göre oluşturulan veri setlerinin PMTK ve POMTK'ya göre model veri uyumunun incelenmesi ve uyumlu oldukları modellerin belirlenmesidir. Bu amaç çerçevesinde oluşturulan veri setlerinin öncelikli olarak PMTK'ya, ardından da POMTK'ya uyumları incelenmiş ve elde edilen sonuçlar sıralı olarak verilmiştir.

Farklı Örneklem Büyüklüklerine Göre Oluşturulan Veri Setlerinin Parametrik Madde Tepki Kuramı Modellerine Uyumunun İncelenmesi

Çalışma kapsamında incelemeye alınan ilk faktör örneklem büyüklüğü faktörüdür. Bu faktör altında yer alan üç koşul olan 500, 1000 ve 3000 kişiden oluşan veri setleri için PMTK kapsamında model veri uyumu incelemeleri yapılmıştır ve elde edilen bulgular bu kısımda sunulmuştur.

500, 1000 ve 3000 Kişilik Veri Setlerinin PMTK Modellerine Uyumunun İncelenmesi

Örneklem büyüklüğü faktörü içinde incelenen koşullar olan 500, 1000 ve 3000 kişilik veri setleri, çalışmanın kapsamında oluşturulan 7242 kişilik temel veri setinden seçkisiz olarak belirlenmiştir. Oluşturulan 500, 1000 ve 3000 kişilik veri setlerinde yer

alan bireylerin toplam puanlarına ilişkin hesaplanan betimsel istatistikler Çizelge 5'te verilmiştir.

Çizelge 5.

500, 1000 ve 3000 Kişilik Veri Setlerinden Hesaplanan Test Puanlarına İlişkin Betimsel İstatistikler

Örneklem Büyüklüğü	500	1000	3000
Test İstatistikleri	Değerler		
Aritmetik Ortalama	15,31	15,46	15,56
Medyan	14,00	15,00	15,00
Mod	11,00	9,00	12,00
Standart Sapma	7,50	7,70	7,49
Varyans	56,35	59,41	56,11
Çarpıklık	,241	,168	,160
Çarpıklığın SH	,107	,075	,044
Basıklık	-1,015	-1,02	-1,01
Basıklığın SH	,214	,150	,089
Ranj	30.00	30.00	30.00
Minimum Değer	0	0	0
Maksimum Değer	30.00	30.00	30.00

Farklı örneklem büyüklüklerinden oluşan veri setleri için hesaplanan test istatistikleri incelendiğinde, genel olarak benzer sonuçlar elde edildiği Çizelge 5'te görülmektedir. Teste ilişkin aritmetik ortalamaların yaklaşık olarak 15 olduğu ve ortanca değerlerinin de 1000 ve 3000 kişilik veri setinde aritmetik ortalamaya çok yakın değerler aldığı, 500 kişilik veri setinde ise bu değerden bir puan düşük olduğu bulunmuştur. Dağılımların tepe değeri değişiklik göstermekte ve en yüksek modun 3000 kişilik veri setinde 12 olarak bulunduğu görülmektedir. Dağılımlara ilişkin merkezi eğilim ölçüleri genel olarak incelendiğinde ise, 3000 kişilik veri setinde diğer örneklem büyüklüklerine göre birbirine daha yakın değerler elde edildiği görülmekte ve bu durum da örnekleme yer alan birey sayısının artmasıyla normal dağılıma daha uygun olduğu sonuçların elde edildiğini göstermektedir. Ayrıca örneklemelerin standart sapma değerleri de birbirine çok yakındır ve her örnekleme testte yer alan tüm maddeleri doğru yanıtlayan bireyler olduğu gibi hiçbir maddeye doğru yanıt veremeyen bireyler de yer almaktadır. Dolayısıyla tüm örneklemelere ilişkin ranjlar eşittir. Son olarak örneklemelerin çarpıklık ve basıklık katsayılarının da ± 1.00 aralığından fazla sapma göstermediği ve her örnekleme test puanlarının normal dağılıma uygun olduğu görülmektedir (Büyüköztürk, Çokluk ve Köklü,

2012). Özetle Çizelge 5'te yer alan değerler incelendiğinde, söz konusu araştırma sorusunu yanıtlamak için seçkisiz olarak oluşturulan örneklemelerin birbirine benzer ve normal dağılımlar gösterdiği sonucuna ulaşılmaktadır.

Oluşturulan veri setleri için PMTK ve POMTK modellerinin uyumu incelenerek madde parametreleri ve birey yetenek kestirimleri yapılmıştır. Buna göre elde edilen ilk bulgu, seçkisiz olarak oluşturulan 500, 1000 ve 3000 kişilik veri setlerinin PMTK modellerine uyumuna ilişkin hesaplanan -2 loglikelihood ve madde ki-kare değerleridir. PMTK kapsamında incelenen 1, 2 ve 3 Parametrelili Lojistik modele ilişkin elde edilen model-veri uyumu bulguları Çizelge 6'da verilmiştir

Çizelge 6.

Seçkisiz Olarak Oluşturulan 500, 100 ve 3000 Kişilik Veri Setlerinin PMTK Modellerine Uyumu

Model Karşılaştırması	500			1000			3000		
	-2LL	sd	p	-2LL	sd	p	-2LL	sd	P
1 PLM-2PLM	100.33	29	<0.001	167.65	29	<0.001	510.98	29	<0.001
2 PLM-3 PLM	57.64	30	>0.001	56.49	30	>0.001	214.18	30	<0.001
1 PLM- 3PLM	62.49	59	<0.001	142.14	59	<0.001	725.16	59	<0.001

Çizelge 6'da verilen model veri uyumu değerleri incelendiğinde, 500 ve 1000 kişilik veri setlerinde, model düzeyinde en yüksek iyileşmenin 2 PLM'de sağlandığı ve bu örneklem için 3 PLM'den yapılan kestirimlerin 2 PLM'den yapılan kestirimlere göre manidar bir iyileşme sağlamadığı görülmektedir. Diğer örneklem büyüklüğü olan 3000 kişilik veri seti için ise modeller arası hesaplanan tüm -2 loglikelihood değerlerinin manidar değişime yol açtığı ancak en yüksek değişimin 3 PLM'de sağlandığı görülmektedir. Dolayısıyla 500 ve 1000 kişilik veri setleri için 2 PLM'nin, 3000 kişilik veri seti için ise 3PLM'nin uyumlu olduğu sonucuna ulaşılmıştır. Model düzeyinde yapılan analizlerin ardından madde düzeyinde ki-kare değerleri incelenerek madde model uyumuna bakılmıştır. Elde edilen sonuçlar Çizelge 7'de verilmiştir.

Çizelge 7.

Seçkisiz Olarak Oluşturulan 500 100 ve 3000 Kişilik Veri Setlerinin PMTK Modellerine Madde Düzeyinde Uyumu

Örneklem Büyükülüğü	1 PLM	2 PLM	3 PLM
	Uyumlu Madde Sayısı	Uyumlu Madde Sayısı	Uyumlu Madde Sayısı
500	14	25	24
1000	14	24	23
3000	4	14	10

Çizelge 7’de yer alan veri setlerinin PMTK modellerine göre uyumlu olduğu belirlenen madde sayıları incelendiğinde, 500 kişilik veri seti için, en yüksek model veri uyumun 25 madde ile 2 PLM için sağlandığı belirlenmiştir. İki parametrelili lojistik modele beş maddenin uyum sağlamadığı bu veri seti için, üç parametrelili lojistik modelin de benzer düzeyde uyumlu olduğu belirlenmiştir. Bu modellere ilişkin olarak hesaplanan ve Çizelge 6’da yer alan -2 loglikelihood değerleri farkının manidar olmadığı göz önünde bulundurularak, kestirimlerin 2 PLM ile yapılmasına karar verilmiştir. Diğer örneklem büyüklüğü olan 1000 kişiden oluşan veri seti için uyumlu madde sayısı 500 kişilik veri setiyle benzerlik göstermektedir ve en iyi model veri uyumu 30 maddeden 24’ünün uyum sağladığı 2 PLM için sağlanmıştır. Bu veri seti için 2 PLM’den sonra model veri uyumunun en yüksek düzeyde sağlandığı model 3 PLM’dir ve 23 madde bu modele uyum sağlamıştır. Bu veri seti için uygun olan modelin belirlenmesi amacıyla, iki ve üç parametrelili lojistik model için hesaplanan -2 loglikelihood değerleri incelenmiş ve değerler farkının manidar olmadığı belirlenerek, 2 PLM ile analizlere devam edilmiştir.

İncelemeye alınan son örneklem büyüklüğü olan 3000 kişilik veri seti için elde edilen bulgular incelendiğinde ise, madde düzeyindeki model veri uyumunun daha küçük örneklem büyüklüğünde daha zayıf olduğu görülmektedir. Bu durumun nedeni ise, maddelerin uyumunu belirten ki-kare değerlerinin örneklem büyüklüğünden etkilenmesi ve küçük farkların dahi manidar çıkmasıdır. Bu nedenle bu örneklem büyüklüğü için öncelikli olarak -2 loglikelihood değerleri incelenmiş ve 2 PLM ile 3 PLM arasındaki farkın manidar olduğu belirlenmiştir. Uyumlu madde sayıları incelendiğinde ise 2 PLM’ye daha fazla uyumlu madde olduğu bulunmuştur. Üç bin kişiden

oluşan veri setinin uyumlu olduğu modelin belirlenmesi adına, Demars (2010) tarafından belirtildiği gibi, model özelliklerine göre incelemeler yapılmış ve 3 PLM kapsamında 2 PLM'den farklı olarak hesaplanan c parametreleri kestirilmiştir. Veri setinde bulunan maddeler için kestirilen şans parametreleri incelendiğinde, 12 madde için kestirilen c değerlerinin 0 olduğu belirlenmiştir. Kestirilen c parametrelerinin ortalaması ise 0.110 olarak hesaplanmış ve hesaplanan en yüksek c parametresinin 0.282 olduğu belirlenmiştir. Maddelerin yarısı için c parametresinin 0'a eşit olması ve bu parametrenin ortalamasının oldukça düşük olması nedeniyle veri setinin 2 PLM'ye daha uygun olduğu sonucuna ulaşılmış ve kestirimler bu modele göre yapılmıştır. Farklı örneklem büyüklüklerinden oluşan veri setlerinin PMTK kapsamında yapılan model veri uyumu incelemeleri sonucunda uyumlu olduğu belirlenen modeller ile madde parametreleri kestirimleri yapılmıştır.

Farklı Örneklem Büyüklüklerine Göre Oluşturulan Veri Setlerinin Parametrik Olmayan Madde Tepki Kuramı Modellerine Uyumunun İncelenmesi

Bu araştırma kapsamında yanıt aranan ilk soruya cevap oluşturmak için farklı örneklem büyüklüğü koşullarına göre oluşturulan veri setlerinin POMTK modellerine uyumu incelenmiş ve veri setlerinin uyum sağladığı modeller belirlenmiştir. POMTK kapsamında model veri uyumu incelemeleri yapılırken, alan yazında belirtilen adımlar izlenmiş ve veri setlerin Çift Monotonluk Modeline uyumu incelenmiş ve elde edilen bulgular analizde izlenmesi gereken adımların sıralaması gözetilerek sunulmuştur.

Madde güçlüklerinin hesaplanması.

POMTK kapsamında yapılan ilk analiz madde güçlüklerinin belirlenmesidir. Madde güçlükleri KTK'da olduğu gibi, maddeyi doğru yanıtlayan birey sayısı/tüm birey sayısı formülünden hesaplanmış ve örneklemelere göre kestirilen madde güçlük indeksleri Çizelge 8'de sunulmuştur.

Çizelge 8.

500, 1000 ve 3000 Kişilik Veri Setlerinin POMTK'ya Göre Hesaplanan Madde Güçlük Parametreleri

Örneklem				Örneklem			
Maddeler	500	1000	3000	Maddeler	500	1000	3000
1	0.71	0.72	0.72	16	0.77	0.73	0.75
2	0.38	0.38	0.39	17	0.32	0.27	0.29
3	0.69	0.71	0.70	18	0.50	0.47	0.46
4	0.69	0.67	0.70	19	0.59	0.56	0.56
5	0.56	0.56	0.55	20	0.51	0.45	0.48
6	0.68	0.71	0.70	21	0.66	0.62	0.62
7	0.61	0.65	0.65	22	0.50	0.46	0.47
8	0.62	0.61	0.63	23	0.22	0.18	0.18
9	0.27	0.27	0.25	24	0.31	0.28	0.27
10	0.62	0.65	0.66	25	0.38	0.31	0.29
11	0.38	0.37	0.36	26	0.47	0.43	0.44
12	0.70	0.69	0.68	27	0.63	0.56	0.57
13	0.54	0.49	0.50	28	0.74	0.68	0.69
14	0.41	0.38	0.40	29	0.43	0.43	0.41
15	0.60	0.57	0.59	30	0.56	0.52	0.51
Ortalama					0.539	0.515	0.518

Çizelge 8'de farklı örneklem büyüklüklerinden oluşan veri setlerinden POMTK'ya göre hesaplanan madde güçlük indeksleri yer almaktadır. Elde edilen değerler örneklem büyüklüklerine göre incelenecek olursa, maddelerin her üç veri setinde de benzer değerler aldığı görülmektedir. Ortalama değerler incelendiğinde ise, en yüksek ortalamanın 500 kişilik veri setinden elde edildiği, 1000 ve 3000 kişilik verilerden kestirilen parametrelere ilişkin hesaplanan ortalamaların ise birbirine yakın olduğu görülmektedir. Madde düzeyinde yapılan incelemelerde ise, 1000 ve 3000 kişilik veri setlerinden kestirilen madde güçlük parametrelerinin 500 kişilik veri setinden kestirilen parametrelere göre birbirlerine daha benzer olduğu görülmektedir. Bu duruma dayanarak ise, örnekleme yer alan birey sayısındaki artışın madde güçlük parametrelerini daha tutarlı hale getirdiği yorumuna ulaşılabilir (Crocker ve Algina, 1986). Madde güçlük parametrelerinin hesaplanmasının ardından, maddelere ilişkin bir diğer psikometrik nitelik olan ve POMTK kapsamında ölçeklenebilirlik katsayısı olarak adlandırılan H_i değerlerinin, bir diğer ifade ile madde ayırt edicilik indekslerinin hesaplanmasına geçilmiştir.

Ölçeklenebilirlik katsayılarının hesaplanması.

Herhangi bir veri setinin POMTK kapsamında yer alan modellerle ölçeklenip ölçeklenemeyeceğini belirlemek için yapılacak olan ilk iş, H_{ij} katsayıları olarak bilenen maddeler arasındaki kovaryansların incelenmesidir. Maddeler arasındaki ikili kovaryansların incelenmesiyle başlayan süreç, aslında POMTK modellerinin madde analizinin ilk aşamasıdır. POMTK kapsamında yer alan her iki model de, madde çiftleri arasındaki kovaryansların pozitif olmasını gerektirir ve madde çiftleri için kovaryanslar hesaplanarak bu değerler incelenir. Madde çiftleri arasındaki kovaryansların incelenmesinin ardından maddenin testin bütünüyle olan kovaryansı H_i incelenir ve bu değer için 0.3 sınır kabul edilir. Son olarak da ölçeğin bütünü için hesaplanan ölçeklenebilirlik değeri, H , incelenir ve ölçeğin Mokken ölçekleme tekniğine uygun olup olmadığı belirlenir (Van Schuur, 2011; Sijtsma ve Molenaar, 2002). Bu çalışmada da oluşturulan 500, 1000 ve 3000 kişilik veri setlerinin POMTK modellerine uyumu inceleneceği için ilk olarak madde çiftleri arasındaki kovaryanslar H_{ij} hesaplanmıştır. Ardından maddeler için H_i katsayıları en son olarak ise ölçeğin bütünü için H katsayısı incelenmiştir. Elde edilen bulgular örneklem büyüklüklerinin sırasıyla sunulmuştur.

1. Madde çiftleri için hesaplanan ölçeklenebilirlik katsayıları

Tüm veri setleri için yapılan ilk analiz madde çiftleri arasındaki kovaryansların incelenmesidir. Toplam 30 maddeden oluşan veri setleri için, madde çiftleri temel alındığında 435 kovaryans hesaplanmıştır. Beş yüz kişilik veri setinden hesaplanan madde çiftleri arasındaki kovaryans değerlerinden, yalnızca biri, madde 1 ile madde 12 arasındaki kovaryans negatif olarak bulunmuş, diğer tüm değerlerin pozitif olduğu belirlenmiştir. Dolayısıyla bu veri seti için söz konusu iki maddenin ölçekten çıkmaya aday maddeler olduğu belirlenmiştir. Bin kişilik veri seti için hesaplanan ikili kovaryanslar incelendiğinde ise, 435 kovaryans değerinden hiçbirinin negatif olmadığı, tüm maddelerin birbirleri ile pozitif kovaryans değerine sahip olduğu belirlenmiştir. İncelemeye alınan son örneklem büyüklüğü olan 3000 kişilik veri seti için yapılan madde çiftleri arasındaki kovaryans incelemesinde ise, negatif bir kovaryans değerine rastlanmamıştır. Sonuç olarak 500, 1000 ve 3000 kişilik veri setlerinde yer alan maddelerden yalnızca 500 kişilik veri setinde yer alan 1 ve 12 nolu maddelerin negatif kovaryans değerine sahip olduğu ve dolayısıyla bu

veri seti için bu maddelerin ölçeklenemeyeceği fakat diğer iki veri setinde yer alan tüm maddelerin POMTK modelleri kapsamında ölçeklemeye uygun olduğu bulunmuştur.

2. Madde ve ölçek düzeyinde hesaplanan ölçeklenebilirlik katsayısı

Madde çiftleri için hesaplanan ölçeklenebilirlik katsayısı ölçek için genel bilgiler vermekte ve maddelerin aynı özelliği ölçüp ölçmediğini belirtmektedir. Madde düzeyinde hesaplanan ölçeklenebilirlik katsayısı ise maddelerin gizil değişkeni ölçme durumlarına ilişkin bilgi vermektedir. Madde düzeyinde hesaplanan ölçeklenebilirlik katsayısı H_i , katsayısı teoride negatif değerler alabilse de, POMTK modelleri kapsamında yapılan ölçeklemede bu değerlerin 0.3'e eşit ve büyük olması beklenmektedir. 0.3 ile 0.4 arasında değişen H_i değerleri maddenin pozitif fakat düşük ayırt edicilikte olduğu, 0.4 ile 0.5 arasında değişen H_i değerleri maddenin orta düzeyde ayırt ediciliğe sahip olduğu, 0.5 ve üzeri değerler ise maddenin yüksek düzeyde ayırt edici nitelikte olduğu anlamına gelmektedir. Hesaplanan H_i değerlerinin sıfırdan manidar olarak farklı olup olmadığını belirlemek için Z istatistiği hesaplanır ve örneklemden elde edilen H_i değerinin evrende de manidar olup olmadığı incelenir (Loevinger, 1948; Molenaar and Sijtsma, 2000). Bu bilgiler doğrultusunda her örneklem büyüklüğü için hesaplanan H_i değerleri, bu değerlere ilişkin Z değerleri ve tüm ölçek düzeyi için hesaplanan H değerleri Çizelge 9'da verilmiştir.

Çizelge 9.

Madde ve Ölçek Düzeyinde Hesaplanan Ölçeklenebilirlik Katsayıları

Örneklemler	500			1000			3000		
	H_i	Sh	Z	H_i	Sh	Z	H_i	Sh	Z
M1	0.262	(0.027)	21.47	0.353	(0.019)	40.37	0.298	(0.012)	58.31
M2	0.419	(0.027)	36.64	0.406	(0.020)	50.18	0.402	(0.011)	84.03
M3	0.303	(0.027)	25.94	0.359	(0.018)	42.68	0.315	(0.012)	63.90
M4	0.419	(0.021)	35.82	0.429	(0.015)	53.77	0.427	(0.009)	88.07
M5	0.360	(0.024)	34.01	0.348	(0.018)	46.72	0.323	(0.011)	73.01
M6	0.348	(0.024)	30.07	0.372	(0.019)	44.04	0.358	(0.011)	73.04
M7	0.330	(0.024)	30.67	0.324	(0.018)	41.57	0.316	(0.011)	68.77
M8	0.365	(0.023)	33.84	0.401	(0.017)	52.76	0.393	(0.010)	86.45
M9	0.550	(0.029)	39.32	0.557	(0.019)	58.38	0.555	(0.012)	92.93
M10	0.276	(0.027)	25.56	0.336	(0.019)	43.03	0.291	(0.011)	62.72
M11	0.310	(0.032)	26.54	0.380	(0.021)	46.46	0.365	(0.013)	73.85

(Devam ediyor)

Çizelge 9.(devam)

Madde ve Ölçek Düzeyinde Hesaplanan Ölçeklenebilirlik Katsayıları

Örneklem	500			1000			3000		
	H _i	Sh	Z	H _i	Sh	Z	H _i	Sh	Z
M12	0.215	(0.031)	18.44	0.235	(0.022)	28.63	0.196	(0.013)	41.15
M13	0.392	(0.023)	36.95	0.379	(0.018)	50.60	0.377	(0.010)	85.05
M14	0.397	(0.026)	34.79	0.424	(0.019)	52.55	0.393	(0.011)	82.96
M15	0.319	(0.025)	30.12	0.341	(0.018)	45.72	0.341	(0.010)	76.33
M16	0.387	(0.029)	28.7	0.395	(0.020)	44.46	0.398	(0.012)	72.54
M17	0.481	(0.031)	36.38	0.521	(0.022)	57.67	0.480	(0.012)	88.12
M18	0.371	(0.025)	34.59	0.386	(0.018)	51.27	0.373	(0.011)	82.77
M19	0.393	(0.022)	37.10	0.416	(0.016)	55.89	0.381	(0.010)	85.98
M20	0.297	(0.028)	27.56	0.365	(0.019)	48.13	0.333	(0.011)	74.72
M21	0.338	(0.024)	30.92	0.357	(0.018)	46.58	0.346	(0.010)	76.27
M22	0.319	(0.027)	29.47	0.369	(0.018)	48.83	0.345	(0.011)	77.06
M23	0.636	(0.035)	35.12	0.669	(0.022)	54.75	0.631	(0.013)	87.04
M24	0.325	(0.040)	23.43	0.315	(0.026)	33.82	0.331	(0.016)	58.83
M25	0.237	(0.036)	19.40	0.300	(0.025)	33.89	0.269	(0.015)	49.34
M26	0.332	(0.028)	29.91	0.365	(0.020)	47.25	0.332	(0.012)	72.64
M27	0.342	(0.024)	32.25	0.391	(0.017)	52.43	0.355	(0.010)	80.14
M28	0.326	(0.027)	27.93	0.360	(0.018)	44.38	0.343	(0.011)	70.63
M29	0.416	(0.028)	36.35	0.416	(0.019)	53.54	0.398	(0.011)	85.11
M30	0.345	(0.025)	32.40	0.328	(0.019)	43.95	0.326	(0.011)	73.64
Ölçek İçin H	0.345	(0.015)	118.93	0.381	(0.011)	182.12	0.360	(0.006)	289.74

Çizelge 9’da yer alan değerlerden ilki madde düzeyindeki, ölçeklenebilirlik katsayılarıdır ve farklı örneklem için hesaplanan H_i değerleri incelendiğinde, kritik değer olan 0.3’ten düşük H_i değerine sahip en fazla maddenin 500 kişilik veri setinde bulunduğu görülmektedir. Bu örneklem büyüklüğü için 1, 10, 12, 20 ve 25 nolu maddelerin H_i değerleri kritik değerinin altında yer almaktadır. 1000 kişilik veri setinde ise, yalnızca 25 nolu maddenin H_i değeri 0.3’ten düşüktür. 3000 kişilik veri setindeki maddelerin H_i değerleri incelendiğinde ise, 500 kişilik veri setinden elde edilen sonuçlara benzer H_i değerlerinin kestirildiği görülmektedir. Bu örneklem için dört maddenin kritik değerinin altında olduğu belirlenmiştir ve bu maddeler, 1, 10, 12 ve 25 nolu maddelerdir. Her üç veri setinde de 25 nolu maddenin eşik değerinin altında H_i değerine sahip olduğu

bulunmuş, dolayısıyla bu maddenin POMTK modellerinde ölçeklenemeyeceği sonucuna ulaşılmıştır. İncelenen H_i katsayılarına göre, 1000 kişilik örneklem büyüklüğünün POMTK modellerine en uygun örneklem büyüklüğü olduğu, 500 ve 3000 kişilik veri setlerinden ise benzer sonuçların elde edildiği sonucuna ulaşılabılır. Kuijpers, van der Ark, Croon ve Sijtsma (2016) çalışmalarında H_i katsayılarına bağlı olmak üzere, bu katsayıların düşük olduğu durumlarda dahi veri setinde yer alan birey sayısının 1000-1500 arasında olmasının analizleri tutarlı hale getirdiğini belirtmektedir. Çalışmada yer alan maddelerin H_i katsayılarının da düşük olduğu göz önünde bulundurulduğunda 500 kişilik veri setinin POMTK uygulamaları için küçük sayılabilecek büyüklükte olduğu sonucuna ulaşılmaktadır. Aynı şekilde 3000 kişilik örneklem büyüklüğü de POMTK algoritmaları için oldukça geniş bir örneklem büyüklüğünü yansıttığı için her iki veri stinden kestirilen değerlerin benzer olduğu sonucuna ulaşılabılır.

Çizelge 9'da yer alan bir başka değer ise, H_i katsayılarının evrende sıfırdan manidar olarak farklılaşp farklılaşmadığını inceleyen Z katsayılarıdır. Z değerleri örneklemeden kestirilen H_i değerlerinin standart sapmasıdır ve sıfırdan farklı olması H_i katsayısının evrende de sıfırdan farklı olduğu anlamına gelmektedir (Sijtsma ve Molenaar, 2002). İncelenen örneklemlerin hepsinde H_i değerlerine ilişkin Z istatistiklerinin sıfırdan farklı olduğu bulunmuş, dolayısıyla H_i değerlerinin sıfırdan manidar olarak farklı olduğu sonucuna ulaşılmıştır.

Çizelge 9'da yer alan bir diğer değer ise H_i katsayılarına ilişkin standart hata değerleridir. Bu değerler incelendiğinde, en yüksek standart hata değerlerinin 500 kişilik veri setinde, en düşük standart hata değerlerinin ise 3000 kişilik veri setinde kestirildiği görülmektedir. Örneklem büyüdükçe standart hata miktarında azalma gözlenmektedir ve bu bulgu da alan yazında belirtilen durumlarla uyum göstermektedir (Meijer, Tenderio ve Wanders, 2015; Sijtsma ve Molenaar, 2002; Van Schuur, 2011) Çizelge 9'da yer alan son değer ise ölçek düzeyi için hesaplanan ölçeklenebilirlik katsayısı olan H değerleridir. Elde edilen H değerleri incelendiğinde, en yüksek H değerinin POMTK modellerine en iyi uyum gösteren 1000 kişilik veri setinden kestirildiği, en düşük H değerinin ise uyumsuz madde sayısı en fazla olan 500 kişilik veri setinden elde edildiği görülmektedir. Ölçek düzeyindeki H katsayıları madde düzeyindeki H_i katsayılarına bağlı olarak hesaplandığı için bu bulgu beklenen bir bulgudur. Üç örneklem büyüklüğü için de kestirilen H değerleri genel olarak

incelendiğinde, 0.3 ile 0.4 arasında değiştiği görülmektedir ve bu değerlere göre her üç veri setinin de Mokken ölçeklemeye zayıf düzeyde uyum sağladığı sonucuna ulaşılmıştır. H katsayılarına ilişkin Z değerlerinin de pozitif ve sıfırdan büyük değerler olduğu Çizelge 9’da görülmekte, dolayısıyla bu değerlerin de sıfırdan manidar olarak farklı olduğu sonucuna ulaşılmıştır. Özetle, madde çiftleri, madde ve ölçek düzeyinde hesaplanan H değerleri incelenmiş ve her üç veri setinin de POMTK modelleri kapsamında ölçeklemeye uygun olduğu sonucuna ulaşılmış ve POMTK modellerine olan uyumun incelenmesi için diğer varsayımların analizine geçilmiştir.

Tek boyutluluk incelemesi.

Mokken modelleri olarak da adlandırılan POMTK modellerinde tek boyutluluğun incelenmesi için faktör analizine alternatif olarak Otomatik Madde Seçim İşlemi (OMSİ-Automated Item Selection Procedure-AISP) kullanılmaktadır. Hem çok kategorili hem de iki kategorili olarak puanlanan maddelere uygulanabilen ve maddeler arası kovaryanslara dayalı olarak hesaplanan bu algoritma ile özellikle ikili puanlanan maddelerde faktör analiziyle gözlenemeyen doğrusal ilişkili yapılarını ortaya çıkarmak olası hale gelmektedir (Meijer, Tenderio ve Wanders, 2015). Bu çalışma kapsamında oluşturulan veri seti için genel olarak MTK’nın varsayımlarının incelenmesi sırasında açımlayıcı ve doğrulayıcı faktör analizleri ile incelenen tek boyutluluk varsayımı, bu aşamada Mokken ölçekleme tekniğine uygun olan OMSİ kullanılarak analiz edilmiştir. Otomatik madde seçim işlemi analizi, hem 7242 kişiden oluşan ve tüm çalışma grubunu kapsayan temel veri seti için, hem de örneklem büyüklüğünün kestirimlere olan etkisinin incelenmesi amacıyla seçkisiz olarak oluşturulan 500, 1000 ve 3000 kişilik veri setleri için gerçekleştirilmiştir. Mokken tarafından önerilen 0.3 katsayısının çok katı bir sınır değer olduğuna dair çeşitli eleştirilerin bulunduğu ve bu değer in doğruluğuna ilişkin alan yazında da çok fazla çalışma olmamasından dolayı araştırmacıların analiz sırasında farklı değerleri de denemeleri çeşitli kaynaklarda önerilmektedir (Sijtsma ve Meijer, 2007; Molenaar, 2001; Wismeijer, Sijtsma, Van Assen ve Vingerhoets, 2008). Bu nedenle bu çalışma kapsamında 0.2 sınır değeri de analize alınarak tek boyutluluk incelemesi yapılmış ve her iki durum arasındaki farklılık incelenmiştir ve elde edilen sonuçlar Çizelge 10’da verilmiştir.

Çizelge 10.

Oluşturulan Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi

Örneklem	Tüm Veri Seti		500		1000		3000	
	H \geq 0.3	H \geq 0.2	H \geq 0.3	H \geq 0.2	H \geq 0.3	H \geq 0.2	H \geq 0.3	H \geq 0.2
M1	1	1	0	0	1	1	1	1
M2	1	1	1	1	1	1	1	1
M3	1	1	1	1	1	1	1	1
M4	1	1	1	1	1	1	1	1
M5	1	1	1	1	1	1	1	1
M6	1	1	1	1	1	1	1	1
M7	1	1	1	1	1	1	1	1
M8	1	1	1	1	1	1	1	1
M9	1	1	1	1	1	1	1	1
M10	1	1	0	1	1	1	1	1
M11	1	1	1	1	1	1	1	1
M12	0	1	0	1	0	1	0	0
M13	1	1	1	1	1	1	1	1
M14	1	1	1	1	1	1	1	1
M15	1	1	1	1	1	1	1	1
M16	1	1	1	1	1	1	1	1
M17	1	1	1	1	1	1	1	1
M18	1	1	1	1	1	1	1	1
M19	1	1	1	1	1	1	1	1
M20	1	1	1	1	1	1	0	1
M21	1	1	1	1	1	1	1	1
M22	1	1	1	1	1	1	1	1
M23	1	1	1	1	1	1	1	1
M24	1	1	1	1	1	1	1	1
M25	0	1	0	1	1	1	0	1
M26	1	1	1	1	1	1	1	1
M27	1	1	1	1	1	1	1	1
M28	1	1	1	1	1	1	1	1
M29	1	1	1	1	1	1	1	1
M30	1	1	1	1	1	1	1	1
Uyumsuz Madde Sayısı	2	-	4	1	1	0	3	1

Çizelge 10'da yer alan H \geq 0.3 ve H \geq 0.2 değerleri OMSİ analizinde kullanılan madde ölçeklenebilirlik katsayılarının sınır değerlerini belirtmektedir. Maddelerin ayırt edicilikleri hakkında bilgi veren bu değer, 0.3 yerine 0.2 olarak alındığında, ölçeklenebilirlik katsayısı 0.2'nin üzerinde olan maddeler de ölçeklenebilmekte, Mokken ölçeklemeye uygun olarak kabul edilmektedir. Çizelge 10'da tanımlanması gereken diğer değerler ise maddelerle ilişkilendirilen 0 ve 1 değerleridir. Sıfır değeri maddenin testin bütünüyle ölçülen psikolojik yapıya yeterli katkıyı sağlamadığı ve bu maddenin ölçek dışında bırakılmasının daha uygun olduğu anlamına gelirken, 1 değeri maddenin ölçülen

gizil deęişkeni yeterli ayırt edicilikte ölçtüęü ve teste katkı sağladığı anlamına gelmektedir. Bir deęeri alan maddeler bir araya getirilerek Mokken ölçekleme tekniğine uygun bir ölçek oluşturmak ve bireyleri yetenek düzeylerine göre, maddeleri de güçlük düzeylerine göre sıralamak mümkün olmaktadır. Çizelgede yer alan ve maddelerle ilişkilendirilen 2 deęeri ise maddelerin testle ölçülen özellikten farklı bir özellięi ölçtüęü anlamına gelmekte ve dolayısıyla bu maddelerin ölçekten çıkarılması gerekmektedir.

Verilen bilgiler ve Çizelge 10'da yer alan deęerler göz önünde bulundurularak farklı örneklem büyüklüklerinin tek boyutluluęu incelendiğinde, tüm veri için yapılan tek boyutluluk analizinde, maddeler için ölçeklenebilirlik katsayısının sınır deęeri 0.3 kabul edildiğinde, 12 ve 25 nolu maddelerin ölçeklemeye uygun olmadığı, sınır deęeri 0.2 olarak alındığında ise tüm maddelerin Mokken ölçeklemeye uygun olduęu görülmektedir. 500 kişilik veri setinde 0.3 deęeri sınır deęeri olarak alındığında, 1, 10, 12 ve 25 nolu maddelerin tek boyutluluęa uygun olmadığı, 0.2 ölçeklenebilirlik deęeri için ise yalnızca 1 nolu maddenin tek boyutluluk özellięi göstermedięi belirlenmiştir. 1000 kişilik veri setinde tek boyutluluk özellięinin 500 kişilik veri setine göre daha yüksek düzeyde olduęu Çizelge 10'da görülmektedir. Bu örneklem büyüklüğünde 0.3 ölçeklenebilirlik deęeri sınır olarak alındığında tek bir madde uyumsuzken, 0.2 deęeri için tüm maddeler uyum sağlamaktadır. İncelemeye alınan son örneklem büyüklüğü 3000 kişilik veri setidir ve bu örneklem büyüklüğü için ölçeklenebilirlik katsayısının sınır deęeri 0.3 alındığında, 12, 20 ve 25 nolu maddeler tek boyutluluk özellięi göstermemekte, H_1 katsayısının sınır deęeri 0.2 olarak alındığında ise sadece 12 nolu maddenin tek boyutluluk özellięi göstermedięi belirlenmiştir.

Mokken ölçekleme teknięinin tek boyutluluk analizi olan OMSİ analizi sonuçlarına genel olarak bakıldığında řu sonuçlara ulařılabilir:

- En yüksek uyum 1000 kişilik veri setinde sağlanmaktadır.
- 500 ve 3000 kişilik veri setlerinden benzer sonuçlar elde edilmiştir. Bu durumun nedeni daha önce de belirtildięi üzere her iki örneklem büyüklüğünün de POMTK'ya uygun olmadığı olabilir.
- 12 ve 25 nolu maddeler birden çok veri setinde tek boyutluluęa uyum sağlamamaktadır. Söz konusu maddeler PMTK modellerine uyumun incelendiğinde AFA incelemesinde de faktör yükü sınır olan 0.32'nin altında deęer

veren maddelerdir. Dolayısıyla bu maddelerin ölçeğin tek boyutlu yapısı ile uygun olmadığı sonucuna ulaşılmıştır.

- Ölçeklenebilirlik katsayısı 0.2 olarak alındığında uyumlu madde sayısı artmaktadır.
- 500 kişilik veri setinde diğer veri setlerinden farklı olarak 1 ve 10 nolu maddeler uyumsuz olarak belirlenmiştir.
- 3000 kişilik veri setinde ise diğer veri setlerinden farklı olarak 20 nolu madde uyumsuz olarak belirlenmiştir.

Parametrik olmayan madde tepki kuramı modelleri için tek boyutluluk varsayımının test edilmesinden sonra yerel bağımsızlık varsayımının incelemesine geçilmiştir. Söz konusu varsayım bu çalışma kapsamında oluşturulan veri setleri için daha önce incelenmiş bir varsayımdır ve bu varsayımın POMTK modelleri için farklı bir inceleme yöntemi alan yazında bulunmamaktadır. Bu nedenle POMTK modelleri için daha önceden POMTK modelleri için yapılan yerel bağımsızlık varsayımı incelemesinin sonuçları geçerli kabul edilmiş ve oluşturulan veri setlerinin yerel bağımsızlık varsayımını karşıladığı kabul edilmiştir. Yerel bağımsızlık varsayımından sonra POMTK modellerine ilişkin incelenmesi gereken bir diğer varsayım Madde Tepki Fonksiyonlarının Monotonluğudur ve bu aşamada oluşturulan veri setleri için bu varsayımın incelenmesine geçilmiştir.

Madde tepki fonksiyonlarının monotonluğunun incelenmesi

Bu çalışma kapsamında yer alan veri setleri için bu varsayım madde kalan puanı regresyonu yöntemi ile incelenmiştir. Her örneklem büyüklüğü için incelenen madde kalan puanı regresyonu eğrilerinin monotonluğu ihlal etme durumları ve söz konusu ihlallerden manidar olanlar Çizelge 11’de verilmiştir.

Çizelge 11.

500, 1000 ve 3000 Kişilik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Maddeler	500 kişilik veri seti				1000 kişilik veri seti				3000 kişilik veri seti			
	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal	Z*
M1	0.26	28	2	0	0.35	28	0	0	0.30	28	0	0
M2	0.42	28	0	0	0.41	28	0	0	0.40	28	0	0
M3	0.30	28	0	0	0.36	28	0	0	0.31	28	0	0

(Devam ediyor.)

Çizelge 11. (devam)

500, 1000 ve 3000 Kişilik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Maddeler	500 kişilik veri seti				1000 kişilik veri seti				3000 kişilik veri seti			
	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal	Z*
M4	0.42	28	1	0	0.43	28	0	0	0.43	28	0	0
M5	0.36	36	0	0	0.35	28	0	0	0.32	28	0	0
M6	0.35	28	1	0	0.37	28	0	0	0.36	28	0	0
M7	0.33	36	1	0	0.32	28	0	0	0.32	21	0	0
M8	0.36	28	0	0	0.40	28	0	0	0.39	21	0	0
M9	0.55	15	1	0	0.56	21	0	0	0.55	28	0	0
M10	0.28	28	1	0	0.34	28	0	0	0.29	28	0	0
M11	0.31	28	0	0	0.38	28	0	0	0.36	28	0	0
M12	0.21	28	1	0	0.23	28	0	0	0.20	28	0	0
M13	0.39	36	1	0	0.38	28	0	0	0.38	28	0	0
M14	0.4	28	0	0	0.42	280	1	0	0.39	28	0	0
M15	0.32	36	1	0	0.34	28	0	0	0.34	21	0	0
M16	0.39	28	0	0	0.40	28	0	0	0.40	21	0	0
M17	0.48	36	1	0	0.52	28	0	0	0.48	28	0	0
M18	0.37	28	2	0	0.39	28	0	0	0.37	21	0	0
M19	0.39	21	0	0	0.42	28	0	0	0.38	28	0	0
M20	0.30	36	2	0	0.37	28	1	0	0.33	28	0	0
M21	0.34	36	1	0	0.36	28	0	0	0.35	28	0	0
M22	0.32	28	1	0	0.37	28	0	0	0.34	28	0	0
M23	0.64	17	2	0	0.67	15	0	0	0.63	28	0	0
M24	0.33	28	1	0	0.32	28	1	0	0.33	28	0	0
M25	0.24	36	3	0	0.30	28	1	0	0.27	28	0	0
M26	0.33	28	1	0	0.36	28	0	0	0.33	28	0	0
M27	0.34	28	0	0	0.39	28	0	0	0.36	21	0	0
M28	0.33	36	1	0	0.36	28	0	0	0.34	21	0	0
M29	0.42	36	1	0	0.42	28	0	0	0.40	28	0	0
M30	0.34	28	1	0	0.33	28	1	0	0.33	28	0	0

Çizelge 11’de verilen maddelere ilişkin özellikler şunlardır: H_i katsayısı, daha önce de belirtildiği üzere maddeye ait ölçeklenebilirlik katsayısıdır ve 0.3 bu katsayı için sınır değer olarak kabul edilmektedir (Sijtsma ve Meijer, 2007). “Çift “olarak belirtilen değer

ise, gözlenen monotonluğun incelenmesi için örneklem büyüklüğüne bağlı olarak oluşturulan madde kalan puanı grup sayısıdır. Bir diğer sütunda yer alan ihlal olarak adlandırılan değerler, oluşturulan madde kalan puanı regresyonunun monotonluğu ihlal ettiği noktaların sayısını belirtmekte ve bu sayının manidarlığı ise Z testi sonuçları ile belirtilmektedir. Z testinin manidarlığındaki değerler, monotonluğun ihlalindeki kaç değer manidar olduğunu göstermektedir (Van Schuur, 2011). Bu açıklamalardan sonra her örneklem büyüklüğü için maddelerin monotonluk varsayımını karşılama durumları incelenebilir.

En küçük örneklem büyüklüğü olan 500 kişilik veri seti için, maddelere ilişkin H_i değerleri incelendiğinde, 1, 10, 12 ve 25 nolu maddelerin H_i değerlerinin 0.30'un altında olduğu belirlenmiştir. Bu maddeler bir önceki varsayım olan tek boyutluluğun OMSİ ile incelenmesi sırasında da, 500 kişilik veri seti için tek boyutluluk özelliği göstermeyen maddelerdir. Dolayısıyla bu maddelerin bu veri setinde Mokken ölçekleme yaklaşımına uygun olmadığı ve testten çıkarılması gerektiği sonucuna ulaşılmaktadır. Bu örneklem büyüklüğü için monotonluk varsayımından yetenek ölçeği boyunca gerçekleşen ihlaller incelendiğinde, en fazla ihlalin 25 nolu maddede olduğu görülmekte fakat bu ihlallerin hiç birinin manidar olmadığı Z değerlerinden anlaşılmaktadır. Dolayısıyla tüm maddelerin monoton azalmayan madde tepki fonksiyonuna sahip olduğu sonucuna ulaşılabılır. 1000 kişilik veri seti için elde edilen sonuçlar incelendiğinde ise maddelerin H_i katsayılarının 500 kişilik veri setine göre daha yüksek olduğu ve sadece bir maddenin 0.30'un altında H_i katsayısına sahip olduğu görülmektedir. Söz konusu madde OMSİ analizinde de ölçeğe uyumlu olmadığı belirlenen 12 nolu maddedir. Yine bu veri seti için monotonluktan ihlallerin sayısı incelendiğinde, 500 kişilik veri setine göre ihlal sayısının daha az olduğu ve söz konusu ihlallerin hiç birinin manidar olmadığı sonucuna ulaşılmaktadır. Bu bulgu da 500 kişilik veri setinden elde edilen sonuçlarla aynıdır ve tüm maddelerin madde tepki fonksiyonlarının monoton azalmayan nitelikte olduğu sonucuna ulaşılmaktadır.

Çalışma kapsamında incelemeye alınan son örneklem büyüklüğü koşulu olan 3000 kişilik veri setinden elde edilen bulgular incelendiğinde, madde tepki fonksiyonlarının monotonluk ihlalinin olmadığı görülmektedir. Maddelere ilişkin H_i değerleri incelendiğinde ise, 12 ve 25 nolu maddenin H değerinin sınır değer olan 0.3'ten düşük olduğu belirlenmiştir. Özetle, oluşturulan tüm veri setleri için madde tepki fonksiyonlarının

monoton azalmayan nitelikte olduğu varsayımı test edilmiş ve üç veri setinde yer alan tüm maddelerin bu varsayımı karşıladığı sonucuna ulaşılmıştır.

Bu aşamaya kadar yapılan analizlerden elde edilen bulgulara göre, 500, 1000 ve 3000 kişilik veri setlerinin POMTK kapsamında yer alan ilk model olan Monoton Homojenlik Modeline uyum sağladığı belirlenmiştir. Madde düzeyinde POMTK'ya göre yapılan uyum incelemeleri sonuçlarına göre yalnızca dört maddenin (M1, M12, M4 ve M25) POMTK kapsamında ölçeklemeye uygun olmadığı belirlenmiştir. Bu bulgu PMTK modelleri için yapılan model veri uyumu incelemesinden elde edilen bulgularla kıyaslandığında, PMTK ile yapılan ölçeklemede, 500 kişilik veri setinde beş madde modele madde düzeyinde uyum sağlamazken, 1000 kişilik veri setinde altı, 3000 kişilik veri setinde ise, 16 madde düzeyinde model veri uyumu sağlanamamıştır. Dolayısıyla PMTK'da ölçeklenemeyen maddelerin büyük çoğunluğu POMTK kapsamında yapılan ölçeklemeye uyum sağlamıştır. Elde edilen bu bulgu da alanyazında belirtilen bulgularla tutarlılık göstermektedir ve PMTK modelleri ile ölçeklenemeyen maddelerin POMTK kapsamında ölçeklenmesi mümkün olabilir (Meijer ve Baneke, 2004). Bu sayede testte yer alana madde sayısını azaltmadan, bir diğer ifadeyle geçerliğe yönelik bir tehdit oluşturmadan, POMTK ile sıralama düzeyinde ölçekleme yapılabilir.

Örneklem büyüklüğü koşullarına göre oluşturulan veri setlerinin POMTK'da yer alan Monoton Homojenlik Modeline uyum sağladığı belirlendikten sonra, daha kısıtlayıcı bir POMTK modeli olan Çift Monotonluk Modeline uyumunu belirleyen varsayım olan değişmez madde sıralaması varsayımını karşılama durumlarının incelemesine geçilmiştir.

Değişmez madde sıralaması varsayımının incelenmesi

Çalışmanın bu aşamasına kadar incelenen varsayımlar POMTK modelleri olan Monoton Homojenlik ve Çift Monotonluk Modeli için ortak olan varsayımlardır. Yapılan incelemeler ve elde edilen sonuçlar gözetildiğinde, oluşturulan veri setlerinin birkaç madde dışında monotonluk, tek boyutluluk ve yerel bağımsızlık varsayımlarını karşıladığı görülmektedir, dolayısıyla tüm veri setlerinin MHM ile uyumlu olduğu sonucuna ulaşılabilir. Söz konusu veri setlerinin ÇMM'ye uyum sağlaması için ise değişmez madde sıralaması ya da kesişmeyen madde tepki fonksiyonları olarak bilinen daha kısıtlayıcı bir varsayımı da karşılaması gerekmektedir. Bu varsayım veri setinde yer alan maddelerin güçlüklerine göre her yetenek düzeyinde aynı şekilde sıralanabileceğini belirtir. Maddelerin

güçlük değerleri grupların yetenek düzeyine göre değişse de, madde güçlük sıralamasının her yetenek düzeyinde aynı olmasını gerektirir ve bu özellik, değişen madde fonksiyonunun belirlenmesinden tanı ve teşhis amaçlı testlerin oluşturulmasına kadar ölçme niteliklerini geliştiren bir özelliktir (Sijtsma ve Molenaar, 2002). Bu varsayımın incelenmesi için önerilen P matrisi ve madde kalan puanı grupları yöntemleri ile incelemeler yapılmış ve elde edilen sonuçlar madde sıralamasının netliği hakkında bilgi veren H^T katsayısı ile birlikte yorumlanmıştır.

1. Madde kalan puanı grubu yöntemi kullanılarak yapılan değişmez madde sıralaması incelemesi sonuçları

Madde kalan puanı yönteminde, her madde çifti için kestirilen madde tepki fonksiyonları madde kalan puanı fonksiyonu ile karşılaştırılır. İkili puanlanan maddeler için madde kalan puanı fonksiyonları x ekseninde madde kalan puanlarını, y ekseninde her kalan grubu için gözlenen yanıtlanma oranlarını gösterir. Kalan gruplar oluşturulurken her grupta güvenilir kestirimlerin yapılmasını sağlayacak kadar birey bir araya getirilir ya da gruplar birleştirilir. Değişmez madde sıralaması sağlandığı durumda, farklı kalan grupları için hesaplanan madde oranları sıralaması tüm grup için hesaplanan madde oranları sıralaması ile aynı olur. Tüm madde çiftleri için kalan grup fonksiyonları hesaplanır ve kalan grup için hesaplanan yanıtlanma oranı tüm grup için belirlenen yanıtlanma oranı ile uyumsuz olduğunda, ihlal olarak raporlaştırılır (Meijer, Tenderio ve Wanders, 2015). Bazı durumlarda manidar olduğu belirtilen ihlallerin pratikteki etkisi oldukça düşük olabilir. Bu durumu belirlemek için, Molenaar ve Sijtsma (2000) ihlallerin etki büyüklüğünü ölçümünü veren Crit değerinin yorumlanmasını önermişlerdir. Bu değer 40'ın altında olması ciddi olmayan ihlalleri, 80 arasında olması ihmal edilebilir ihlalleri ve 80'den yüksek olması ise ciddi ihlaller olduğunu ve değişmezliğin sağlanmadığı anlamına gelir. Tüm bu bilgilere göre her üç örneklem büyüklüğünde yer alan maddelere ilişkin yapılan değişmezlik incelemeleri sonuçları Çizelge 12'de sunulmuştur.

Çizelge 12.

Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi

Maddeler	500 kişilik veri seti				1000 kişilik veri seti				3000 kişilik veri seti			
	H _i	İhlal	Crit	Z *	H _i	İhlal	Cirt	Z*	H _i	İhlal	Crit	Z*
M1	0.26	18	101	4	0.35	8	50	1	0.30	10	87	6
M2	0.42	19	78	2	0.41	7	42	2	0.40	6	62	4
M3	0.30	22	75	1	<u>0.36</u>	<u>8</u>	<u>31</u>	<u>0</u>	0.31	15	96	9
M4	0.42	25	114	7	0.43	18	88	5	0.43	23	137	19
M5	0.36	20	62	1	0.35	11	61	2	0.32	9	74	6
M6	0.35	22	83	3	0.37	7	41	1	0.36	13	80	9
M7	0.33	24	98	7	0.32	14	69	3	0.32	10	95	7
M8	0.36	21	72	2	0.40	14	63	3	0.39	18	114	12
M9	0.55	8	51	2	0.56	5	69	3	0.55	4	76	4
M10	0.28	38	104	5	0.34	18	50	1	0.29	14	95	9
M11	<u>0.31</u>	<u>17</u>	<u>47</u>	<u>0</u>	0.38	6	37	1	0.36	5	52	3
M12	0.21	41	156	13	<u>0.23</u>	38	150	17	<u>0.20</u>	39	197	33
M13	0.39	24	93	5	0.38	20	89	10	0.38	15	111	15
M14	0.4	19	77	2	0.42	13	72	6	0.39	6	61	4
M15	0.32	23	63	1	0.34	12	52	2	0.34	9	75	4
M16	0.39	12	69	1	0.40	6	56	2	0.40	6	68	3
M17	0.48	7	48	1	0.52	5	62	2	0.48	7	87	7
M18	0.37	24	89	3	0.39	9	54	3	0.37	4	51	3
M19	0.39	29	106	7	0.42	19	83	5	0.38	8	85	5
M20	0.3	27	107	6	0.37	12	63	4	0.33	7	77	4
M21	0.34	26	78	2	0.36	10	49	1	0.35	11	82	7
M22	<u>0.32</u>	<u>21</u>	<u>55</u>	<u>0</u>	0.37	10	56	4	0.34	5	52	2
M23	<u>0.64</u>	<u>1</u>	<u>-8</u>	<u>0</u>	0.67	2	33	2	0.63	2	37	2
M24	0.33	14	62	1	0.32	12	100	6	0.33	11	98	9
M25	0.24	30	122	8	0.30	14	96	7	<u>0.27</u>	14	122	10
M26	0.33	17	59	1	0.36	8	46	2	0.33	6	67	5
M27	0.34	20	71	2	0.39	13	56	2	0.36	7	64	3
M28	0.33	19	66	1	0.36	10	60	4	0.34	9	73	6
M29	0.42	27	76	2	0.42	13	65	4	0.40	9	57	5
M30	0.34	27	88	4	0.33	12	81	5	0.33	14	110	11

Çizelge 12’de bulunan değerler genel olarak incelendiğinde, tüm örneklem büyüklüklerinde değişmez madde sıralaması varsayımından neredeyse tüm maddeler için ihlaller olduğu görülmektedir. Beş yüz kişilik veri setinde sadece 11, 22 ve 23 nolu maddeler düzeyinde oluşan ihlaller manidar değilken, 1000 kişilik veri setinde ise 3 nolu maddenin manidar olmayan ihlali bulunurken, 3000 kişilik veri setinde yer alan tüm maddeler değişmez madde sıralaması varsayımını manidar düzeyde ihlal etmiştir. Varsayımından manidar düzeyde ihlal göstermeyen maddeler çizelgede altı çizili olarak gösterilmiştir. Neredeyse tüm maddelerin manidar düzeyde ihlaller gösterdiği göz önünde

bulundurularak, söz konusu ihlallerin etki büyüklüklerini belirten Crit değerlerinin incelenmesine gerek duyulmuştur. Molenaar ve Sijtsma (2000) tarafından önerildiği şekilde Crit değerleri için 80 ve üzerinde olan ihaller ciddi olarak kabul edilmiş ve bu maddeler çizelgede italik yazılarak gösterilmiştir. Veri setleri düzeyinde hesaplanan Crit değerleri Çizelge 13'te özetlenmiştir.

Çizelge 13.

Örneklere Göre Maddelerin Crit Değerlerinin İncelenmesi

Örneklemler	500	1000	3000
Crit Değerleri	Madde Sayıları		
<40	1	3	1
<80	16	20	14
≥80	13	7	15

Çizelge 13'te yer alan sonuçlar genel olarak incelendiğinde, değişmez madde sıralaması varsayımını en iyi düzeyde karşılayan örneklemin 1000 kişilik veri seti olduğu görülmektedir. Bu veri setindeki 20 madde ihmal edilebilir ihlaller göstermekte, 7 madde ise ciddi ihlaller göstermektedir. Beş yüz ve 3000 kişilik veri setlerinden elde edilen sonuçların ise benzer olduğu görülmektedir. Her iki örneklem için de Crit değeri 80 ve üzeri olan madde sayısı 10'dan fazladır ve ihmal edilebilir ihlal sayısına sahip olan madde sayısı da birbirine yakındır. Dolayısıyla değişmez madde sıralaması varsayımının 500 kişilik veri setinde 17 madde için, 1000 kişilik veri setinde 23 madde için, 3000 kişilik veri setinde ise 15 madde için sağlandığı sonucuna ulaşılabılır. Bu değerlere göre 1000 kişilik veri setinde için kısmi düzeyde karşılanan değişmez madde varsayımı 500 ve 3000 kişilik veri setlerinde karşılanamamıştır. Madde kalan grup puanı kullanılarak yapılan değişmez madde sıralaması varsayımı sonuçları bir diğer yöntem olan P matrisi yöntemiyle de incelenmiş ve her iki yöntemden elde edilen sonuçlar kıyaslanmıştır.

2. P matrisi yöntemi yöntemi kullanılarak yapılan değişmez madde sıralaması incelemesi sonuçları

Bir önceki aşamada madde kalan grubu yöntemiyle test edilen kesişmeyen madde tepki fonksiyonu varsayımı p matrisi yöntemi kullanılarak da test edilmiş ve sonuçlar Çizelge 14'te verilmiştir.

Çizelge 14.

Kesişmeyen Madde Tepki Fonksiyonu Varsayımının P Matrisi Yöntemiyle Test Edilmesi

Maddeler	500 kişilik veri seti				1000 kişilik veri seti				3000 kişilik veri seti			
	H _i	İhlal	Crit	Z *	H _i	İhlal	Cirt	Z*	H _i	İhlal	Crit	Z*
M1	0.26	8	57	7	0.35	0	0	0	0.30	0	0	0
M2	0.42	4	38	4	0.41	0	0	0	0.40	0	0	0
M3	0.30	6	38	3	0.36	0	0	0	0.31	2	48	2
M4	0.42	28	81	20	0.43	9	60	9	0.43	16	88	16
M5	0.36	1	23	1	0.35	2	36	2	0.32	0	0	0
M6	0.35	9	55	8	0.37	3	39	3	0.36	2	44	2
M7	0.33	2	31	2	0.32	1	28	1	0.32	0	0	0
M8	0.36	6	50	5	0.40	1	25	1	0.39	1	40	1
M9	0.55	11	50	10	0.56	6	45	6	0.55	1	25	1
M10	0.28	11	68	10	0.34	0	0	0	0.29	0	0	0
M11	0.31	4	40	3	0.38	0	0	0	0.36	0	0	0
M12	0.21	27	91	20	0.23	13	79	13	0.20	16	99	16
M13	0.39	2	27	2	0.38	0	0	0	0.38	1	30	1
M14	0.40	5	43	5	0.42	0	0	0	0.39	1	33	1
M15	0.32	6	46	4	0.34	2	31	2	0.34	0	0	0
M16	0.39	1	21	1	0.40	0	0	0	0.40	0	0	0
M17	0.48	5	30	3	0.52	3	36	3	0.48	13	90	13
M18	0.37	5	36	3	0.39	0	0	0	0.37	0	0	0
M19	0.39	6	43	5	0.42	3	37	3	0.38	0	0	0
M20	0.3	7	46	5	0.37	1	25	1	0.33	0	0	0
M21	0.34	2	32	2	0.36	1	27	1	0.35	0	0	0
M22	0.32	1	25	1	0.37	0	0	0	0.34	0	0	0
M23	0.64	0	0	0	0.67	0	0	0	0.63	0	0	0
M24	0.33	11	61	10	0.32	8	63	8	0.33	2	65	2
M25	0.24	13	70	11	0.30	1	32	1	0.27	12	98	12
M26	0.33	0	0	0	0.36	1	28	1	0.33	0	0	0
M27	0.34	3	30	2	0.39	1	23	1	0.36	0	0	0
M28	0.33	9	53	7	0.36	4	45	4	0.34	3	48	3
M29	0.42	7	48	6	0.42	2	31	2	0.40	1	32	1
M30	0.34	2	30	2	0.33	0	0	0	0.33	1	32	1

Çizelge 14’te yer alan değerler incelendiğinde, bir önceki yöntem olan madde kalan puanı gruplarına dayalı olarak elde edilen sonuçlarla benzerlik gösterdiği görülmektedir. Söz konusu çizelgede, her veri setinde yer alan maddelerin ölçeklenebilirlik katsayıları (H_i), maddelere ilişkin oluşturulan madde tepki fonksiyonlarının varsayımı ihlal ettiği durumların sayısı, bu ihlallerin manidarlığını belirten Z testi puanları ve ihlallerin etki büyüklüğünü belirten Crit değerleri yer almaktadır. Elde edilen sonuçlar genel olarak incelendiğinde, örneklemere göre değişmez madde sıralaması varsayımından manidar olarak sapma gösteren maddelerin sayısının değiştiği görülmektedir. En fazla ihlalin 500 kişilik veri setinde yer alan maddeler için olduğu görülmekte ve bu örnekte yer alan maddelerden yalnızca ikisi için (M23 ve M26) var olan ihlallerin manidar olmadığı, kalan tüm maddelerin söz konusu varsayımı manidar düzeyde ihlal ettiği görülmektedir. 1000 kişilik veri setinde için ise durum 500 kişilik veri setinden elde edilen sonuçlara göre çok farklıdır. 1000 kişilik veri setinde yer alan maddelerden, 18’i incelenen varsayımdan manidar olarak ihlaller göstermekte, 12 tanesi için ise var olan ihlaller manidar olmadığı görülmektedir. 3000 kişilik veri setinden elde edilen sonuçlar da 1000 kişilik veri setinden elde edilen sonuçlarla paralellik göstermekte ve bu örnekte yer alan 30 maddenin 14’ü varsayımdan manidar ihlaller göstermekte, 16’sı ise varsayımına uyum sağlamaktadır. Daha önce de belirtildiği gibi söz konusu ihlallerin etki büyüklükleri incelenerek pratikteki öneminin araştırılması gerekmektedir. Bu nedenle çizelgede yer alan Crit değerleri örneklemere göre incelenmiş ve elde edilen sonuçlar Çizelge 15’te sunulmuştur.

Çizelge 15.

Örneklere Göre Maddelerin Crit Değerlerinin İncelenmesi

Örneklemler	500	1000	3000
Crit Değerleri	Madde Sayıları		
<40	14	25	21
<80	14	5	5
≥80	2	-	4

Çizelge 15’te maddeler, örneklemere göre hesaplanan değişmez madde sıralaması varsayımından olan ihlallerin etki büyüklüklerine göre gruplandırılmıştır. 40’in altında olan Crit değerlerinin önemsiz ihlalleri, 80’in altında olan ihlallerin göz ardı edilebilecek küçük ihlalleri, 80’in üzerinde olan değerlerin ise ciddi ihlalleri belirttiği göz önünde bulundurulursa, 500 ve 3000 kişilik veri setlerinde ciddi ihlallerin sayısının oldukça düşük

olduğu, 1000 kişilik veri setindeki ihlallerin ise büyük çoğunluğunun ciddi olmayan ihlaller olduğu görülmektedir. Dolayısıyla bu yöntem kullanılarak yapılan değişmez madde sıralaması varsayımının incelenmesinde, 1000 kişilik veri seti için bu varsayımın karşılandığı, 500 ve 1000 kişilik veri setleri için ise varsayımı ihlal eden maddelerin çıkarılması ile bu varsayımın karşılanacağı sonucuna ulaşılmaktadır. Elde edilen sonuçlar aynı varsayımın test edildiği bir diğer yöntem olan madde kalan puanı grupları regresyonu sonuçları ile kıyaslandığında, söz konusu yöntemden elde edilen sonuçlara göre bu yöntemle elde edilen sonuçların, varsayımın karşılanma durumuna göre daha olumlu olduğu belirlenmiştir. Madde kalan puanı grupları yönteminde de varsayımı en iyi düzeyde karşılayan veri seti 1000 kişilik veridir ancak her üç örnekleme de ciddi ihlallerin sayısı daha fazladır. Her iki yöntemden elde edilen değerler arasında farklılıklar gözlemlendiği için maddelerin değişmez madde sıralaması varsayımından olan ihlallerini değerlendirmek için Lingvoet, van der Ark, Marvelde ve Sijtsma (2011) tarafından başvurulması gerekli görülen ve değişmeyen madde sıralamasının netliğini belirten H^T katsayısı tüm veri setleri için hesaplanmıştır. 0 ile 1 arasında değer alan H^T katsayısının düşük değerler alması, madde tepki fonksiyonlarının birbirine yakın olduğu anlamına gelirken, 1'e yakın değerler alması ise madde tepki fonksiyonlarının birbirinden oldukça uzak olduğu, dolayısıyla kesişmediği anlamına gelir. Sijtsma ve Molenaar (1992), bu katsayının tüm maddeler için ortak bir katsayı olduğu belirtmekte ve madde kalan puanı grupları ve P matrisinden elde edilen bilgilerin bu katsayı ile birleştirilerek yorumlanmasını önermektedir. İncelemeye alınan tüm örneklem büyüklükleri için hesaplanan H^T değerleri ve değişmeyen madde sıralaması varsayımını ihlal eden madde sayısı Çizelge 16'da verilmektedir.

Çizelge 16.

Örneklem İçin Hesaplanan H^T Katsayıları

Örneklem	500	1000	3000
H^T Katsayısı	0.285	0.289	0.285
Ciddi İhlal Sayısı	19	7	15

Çizelge 16'da ölçek düzeyinde hesaplanan H^T katsayıları ve her örneklem büyüklüğü için kesişmeyen madde sıralaması varsayımını ciddi düzeyde ihlal eden madde sayısı görülmektedir. H^T katsayıları incelenecek olursa, hiçbir örneklem büyüklüğünden elde edilen H^T katsayılarının 0.3'ten yüksek olmadığı görülmektedir. Bu durumda

maddelerin deęişmeyen madde sıralaması varsayımını en alt düzeyde dahi karşılamadığını göstermektedir. Ayrıca her örneklem düzeyinden söz konusu varsayımı ciddi düzeyde ihlal eden madde sayıları incelendiğinde, madde kalan puanı gruplarından elde edilen bulgulara benzer deęerlerin elde edildiđi Elde edilen bulgular deęerlendirildiğinde, bu varsayımın 500 ve 3000 kişilik veri setlerinde büyük oranda sağlanmadığı, 1000 kişilik veri setinde ise çeşitli revizyonlar yapılarak bu varsayımın karşılanabileceđi sonucuna ulaşılmıştır.

Son olarak incelenen veri setleri için POMTK modellerinden MHM modelinin uygulanabileceđi sonucuna ulaşılmış ve çalışma kapsamında madde parametreleri ve birey yetenek kestirimlerine etkisi incelenen bir diđer faktör olan test uzunluđu faktörü koşullarına göre oluşturulan veri setlerinin PMTK ve POMTK modellerine ilişkin yapılan model veri uyumu incelemelerine geçilmiştir.

Farklı Test Uzunluklarına Göre Oluşturulan Veri Setlerinin Parametrik Madde Tepki Kuramı Modellerine Uyumunun İncelenmesi

Araştırma kapsamında madde parametreleri ve birey yetenek kestirimlerine etkisi incelenen iki faktör olan örneklem büyüklüđu ve test uzunluđu faktörleri altında yer alan koşulların çaprazlanması ile oluşturulan dokuz veri seti öncelikli olarak PMTK kapsamında incelenmiştir. Yapılan ilk analiz veri setlerinin PMTK modellerine uyumunun incelenmesidir. Bu amaçla beş, 15 ve 25 maddeden oluşan 500, 1000 ve 3000 kişilik dokuz veri setinin BILOG-MG programı ve R programında yer alan “irtos” paketi kullanılarak bir, iki ve üç parametrelili lojistik modele uyumları incelenmiştir. Her veri seti ve model için hesaplanan -2 loglikelihood deęerleri ile söz konusu modele uyum sağlayan madde sayısı Çizelge 17’de verilmiştir.

Çizelge 17.

Farklı Test Uzunlukları ve Örneklem Büyüklüđüne Göre Oluşturulan Veri Setlerinin PMTK Modellerine Uyumu

Örneklem Büyüklüđu	1 PLM-2 PLM				2 PLM-3PLM			3PLM
	Madde Sayısı	-2LL	Sd	Uyumlu Madde sayısı-1plm	-2LL	Sd	Uyumlu Madde sayısı-2plm	Uyumlu Madde sayısı-3plm
500	5	7.75	4	0	5.63	5	0	0
	15	39.15	14	5	22.41	15	10	7
	25	20.60	24	9	40.12	25	22	17

(Devam ediyor.)

Çizelge 17.(devam)

Farklı Test Uzunlukları ve Örneklem Büyüklüğüne Göre Oluşturulan Veri Setlerin PMTK Modellerine Uyumu

Örneklem Büyüklüğü	1 PLM-2 PLM			2 PLM-3PLM			3PLM	
	Madde Sayısı	-2LL	Sd	Uyumlu Madde sayısı-1plm	-2LL	Sd	Uyumlu Madde sayısı-2plm	Uyumlu Madde sayısı-3plm
1000	5	10.12	4	0	9.25	5	0	0
	15	74.56	14	5	23.10	15	6	4
	25	187,12	24	12	48,52	25	16	17
3000	5	900.13	4	0	10.68	5	0	0
	15	288,12	14	0	18,20	15	1	0
	25	500,45	24	4	123,00	25	0	2

Çizelge 17 yer alan model uyum istatistikleri, -2 loglikelihood farklarına göre incelendiğinde, beş maddelik veri setleri için 500 kişilik veri setinde, ilgili serbestlik derecesine göre hiçbir parametrik modelin model veri uyum düzeyinde manidar bir iyileşme sağlamadığı belirlenmiştir. Ayrıca bu veri seti için madde düzeyinde uyum incelemesi sonuçlarına bakıldığında ise, hiçbir PMTK modeli için uyumlu maddenin olmadığı görülmektedir. Dolayısıyla 500 kişiden oluşan beş maddelik veri seti için çalışma kapsamında incelenen modellerden hiçbiri için uyum sağlanamamıştır. Beş maddeden oluşan 1000 kişilik veri seti için hesaplanan 1 PLM ile 2 PLM arasındaki -2 loglikelihood farkı modelin manidar yönde iyileştiğini belirtmektedir. Dolayısıyla bu veri seti için 2PLM'nin uyumlu olduğu sonucuna varılmıştır. Beş maddeden oluşan son veri ise 3000 kişilik veri setidir ve elde edilen sonuçlar bir önceki örneklem büyüklüğünden elde edilen değerlerle oldukça benzerdir. Loglikelihood değerleri arasındaki farklar incelendiğinde, model veri uyumunun 1 PLM'ye göre 2 PLM'de daha iyi olduğu bulunmuştur.

Test uzunluğu faktörü altında yer alan bir diğer koşul olan 15 maddelik veri setleri için elde edilen sonuçlar incelendiğinde ise, -2 loglikelihood değerlerine göre en yüksek iyileşme düzeyinin 1PLM ile 2PLM arasında olduğu görülmektedir. Tüm örneklem büyüklükleri için 2 PLM ve 3 PLM arasındaki -2 loglikelihood farkının manidar olmadığı bulunmuştur. Dolayısıyla bu değerlere göre 15 maddelik veri setlerinin 2 PLM'ye uygun olduğu sonucuna ulaşılmış ve uyum istatistikleri madde düzeyinde incelenmiştir. Madde

düzeyinde uyumlar incelendiğinde ise, 500 ve 1000 kişilik veri setleri için en yüksek sayıda uyumlu maddenin 2 PLM’de olduğu görülmektedir. 3000 kişilik veri setinde ise uyumlu madde sayısı 1 PLM ve 3 PLM için 0 iken, 2 PLM için birdir. Bu durumun nedeni daha önce de belirtildiği üzere ki-kare testinin örneklem büyüklüğüne duyarlı olmasıdır. Elde edilen bu bulgulara göre 15 maddelik veri setlerinin hem -2 loglikelihood değerlerindeki farkın manidarlığı hem de uyumlu madde sayısının diğer modellere nispeten yüksek olması nedeniyle 2 PLM’ye uyum sağladığı sonucuna ulaşılmıştır.

İncelemeye alınan son test uzunluğu koşulu olan 25 maddelik veri setleri için hesaplanan model veri uyumu istatistikleri genel olarak incelendiğinde ise, madde düzeyindeki en iyi model veri uyumunun 500 kişilik veri setinde ve 2 PLM için sağlandığı ve örnekleme yer alan birey sayısı arttıkça madde düzeyindeki uyumun azaldığı görülmektedir. Yirmi beş maddeden oluşan 500 kişilik veri setinin 22 maddesi 2 PLM’ye uyumluyken, 1000 kişilik veri setinde ise en iyi uyumun 3 PLM için sağlandığı belirlenmiştir ve bu model için 2PLM ile 3PLM arasında hesaplanan -2 loglikelihood farkının manidar olduğu belirlenmiştir. Dolayısıyla 25 maddelik veri setlerinden 500 ve 1000 kişilik veri setleri için 2 PLM’nin, 3000 kişilik veri seti için ise 3 PLM’nin uyumlu olduğu bulunmuştur.

Farklı madde sayılarına ve örneklem büyüklüklerine göre PMTK modellerinin model veri uyumu incelendiğinde elde edilen bulgular şu şekilde özetlenebilir:

- Beş maddelik veri setleri hiçbir örneklem büyüklüğünde PMTK modellerine uyum sağlamamıştır. Bu durum beş maddenin PMTK kestirimleri için yetersiz olduğu sonucunu ortaya çıkarmaktadır. Madde sayısının az olduğu durumlarda PMTK ile yapılan kestirimlerin tutarsız sonuçlar verdiği, MTK’nın avantajı olarak kabul edilen bireyden bağımsız madde parametrelerine ulaşamadığı bulunmuştur.
- Madde sayılarına göre PMTK uyumları genel olarak incelendiğinde en yüksek model veri uyum en fazla maddenin bulunduğu 25 maddelik veri setinde gözlenmiştir. Dolayısıyla madde sayısının artması PMTK modellerine olan uyumu artırmıştır ve bu durum da alan yazında kabul edilmiş bulgudur (Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1985).

Farklı Test Uzunluklarına Göre Oluşturulan Veri Setlerinin Parametrik Olmayan Madde Tepki Kuramı Modellerine Uyumunun İncelenmesi

Öncelikli olarak PMTK'ya uyumu incelenen farklı test uzunluklarından ve örneklem büyüklüklerinden oluşan veri setleri, bu aşamada POMTK kapsamında incelenmiş ve uyumlu model belirlenmiştir. POMTK'ya uyumlu modelin belirlenmesi için daha önce de belirtilen adımlar izlenmiş ve ilk olarak maddelerin ortalama güçlükleri hesaplanmıştır.

Beş, On Beş ve Yirmi Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Güçlükleri.

Oluşturulan farklı örneklem büyüklüğüne ve madde sayısına sahip üç veri seti için de madde güçlük parametreleri hesaplanmış ve Çizelge 18, 19 ve 20'de sunulmuştur.

Çizelge 18.

Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Değerleri

Maddeler	Örneklem Büyüklüğü		
	500	1000	3000
M1	0.71	0.72	0.72
M2	0.38	0.38	0.39
M3	0.69	0.71	0.71
M4	0.69	0.67	0.70
M5	0.56	0.55	0.55
Ortalama	0.60	0.61	0.61

Çizelge 18'de yer alan değerler incelendiğinde, tüm örneklemelere ilişkin hesaplanan ortalama güçlük değerlerinin birbirine çok yakın olduğu görülmektedir. Madde düzeyinde elde edilen parametrelerin de çok benzer olduğunu da çizelgede görmek mümkündür. Kestirilen değerlere göre, beş maddenin 3'ü kolay (M1, M3 ve M4), biri orta güçlükte (M5), kalan biri ise (M2) zordur.

İncelemelere bir diğer test uzunluğu koşulu olan 15 maddelik veri setleri ile devam edilmiş ve bu veri setlerinde yer alan maddeleri için hesaplanan madde güçlük değerleri Çizelge 19'da verilmiştir.

Çizelge 19.

On Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Değerleri

Örneklem Büyüklüğü				Örneklem Büyüklüğü			
Maddeler	500	1000	3000	Maddeler	500	1000	3000
M1	0.71	0.73	0.72	M9	0.26	0.27	0.25
M2	0.38	0.38	0.39	M10	0.61	0.65	0.66
M3	0.69	0.71	0.70	M11	0.37	0.37	0.36
M4	0.69	0.67	0.70	M12	0.69	0.69	0.68
M5	0.55	0.56	0.55	M13	0.52	0.50	0.50
M6	0.68	0.71	0.70	M14	0.39	0.38	0.40
M7	0.60	0.65	0.64	M15	0.57	0.57	0.59
M8	0.61	0.61	0.63	Ortalama	0.56	0.56	0.57

Çizelge 19’da verilen madde güçlük parametreleri genel olarak incelendiğinde, farklı örneklemlerde benzer değerlerin elde edildiği görülmektedir. Testlerin güçlük ortalaması 0.56 ve 0.57 olarak bulunmuştur, dolayısıyla testlerin orta güçlükte olduğu sonucuna ulaşılabilir ve bu bulgu PMTK kapsamında hesaplanan test güçlük düzeyleri ile paralellik göstermektedir. Elde edilen değerler madde düzeyinde incelendiğinde, madde güçlük indeksi düşük olan üç madde bulunmaktadır. Bunlar en düşüğe doğru; M9, M2 ve M11’dir.

Test uzunluğu faktörü altında yer alan son koşul olan 25 maddelik testlere ilişkin madde güçlük parametreleri de hesaplanmış ve elde edilen değerler Çizelge 20’de verilmiştir.

Çizelge 20.

Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Değerleri

Örneklem Büyüklüğü				Örneklem Büyüklüğü			
Maddeler	500	1000	3000	Maddeler	500	1000	3000
M1	0.71	0.73	0.72	M14	0.39	0.38	0.40
M2	0.38	0.38	0.39	M15	0.57	0.57	0.59
M3	0.69	0.71	0.70	M16	0.75	0.73	0.75
M4	0.69	0.67	0.70	M17	0.29	0.27	0.30
M5	0.56	0.56	0.55	M18	0.48	0.48	0.46
M6	0.68	0.71	0.70	M19	0.56	0.56	0.56
M7	0.61	0.65	0.65	M20	0.47	0.46	0.48
M8	0.61	0.61	0.63	M21	0.62	0.62	0.62
M9	0.26	0.27	0.25	M22	0.46	0.47	0.48
M10	0.61	0.65	0.66	M23	0.17	0.18	0.18
M11	0.37	0.37	0.36	M24	0.26	0.28	0.27
M12	0.67	0.69	0.68	M25	0.33	0.32	0.30
M13	0.52	0.50	0.51	Ortalama	0.51	0.51	0.51

Çizelge 20’de görülen madde güçlük parametreleri genel olarak incelendiğinde, örneklemelere göre çok benzer sonuçlar elde edildiği ve test ortalamalarının aynı olduğu görülmektedir. Her üç test de tam olarak orta güçlüktedir, ancak hesaplanan ortalama güçlük değerleri daha önceki test uzunluğu koşulu olan beş, 15 maddelik veri setlerinden hesaplanan değerlerden daha düşüktür. Bu durumun nedeni ise, söz konusu veri setlerinden farklı olarak bu veri setinde farklı 10 maddenin içinde güçlük düzeyleri düşük olan maddelerin yer almasıdır. Bu maddelerden biri test içindeki en düşük güçlük değerine sahip 23 nolu maddedir. Ayrıca 17, 24 ve 25 nolu maddelerin güçlük değerleri de test ortalamalarına göre oldukça düşüktür. Elde edilen sonuçlar örneklemelere göre incelendiğinde maddelerin farklı örneklem büyüklüklerinde benzer güçlük değerlerine sahip olduğu görülmektedir. Sonuç olarak, POMTK kapsamında hesaplanan madde güçlük indeksleri farklı örneklemelere ve test uzunluklarına göre incelendiğinde benzer sonuçların elde edildiği bulunmuş ve bu bulgu da PMTK’da elde edilen bulguyla tamamen örtüşmektedir. Her iki yaklaşım kapsamında da farklı test uzunlukları ve örneklem büyüklükleri için kestirilen madde güçlük parametreleri birbirine benzerlik göstermektedir. Madde güçlüklerinin hesaplanmasını takiben, POMTK analizlerinin ikinci aşaması olan ve madde ayırt edicilik parametresine karşılık gelen ölçeklenebilirlik katsayılarının hesaplanmasına geçilmiştir.

Beş, on beş ve yirmi beş maddelik veri setleri için ölçeklenebilirlik katsayılarının hesaplanması.

Farklı test uzunluklarından oluşan veri setleri için madde güçlüklerinin hesaplanmasının ardından, POMTK kapsamında ölçekleme yapmanın uygun olup olmadığını belirlemek için ölçeklenebilirlik katsayısı olarak bilinen H katsayılarının hesaplanmasına geçilmiştir. Ölçeklenebilirlik katsayılarının hesaplanmasına ilk olarak madde çiftleri için kestirilen H_{ij} katsayılarını hesaplanması ile başlanmış, daha sonra madde ve ölçek düzeyi için ölçeklenebilirlik katsayıları hesaplanmıştır.

1. Madde çiftleri için hesaplanan ölçeklenebilirlik katsayıları

Oluşturulan tüm veri setleri için madde çiftleri arasındaki kovaryanslar hesaplanmış ve negatif değerler olup olmadığına bakılmıştır. Beş maddeden oluşan 500, 1000 ve 3000 kişilik veri setleri incelendiğinde, her üç veri setinde de madde çiftleri arasındaki tüm kovaryansların pozitif olduğu belirlenmiştir. On beş maddeden oluşan veri setleri için

hesaplanan H_{ij} katsayıları incelendiğinde ise, 500 kişilik veri setinde 1 ile 12 nolu maddeler için hesaplanan H_{ij} değerinin negatif olduğu bulunurken, 1000 ve 3000 kişilik veri setleri için tüm H_{ij} değerlerinin pozitif olduğu belirlenmiştir. En uzun test olarak incelemeye alınan 25 maddelik veri setleri için hesaplanan H_{ij} değerleri incelendiğinde ise, yalnızca 500 kişilik veri setinde 1 ile 12 nolu madde çifti için hesaplanan H_{ij} katsayısının negatif olduğu, 1000 ve 3000 kişilik veri setleri için hesaplanan tüm H_{ij} katsayılarının ise pozitif olduğu belirlenmiştir.

Sonuç olarak yalnızca 15 ve 25 maddeden oluşan 500 kişilik veri setlerinde yer alan 1 ve 12 nolu maddeye ilişkin hesaplanan H_{ij} katsayısının negatif olduğu ve dolayısıyla bu veri setleri için bu maddelerin ölçeklenemeyeceği fakat diğer örneklerde yer alan tüm maddelerin POMTK modelleri kapsamında ölçeklemeye uygun olduğu bulunmuştur. Ayrıca bu sonuç aynı araştırma sorusunun a maddesinde incelenen 30 maddeden oluşan veri seti için elde edilen bulgularla oldukça benzerdir. Söz konusu analizlerde de 500 kişilik veri setinde yalnızca 1 ile 12 nolu madde arasındaki H_{ij} katsayısının negatif olduğu belirlenmiştir. Özetle, H_{ij} katsayıları bakımından beş maddelik veri setine göre, 15, 25 ve 30 maddelik veri setleri için benzer bulgular elde edildiği sonucuna ulaşılabilir.

2. Madde ve ölçek düzeyinde hesaplanan ölçeklenebilirlik katsayısı

Test uzunluğu faktörü altında oluşturulan beş, 15 ve 25 maddelik veri setlerinde yer alan her örneklem büyüklüğü için H_i , H ve Z değerleri, hesaplanmış ve farklı sayıda maddeden oluşan veri setlerine göre gruplandırılarak verilmiştir. Beş maddeden oluşan veri setleri için hesaplanan H_i ve Z katsayıları ile ölçeğin tümü için hesaplanan H katsayısı Çizelge 21’de verilmiştir.

Çizelge 21.

Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları

Veri Setleri	500		1000		3000	
Maddeler	H_i (Se)	Z	H_i (Se)	Z	H_i (Se)	Z
M1	0.265 (0.035)	9.32	0.334 (0.025)	16.18	0.279 (0.015)	23.38
M2	<u>0.520 (0.042)</u>	13.55	<u>0.518 (0.032)</u>	18.83	<u>0.502 (0.017)</u>	31.36
M3	0.296 (0.034)	10.76	0.331 (0.025)	16.51	0.299 (0.015)	25.81
M4	0.367 (0.033)	13.35	0.410 (0.023)	20.71	0.397 (0.013)	34.35
M5	0.385 (0.037)	12.72	0.407 (0.026)	19.13	0.374 (0.016)	29.35
H Katsayısı	0.359 (0.027)	18.92	0.395 (0.019)	28.93	0.364 (0.012)	45.69

Çizelge 21 incelendiğinde, en düşük H_i katsayısının her üç veri seti için de bir nolu madde için hesaplandığı görülmekte ve 500 ve 3000 kişilik veri setlerinde bu katsayıların 0.30 olan sınır değer altına düştüğü, 1000 kişilik veri setinde ise 0.334 olarak hesaplandığı görülmektedir. Üç nolu madde de bir nolu maddeye benzer ölçeklenebilirlik katsayısı değerlerine sahiptir ve bu madde 500 ve 3000 kişilik veri setlerinde sınır değere çok yakın değerler alırken, 1000 kişilik veri setinde ise sınır değer üstünde olduğu belirlenmiştir. Söz konusu iki madde dışında kalan üç madde de sınır değer üstünde değerler almıştır. Ayrıca tüm örneklemelerden hesaplanan H_i değerlerine ilişkin standart hata değerlerinin de sıfıra yakın olduğu bulunmuştur. H_i değerlerinin evrende sıfırdan farklı olup olmadığını inceleyen Z değerleri de tüm maddeler için sıfırdan farklı bulunmuş, dolayısıyla H_i katsayılarının evrende de sıfırdan manidar olarak farklı olduğu sonucuna gidilmektedir.

Çizelge 21’de yer alan son değer ise ölçek düzeylerinde hesaplanan H katsayılarıdır. Ölçeğin tümü için hesaplanan H katsayılarının maddeler için H_i katsayılarına bağlı olarak hesaplandığı göz önünde bulundurulduğunda, en düşük H katsayısının 500 kişilik veri setinden kestirilmesi beklenen bir durumdur. En yüksek H katsayısı ise madde ölçeklenebilirlik katsayılarının en yüksek olduğu 1000 kişilik veri setinden hesaplandığı görülmekte ve bu katsayının 0.4’e yakınlığı göz önünde bulundurulduğunda, bu veri setinin Mokken ölçekleme tekniğine orta düzeyde uyum sağladığı, 500 ve 3000 kişilik veri setlerinin ise zayıf uyum sağladığı sonucuna ulaşılmıştır. Bu bulgu bir önceki aşamada hesaplanan 30 maddelik veri seti için elde edilen bulgularla paraleldir, 30 maddelik veri setleri içinde de Mokken ölçekleme tekniğine en iyi uyum 1000 kişilik veri seti için sağlanmıştır.

İncelemeye alınan bir diğer test uzunluğu koşulu olan 15 maddelik veri setleri için hesaplanan katsayılar Çizelge 22’de verilmiştir.

Çizelge 22.

On Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları

Veri setleri	500		1000		3000	
Maddeler	H_i (Se)	Z	H_i (Se)	Z	H_i (Se)	Z
M1	0.239 (0.030)	14.76	0.337 (0.020)	29.08	0.281 (0.012)	41.67
M2	0.444 (0.029)	25.42	0.427 (0.022)	34.46	0.422 (0.013)	57.39

(Devam ediyor.)

Çizelge 22. (devam)

On Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları

Veri setleri	500		1000		3000	
Maddeler	H_i (Se)	Z	H_i (Se)	Z	H_i (Se)	Z
M3	0.287 (0.029)	18.52	0.331 (0.020)	29.80	0.302 (0.012)	46.71
M4	0.408 (0.023)	26.31	0.410 (0.017)	38.47	0.410 (0.010)	64.14
M5	0.366 (0.027)	24.38	0.351 (0.020)	32.69	0.326 (0.012)	50.99
M6	0.340 (0.026)	22.14	0.360 (0.019)	32.23	0.343 (0.011)	53.12
M7	0.343 (0.026)	23.33	0.328 (0.019)	31.12	0.313 (0.011)	50.53
M8	0.376 (0.025)	25.53	0.418 (0.017)	39.63	0.390 (0.010)	62.94
M9	<u>0.606 (0.030)</u>	27.00	<u>0.608 (0.020)</u>	39.53	<u>0.602 (0.013)</u>	62.06
M10	0.283 (0.028)	19.23	0.321 (0.020)	30.47	0.286 (0.011)	46.08
M11	0.345 (0.035)	19.22	0.394 (0.024)	31.39	0.382 (0.015)	49.84
M12	0.202 (0.031)	13.10	0.230 (0.022)	21.15	0.184 (0.013)	29.29
M13	0.406 (0.026)	26.40	0.412 (0.020)	36.79	0.399 (0.011)	60.24
M14	0.422 (0.030)	24.21	0.446 (0.020)	36.08	0.412 (0.012)	56.69
M15	0.306 (0.028)	20.61	0.353 (0.020)	33.22	0.334 (0.011)	53.27
H Katsayısı	0.352 (0.016)	60.35	0.376 (0.012)	90.65	0.352 (0.007)	143.369

Çizelge 22’de yer alan değerler incelendiğinde 12 nolu maddenin her üç veri setinde de sınır değer olan 0.3’ten daha düşük H_i değerine sahip olduğu ve en düşük değeri 3000 kişilik veri setinde 0.184 olarak aldığı görülmektedir. Örneklemelere göre inceleme yapıldığında ise, 500 kişilik veri setinde dört maddenin (M1, M3, M10 ve M12), 1000 kişilik veri setinde bir maddenin (M12), 3000 kişilik veri setinde ise üç maddenin (M1, M10 ve M12) 0.3’ten daha düşük değerler aldığı görülmektedir. En yüksek H_i değerine sahip olan madde ise her üç veri setinde de aynıdır, M9, ve çizelgede kalın ve italik yazılarak gösterilmiştir. Söz konusu maddenin ayırt edicilik değeri olarak yorumlanan H_i katsayısının çok yüksek olduğu görülmektedir. Maddeler genel olarak incelendiğinde ise, en yüksek H_i değerlerinin 1000 kişilik veri setinden elde edildiği, en düşük H_i değerlerinin ise 500 kişilik veri setinden elde edildiği görülmektedir. Ölçeklenebilirlik katsayılarının manidarlığı hakkında bilgi veren Z katsayıları ise tüm maddeler için sıfırdan yüksektir, dolayısıyla tüm maddelerin ölçeklenebilirlik katsayısının örneklemelerde olduğu gibi evrende de sıfırdan farklı olduğu sonucuna ulaşılabilir. Maddelerin H_i değerlerine ilişkin

incelenmesi gereken bir başka değer ise standart hata değerleridir. Çizelge 22 incelendiğinde veri setinde bulunan birey sayısı arttıkça standart hata değerlerinin düştüğü ve bu durumun tüm maddeler için geçerli olduğu görülmektedir. Bu bulgu alanyazında elde edilen bulgularla örtüşmektedir (Kuijpers, van der Ark, Croon ve Sijtsma, 2016). Belirtilen çalışmada da örnekleme yer alan birey sayısının artmasının, madde düzeyinde hesaplanana ölçeklenebilirlik katsayılarının daha yansız kestirildiği bulunmuştur. Bu çalışma için aynı durum ortaya çıkmış, örnekleme yer alan birey sayısı arttıkça, ölçeklenebilirlik katsayılarına ilişkin standart hata değerleri düşmüştür.

Çizelge 22’de yer alan son değer olan H katsayıları incelendiğinde ise tüm veri setleri için 0.3 ile 0.4 arasında olduğu ve 500 ile 3000 kişilik veri setlerinin aynı H değerine, 0.352, sahip olduğu görülmektedir. 1000 kişilik veri setinin H değeri diğer örneklere göre daha yüksek bulunmuştur ve bu durumun nedeni ise örnekleme yer alan maddelerin H_i değerlerinin diğer örneklemlerdekilere göre daha yüksek olmasıdır. Özetle H değerleri incelendiğinde, tüm ölçeklerin Mokken ölçeklemeye zayıf düzeyde uyum sağladığı sonucuna ulaşılmaktadır.

İncelemeye alınan son test uzunluğu koşulu ise 25 maddelik testler için hesaplanan H_i değerleri, bu değerlere ait Z istatistikleri ve ölçeklerin bütününe ilişkin H katsayıları Çizelge 23’te yer almaktadır.

Çizelge 23.

Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları

Veri setleri	500		1000		3000	
Maddeler	H_i (Se)	Z	H_i (Se)	Z	H_i (Se)	Z
M1	0.262 (0.028)	19.44	0.352 (0.019)	36.51	0.297 (0.012)	52.73
M2	0.427 (0.026)	33.25	0.413 (0.020)	45.69	0.411 (0.011)	76.63
M3	0.301 (0.028)	23.35	0.356 (0.019)	38.39	0.317 (0.012)	58.32
M4	0.424 (0.022)	32.87	0.438 (0.016)	49.65	0.436 (0.009)	81.44
M5	0.371 (0.024)	31.41	0.353 (0.018)	42.26	0.332 (0.011)	66.90
M6	0.352 (0.024)	27.53	0.382 (0.019)	40.92	0.367 (0.011)	67.81
M7	0.338 (0.025)	28.43	0.332 (0.019)	38.43	0.324 (0.011)	63.79

(Devam ediyor.)

Çizelge 23. (devam)

Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan H_i ve Z Katsayıları

Veri setleri	500		1000		3000	
Maddeler	H_i (Se)	Z	H_i (Se)	Z	H_i (Se)	Z
M8	0.379 (0.023)	31.78	0.417 (0.017)	49.31	0.405 (0.010)	80.52
M9	0.554 (0.028)	35.91	0.558 (0.019)	53.36	0.554 (0.012)	84.53
M10	0.274 (0.028)	22.92	0.337 (0.019)	39.04	0.297 (0.011)	57.89
M11	0.318 (0.032)	24.26	0.394 (0.024)	41.87	0.366 (0.013)	66.33
M12	0.220 (0.032)	17.08	0.241 (0.022)	26.63	0.197 (0.013)	37.55
M13	0.396 (0.024)	33.33	0.389 (0.018)	46.03	0.388 (0.010)	77.60
M14	0.399 (0.027)	31.08	0.425 (0.019)	47.04	0.399 (0.011)	74.89
M15	0.318 (0.025)	26.97	0.351 (0.018)	41.93	0.343 (0.011)	68.78
M16	0.388 (0.030)	26.21	0.391 (0.021)	40.00	0.397 (0.012)	65.74
M17	0.468 (0.030)	32.02	0.512 (0.022)	49.00	0.476 (0.012)	79.28
M18	0.376 (0.025)	31.28	0.398 (0.018)	46,95	0.387 (0.011)	76.13
M19	0.402 (0.023)	34.10	0.423 (0.017)	50.55	0.389 (0.010)	78.34
M20	0.301 (0.029)	24.96	0.373 (0.020)	43.57	0.339 (0.011)	67.42
M21	0.340 (0.025)	28.17	0.355 (0.018)	41.74	0.347 (0.011)	69.02
M22	0.314 (0.028)	25.87	0.370 (0.019)	43.43	0.347 (0.011)	68.79
M23	<u>0.625 (0.035)</u>	31.45	<u>0.657 (0.022)</u>	49,15	<u>0.624 (0.013)</u>	78.59
M24	0.320 (0.040)	20.92	0.305 (0.027)	29.82	0.329 (0.016)	53.15
M25	0.243 (0.036)	17.82	0.295 (0.025)	30.13	0.263 (0.015)	43.89
H Katsayısı	0.357 (0.015)	97.85	0.386 (0.011)	150.02	0.367 (0.006)	239.71

Çizelge 23 incelendiğinde tüm örneklem için hesaplanan H_i katsayılarının büyük çoğunluğunun 0.3 değerinin üzerinde olduğu, en yüksek H_i katsayısının her üç örnekleme de 23 nolu madde için hesaplandığı, en düşük H_i katsayısının ise yine üç örnekleme de ortak olarak 12 nolu madde için hesaplandığı görülmektedir. Tüm maddelere için hesaplanan Z katsayıları sıfırdan farklıdır ve H_i katsayılarına ilişkin hesaplanan standart hata değerlerinin en düşük 3000 kişilik veri setinde, en yüksek ise 500 kişilik veri setinde kestirildiği görülmektedir. Daha önce de belirtildiği üzere örnekleme yer alan birey sayısı arttıkça H_i katsayılarına ilişkin kestirilen standart hatalar azalmaktadır. Örneklemelere göre H_i katsayıları incelenecek olursa, 500 ve 3000 kişilik veri setlerinde üç maddenin H_i

değerinin 0.3'ün altında olduğu, 1000 kişilik veri setinde ise yalnızca bir maddenin sınır değerden düşük H_i değerine sahip olduğu görülmektedir. Çizelge 23'te koyu yazılarak gösterilen bu maddeler, tüm veri setlerinde aynı maddelerdir. Bu durumda örneklerde yer alan birey sayısı gözetilmeksizin maddelerin benzer H_i değerlerine sahip olduğu sonucuna ulaşılabilir. Ayrıca H_i değerlerinin tümü için hesaplanan Z katsayılarının sıfırdan farklı olduğu görülmektedir. Son olarak ölçekler için hesaplanan H katsayıları incelendiğinde ise, üç değerinde 0.3 ile 0.4 arasında yer aldığı görülmekte, dolayısıyla ölçeklerin Mokken ölçekleme tekniğine zayıf uyum gösterdiği sonucuna ulaşılmaktadır. H katsayılarına ilişkin hesaplanan standart hata değerleri incelendiğinde ise, tüm değerlerin düşük olduğu fakat yine en yüksek değer 500 kişilik veri setinde, en düşük değer ise 3000 kişilik veri setinde kestirildiği görülmektedir.

Farklı madde sayısı ve örneklem büyüklüklerinden oluşan veri setlerinde yer alan maddeler, madde çiftleri ve veri setlerinin tümü için hesaplanan ölçeklenebilirlik katsayılarına ilişkin elde edilen bulgular özetlenecek olursa şu sonuçlara ulaşılabilir;

- Maddeler arası ikili kovaryanslara dayalı olarak hesaplanan H_{ij} katsayılarından, yalnızca 15 ve 25 maddeden oluşan 500 kişilik veri setinde yer alan 1 ve 12 nolu maddeye ilişkin hesaplanan H_{ij} katsayısı negatif değer almıştır. Dolayısıyla bu veri setleri için bu maddelerin ölçeklenemeyeceği fakat diğer veri setlerinde yer alan tüm maddelerin POMTK modelleri kapsamında ölçeklemeye uygun olduğu bulunmuştur. Bu bulgu araştırma sorusunun a seçeneğinde incelenen farklı örneklem büyüklüklerinden oluşan veri setleri için elde edilen bulgularla benzerdir. Mokken ölçeklemeye uyum sağlamayan maddeler her iki analizde de aynı maddelerdir.
- Farklı madde sayılarından oluşan veri setleri içinde Mokken ölçekleme tekniğine en iyi düzeyde uyum sağlayan veri setinin, 1000 kişilik veri seti olduğu bulunmuştur. Tüm farklı madde sayısı koşullarında 1000 kişilik veri setindeki maddeler için hesaplanan H_i değerleri daha yüksek bulunmuş, dolayısıyla bu veri setleri için hesaplanan H değerleri de daha yüksek bulunmuştur. Bu bulgu da alanyazında Mokken ölçekleme için gereken en düşük örneklem büyüklüğünü belirlemeyi amaçlayan çalışmada elde edilen bulgularla paralellik göstermektedir. Maddelere ilişkin ayırt edicilik değerinin veren H_i katsayılarının 0.3'e yakın olması durumunda 1500'den küçük örneklerde yanlış sınıflama oranının yüksek olduğu çalışmada belirtilen sonuçlardandır (Straat, van der

Ark ve Sijtsma, 2014). Bu nedenle 500 kişilik örneklem en düşük uyumun elde edildiği veri seti olmuştur.

- H_{ij} katsayıları negatif bulunan 1 ve 12 nolu maddeden, 1 nolu maddenin H_i değerleri 500 ve 3000 kişilik tüm veri setlerinde ve 12 nolu maddenin ise tüm veri setlerinde 0.3'ten düşük olduğu görülmüştür. Dolayısıyla bu maddelerin veri seti dışında tutulmasıyla Mokken ölçeklemeye olan uyumun artacağı sonucuna ulaşılmıştır. Ayrıca bu bulgu 30 maddelik veri setleri için elde edilen bulguyla aynıdır, dolayısıyla madde sayısındaki değişiklik bu maddelere ilişkin hesaplanan H_i değerlerini etkilememiştir. Bu bulgu da alanyazında elde edilen bulgularla benzerlik göstermekte, test uzunluğunun madde düzeyindeki ölçeklenebilirlik katsayıları üzerindeki etkisi ihmal edilebilir düzeyde kalmıştır (Kuijpers, van der Ark, Croon ve Sijtsma, 2016).
- Farklı madde sayısı ve örneklemelerden oluşan veri setleri için kestirilen tüm ölçeklenebilirlik katsayıları (H_{ij} , H_i ve H) için Z katsayıları hesaplanmış ve tüm değerlerin sıfırdan farklı olduğu, dolayısıyla bu katsayıların sıfırdan manidar şekilde farklılaştığı sonucuna ulaşılmıştır.
- Tüm ölçeklenebilirlik katsayıları için hesaplanan standart hata değerleri düşüktür fakat madde ve örneklemde yer alan birey sayısındaki artışın standart hata değerlerini düşürdüğü sonucuna ulaşılmıştır.

Ölçeklenebilirlik katsayılarına ilişkin yapılan incelemelerden sonra veri setinin Mokken ölçeklemenin bir başka varsayımı olan tek boyutluluğa olan uyumunun incelenmesine geçilmiş ve elde edilen bulgular aşağıda sunulmuştur.

Tek boyutluluk incelemesi

Parametrik Olmayan Madde Tepki Kuramı yaklaşımı kapsamında tek boyutluluk analizi için Otomatik Madde Seçim İşlemi (OMSİ-Automated Item Selection Procedure-AISP) kullanılmaktadır. PMTK kapsamında Açıklayıcı Faktör Analizi ile incelenen bu varsayım bu aşamada OMSİ ile incelenmiştir. Alan yazında önerildiği şekilde ölçeklenebilirlik katsayıları 0.3, 0.2 ve 0 olarak değiştirilerek ve Mokken analizinde default olarak uygulanan “normal” yöntemine ek olarak “Genetic Algorithm (ga)” yöntemi de kullanılarak tek boyutluluk analizi yapılmıştır. Elde edilen bulgular veri setlerinde yer alan madde sayılarına göre verilmiş ve Çizelge 24’te beş maddeden oluşan 500, 1000 ve 3000 kişilik veri setleri için yapılan OMSİ analiz sonuçları yer almaktadır.

Çizelge 24.

Beş Maddelik Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi

Veri Setleri	500				1000				3000			
Yöntem	nor	ga	nor	ga	nor	ga	nor	ga	nor	ga	nor	ga
$H \geq$	0.3	0.3	0.2	0.2	0.3	0.3	0.2	0.2	0.3	0.3	0.2	0.2
M1	0	0	1	1	1	1	1	1	0	1	1	1
M2	1	1	1	1	1	1	1	1	1	1	1	1
M3	1	1	1	1	1	1	1	1	1	0	1	1
M4	1	1	1	1	1	1	1	1	1	1	1	1
M5	1	1	1	1	1	1	1	1	1	1	1	1
Uyumsuz Madde Sayısı	1	1	0	0	0	0	0	0	1	1	0	0

Çizelge 24’ te yer alan “nor” ifadesi OMSİ’nin default yöntemini belirtmekteyken, “ga” değerleri generic algorithm yönteminden elde edilen sonuçları göstermektedir. Çizelgede yer alan değerler kısaca belirtilecek olursa, maddelere ilişkin 1 değeri, maddenin ölçülen gizil değişkeni yeterli ayırt edicilikte ölçtüğü ve teste katkı sağladığı anlamına gelirken, 0 değeri bu durumun tam tersini ifade etmektedir. Sıfır değerini alan maddelerin testle ölçülmek istenen yapıya yeterince hizmet etmediği düşünülmektedir. Bu bilgiler ışığında beş maddelik veri setleri için elde edilen bulgular incelendiğinde, incelenen her iki yöntemden de elde edilen sonuçların aynı olduğu, yalnızca sınır değer 0.3’ten 0.2’ye düşürüldüğü durumda bekleneildiği üzere uyumlu madde sayısının arttığı görülmektedir. Örneklemelere göre elde edilen sonuçlar incelenecek olursa, tüm maddelerin her iki yöntemde de tek boyutluluğu sağladığı veri setinin 1000 kişilik veri seti olduğu görülmektedir. Diğer iki örneklem büyüklüğü koşulu olan 500 ve 3000 kişilik veri setlerinde ise 0.3 sınırına ve normal yöntemine göre yapılan kestirimlerde 1 nolu madde uyumsuz olarak belirlenmiştir. Ayrıca 3000 kişilik veri setinde 1 nolu maddeden farklı olarak ga yöntemiyle 3 nolu madde uyumsuz olarak belirlenmiştir. Özetle, beş maddelik veri setleri için yapılan tek boyutluluk analizinde 1 nolu maddenin iki veri setinde, 3 nolu maddenin ise yalnızca bir veri setinde uyumsuz bulunduğu belirlenmiştir.

Araştırmada incelenen bir diğer madde sayısı koşulu olan 15 maddelik veri setleri için de OMSİ analizleri yapılmış ve elde edilen sonuçlar Çizelge 25’te verilmiştir.

Çizelge 25.

On Beş Maddelik Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi

Veri Setleri	500				1000				3000			
	nor	ga	nor	ga	nor	ga	nor	ga	nor	ga	nor	ga
H_≥	0.3	0.3	0.2	0.2	0.3	0.3	0.2	0.2	0.3	0.3	0.2	0.2
M1	0	0	1	1	1	1	1	1	1	1	1	1
M2	1	1	1	1	1	1	1	1	1	1	1	1
M3	1	1	1	1	1	1	1	1	1	1	1	1
M4	1	1	1	1	1	1	1	1	1	1	1	1
M5	1	1	1	1	1	1	1	1	1	1	1	1
M6	1	1	1	1	1	1	1	1	1	1	1	1
M7	1	1	1	1	1	1	1	1	1	1	1	1
M8	1	1	1	1	1	1	1	1	1	1	1	1
M9	0	1	1	1	1	1	1	1	1	1	1	1
M10	1	0	1	1	1	1	1	1	1	1	1	1
M11	1	1	1	1	1	1	1	1	1	1	1	1
M12	0	0	0	0	0	0	1	1	0	0	0	1
M13	1	1	1	1	1	1	1	1	1	1	1	1
M14	1	1	1	1	1	1	1	1	1	1	1	1
M15	1	1	1	1	1	1	1	1	1	1	1	1
Uyumsuz Madde Sayısı	3	3	1	1	1	1	0	0	1	1	0	0

Çizelge 25'te belirtilen ki yönteme ve iki ayrı sınır değere göre hesaplanan değerler incelendiğinde, dört ayrı maddenin, 1, 9, 10 ve 12 nolu maddelerin, farklı veri setlerinde uyumsuz olarak belirlendiği görülmektedir. Bu maddelerden 12 nolu maddenin 0.3 sınır değerinde 1000 ve 3000 kişilik veri setlerinde, 500 kişilik veri setinde ise hem 0.3 hem de 0.2 sınır değerlerinde uyumsuz olduğu bulunmuştur. Uyumsuz olduğu belirlenen diğer tüm maddelerin 500 kişilik veri setinde yer aldığı görülmektedir, dolayısıyla 15 maddelik veri setleri için en fazla sayıda uyumsuz maddenin 500 kişilik veri setinde olduğu belirlenmiştir. 1000 ve 3000 kişilik veri setleri için yalnızca 12 nolu maddenin uyumsuz olduğu görülmektedir. Elde edilen bulgular özetlenecek olursa, 12 nolu maddenin neredeyse tüm örneklerde uyumsuz olarak belirlendiği görülmekte, 1, 9 ve 10 nolu maddeler ise yalnızca 500 kişilik veri setinde uyumsuzluk göstermektedir. Dolayısıyla 12 nolu maddenin ölçek dışında bırakılmasının gerekli olduğu ve diğer maddeler için veri seti düzeyinde revizyonlar yapılması gerektiği sonucuna ulaşılabılır.

İncelemeye alınan son test uzunluğu koşulu olan 25 maddelik veri setleri için uygulanan OMSİ analizi sonuçları Çizelge 26'da verilmiştir.

Çizelge 26.

Yirmi Beş Maddelik Veri Setlerinin Otomatik Madde Seçim İşlemi Yöntemiyle Tek Boyutluluk Analizi

Veri setleri	500				1000				3000			
Yöntem	Nor	ga	nor	ga	nor	ga	nor	ga	nor	ga	nor	ga
H_i ≥	0.3	0.3	0.2	0.2	0.3	0.3	0.2	0.2	0.3	0.3	0.2	0.2
M1	2	0	1	0	1	1	1	1	1	1	1	1
M2	1	1	1	1	1	1	1	1	1	1	1	1
M3	1	1	1	1	1	1	1	1	1	1	1	1
M4	1	1	1	1	1	1	1	1	1	1	1	1
M5	1	1	1	1	1	1	1	1	1	1	1	1
M6	1	1	1	1	1	1	1	1	1	1	1	1
M7	1	1	1	1	1	1	1	1	1	1	1	1
M8	1	1	1	1	1	1	1	1	1	1	1	1
M9	1	1	1	1	1	1	1	1	1	1	1	1
M10	0	0	1	1	1	1	1	1	1	1	1	1
M11	1	1	1	1	1	1	1	1	1	1	1	1
M12	0	0	0	1	0	0	1	1	0	0	0	0
M13	1	1	1	1	1	1	1	1	1	1	1	1
M14	1	1	1	1	1	1	1	1	1	1	1	1
M15	1	1	1	1	1	1	1	1	1	1	1	1
M16	1	1	1	1	1	1	1	1	1	1	1	1
M17	1	1	1	1	1	1	1	1	1	1	1	1
M18	1	1	1	1	1	1	1	1	1	1	1	1
M19	1	1	1	1	1	1	1	1	1	1	1	1
M20	1	1	1	1	1	1	1	1	1	1	1	1
M21	1	1	1	1	1	1	1	1	1	1	1	1
M22	1	1	1	1	1	1	1	1	1	1	1	1
M23	1	1	1	1	1	1	1	1	1	1	1	1
M24	1	1	1	1	1	1	1	1	1	1	1	1
M25	2	0	1	1	0	1	1	1	0	0	1	1
Uyumsuz Madde Sayısı	4	4	1	1	2	1	0	0	2	2	1	1

Çizelge 26'da yer alan bulgular genel olarak incelendiğinde, bu test uzunluğu koşulu için, uyumsuz maddelerin, daha az madde içeren veri setlerine göre, daha fazla olduğu görülmektedir. Özellikle 500 kişilik veri setlerinde H_i katsayısı sınır değerinin 0.3 olarak alındığı durumda, dört maddenin, 1,20, 12 ve 25 nolu maddeler, uyumsuz olarak belirlendiği, hatta diğer test uzunluğu koşullarına ilişkin yapılan incelemelerden farklı olarak 1 ve 25 nolu maddenin testin bütününden farklı bir özelliği ölçmeye yönelik olduğu görülmektedir. Söz konusu iki maddeye ilişkin değerler 1 ve 0'dan farklı olarak 2 olarak

gösterilmiştir. Beş yüz kişilik veri seti için elde edilen sonuçlar incelendiğinde ise, H_i katsayısı sınır değerinin 0.2 olarak alındığı durumda her iki yöntemde de birer madde uyumsuz olarak görülmektedir. Normal yöntemi kullanıldığında 12 nolu madde, ga yöntemi kullanıldığında ise 1 nolu madde uyumsuz olarak bulunmuştur. Söz konusu maddeler çalışmanın daha önceki aşamalarında da gerek maddeler arası kovaryanslarının negatif olması ile gerekse ölçeklenebilirlik katsayılarının düşük olması nedeniyle 500 kişilik veri seti için uyumsuz olduğu belirlenen maddelerdir. Sınır değerin 0.2'ye çekilmesi uyumlu madde sayısını artırmıştır. Bin kişilik veri setinde de aynı durum geçerlidir, sınır değeri 0.2 olarak alındığında her iki yöntemde de tüm maddeler uyumlu olarak bulunmuştur. H_i katsayısı sınır değerinin 0.3 olarak alındığı durumda ise normal yöntemine göre 25 nolu madde de uyumsuz olarak belirlenmiştir. 3000 kişilik veri setinde ise H_i katsayısı sınır değerinin 0.3 olarak alındığı durumda 12 ve 25 nolu maddeler, 0.2 olarak alındığı durumda ise yalnızca 12 nolu madde uyumsuz olarak belirlenmiştir.

Farklı madde sayılarından ve örneklem büyüklüklerinden oluşan veri setleri için POMTK kapsamında yapılan OMSİ analizlerinden elde edilen bulgular özetlenecek olursa;

- Madde sayılarına ve örneklere göre uyumlu madde sayısı genel olarak benzer olmakla birlikte, en fazla uyumsuz maddenin, tüm farklı madde sayısı koşullarında, 500 kişilik veri setlerinde yer aldığı bulunmuştur.
- Farklı veri setlerinden uyumsuz olarak belirlenen maddeler genel olarak aynıdır ve 12 ile 25 nolu maddenin birden çok veri setinde ve incelenen iki yöntem de de ortak olarak uyumsuz olarak belirlenmiştir. Söz konusu maddeler araştırma kapsamında daha önce yapılan OMSİ analizinde de uyumsuz olarak buldukları için, bu maddeler için Mokken ölçekleme tekniğine uyum sağlamadığı sonucuna ulaşılabilir ve bu maddelerin ölçek dışında tutulması söz konusu olabilir.
- POMTK analizlerinde bu yaklaşıma göre ölçeklemeye uyumsuz olduğu belirlenen madde sayısı POMTK kapsamında yapılan model veri uyumu analizlerinde uyumsuz olarak belirlenen maddelere göre daha azdır. Dolayısıyla raporda daha önce belirtildiği gibi, alan yazındaki bulgulara da benzer olarak, POMTK'nın doğrulayıcı yaklaşımından dolayı ölçeklenemeyen maddelerin POMTK kapsamında ölçeklenmesi mümkün olmaktadır (Meijer ve Baneke, 2004; Sijtsma ve Junker, 2000).

- Özellikle 500 kişilik veri setlerinde, H_i sınır değerinin 0.3 olarak alındığı durumlarda diğer veri setlerinden farklı olarak 9 ve 10 nolu maddelerinde uyumsuz olarak belirlendiği durumlar olmuştur. Bu örneklem büyüklüğü için H_i sınır değeri 0.2 alınarak, söz konusu maddelerin ölçeklenebileceği sonuca ulaşılmaktadır. Bu bulgu da alanyazında yapılan çalışmalardan elde edilen bulgular ile uyum göstermektedir. Örneklem büyüklüğünün 1500'ün altında olduğu durumlarda, madde ölçeklenebilirlik katsayıları düşük olan maddelerin POMTK ile ölçeklenme olasılığı düşmektedir (Straat, van der Ark ve Sijtsma, 2014).
- Tüm veri setlerinin birkaç revizyonla Mokken ölçeklemeye uygun hale geldiği ve tek boyutluluk varsayımını sağladığı sonucuna ulaşılabilir.

Farklı madde sayısı ve örneklemelerden oluşan veri setleri POMTK kapsamında yapılan tek boyutluluk incelemesini takiben, POMTK'nın bir diğer varsayımı olan madde tepki fonksiyonlarının monotonluk varsayımının incelenmesine geçilmiştir.

Madde tepki fonksiyonlarının monotonluğunun incelenmesi

POMTK modelleri kapsamında oluşturulan madde tepki fonksiyonlarının ve madde adım fonksiyonlarının gizil değişken boyunca azalmayan bir fonksiyon göstermesi Kernel düzgünleştirmesi ve madde kalan puanı regresyon grafiğinin oluşturulması şeklinde iki farklı yöntemle incelenebilir. Bu aşamada madde kalan regresyon eğrileri oluşturularak farklı madde sayısı ve örneklem büyüklüklerinden oluşan veri setlerinde yer alan maddelerin monotonluğu incelenmiş ve elde edilen bulgular veri setlerindeki madde sayıları temel alınarak sunulmuştur. Öncelikli olarak beş maddelik veri setleri için incelenen madde kalan puanı regresyon eğrilerinin monotonluğu ihlal etme durumları ve söz konusu ihlallerden manidar olanlar Çizelge 27'de verilmiştir

Çizelge 27.

Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Örneklem Büyüküğü	500				1000				3000			
	H_i	Çift	İhlal	Z*	H_i	Çift	İhlal	Z*	H_i	Çift	İhlal	Z*
M1	0.27	10	1	0	0.33	10	0	0	0.28	10	0	0
M2	0.52	6	0	0	0.52	6	0	0	0.50	6	0	0

(Devam ediyor)

Çizelge 27.(devam)

Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Örneklem Büyüküğü	500				1000				3000			
	H _i	Çift	İhlal	Z *	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal	Z*
M3	0.30	6	0	0	0.33	10	0	0	0.30	6	0	0
M4	0.37	6	0	0	0.41	6	0	0	0.40	6	0	0
M5	0.38	6	0	0	0.41	6	0	0	0.37	6	0	0

Çizelge 27’de yer alan değerler incelendiğinde maddelere ilişkin H_i katsayıları, Çift olarak belirtilen ve gözlenen monotonluğun incelenmesi için örneklem büyüğüne bağlı olarak oluşturulan madde kalan puanı grup sayıları ve ihlal olarak adlandırılan madde kalan puanı regresyonunun monotonluğu ihlal ettiği noktaların sayısı görülmektedir. Ayrıca monotonluktan gerçekleşen ihlallerin ve bu ihlallerin manidarlığı ise Z testi sonuçları ile belirtilmektedir. Z testinin manidarlığındaki değerler, monotonluğun ihlalindeki kaç değer manidar olduğunu göstermektedir (Meijer, Tenderio ve Wanders, 2015). Bu açıklamalardan sonra beş maddelik veri setlerinde yer alan maddelerin monotonluk varsayımını karşılama durumları incelenebilir. Genel olarak bir değerlendirme yapılacak olursa, incelemeye alınan beş maddenin dördünde monotonluk varsayımı ihlali olmamış, yalnızca 500 kişilik veri setindeki bir nolu madde için bir ihlal belirlenmiştir, fakat bu ihlale ilişkin Z katsayısı incelendiğinde, söz konusu ihlalin manidar olmadığı belirlenmiştir. Dolayısıyla beş maddelik veri setleri için monotonluk varsayımının karşılandığı belirlenmiş ve on beş maddelik veri setleri için incelemelere geçilmiş ve bu veri setlerine ilişkin elde edilen bulgular Çizelge 28’de verilmiştir.

Çizelge 28.

On Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Örneklem Büyüküğü	500				1000				3000			
	H _i	Çift	İhlal	Z *	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal	Z*
M1	0.24	21	1	0	0.34	21	0	0	0.28	15	0	0
M2	0.44	15	0	0	0.43	15	0	0	0.42	15	0	0

(Devam ediyor.)

Çizelge 28.(devam)

On Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Örneklem Büüklüğü	500				1000				3000			
	Maddeler	H _i	Çift	İhlal	Z *	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal
M3	0.29	21	0	0	0.33	15	0	0	0.30	15	0	0
M4	0.41	15	0	0	0.41	15	0	0	0.41	15	0	0
M5	0.37	15	0	0	0.35	15	0	0	0.33	15	0	0
M6	0.34	15	0	0	0.36	15	0	0	0.34	15	0	0
M7	0.34	15	0	0	0.33	15	0	0	0.31	15	0	0
M8	0.38	15	0	0	0.42	15	0	0	0.39	15	0	0
M9	0.61	21	0	0	0.61	16	0	0	0.60	15	0	0
M10	0.28	15	1	0	0.32	15	0	0	0.29	15	0	0
M11	0.34	15	0	0	0.39	15	0	0	0.38	15	0	0
M12	0.20	15	0	0	0.23	21	0	0	0.18	15	0	0
M13	0.41	15	0	0	0.41	15	0	0	0.40	15	0	0
M14	0.42	15	0	0	0.45	15	1	0	0.41	15	0	0
M15	0.31	15	0	0	0.35	15	0	0	0.33	15	0	0

Çizelge 28’de yer alan değerler incelendiğinde, yalnızca 500 kişilik veri setinde 1 ve 10 nolu maddelerde monotonluk varsayımından ihlallerin olduğu, ancak bu ihlallere ilişkin Z katsayılarının da manidar olmadığı görülmektedir. 1000 ve 3000 kişilik veri setlerinde yer alan maddelere ilişkin oluşturulan madde kalan regresyonlarının hiçbirinde herhangi bir ihlalin olmadığı belirlenmiştir. Dolayısıyla 15 maddeden oluşan veri setleri için monotonluk varsayımının karşılandığı sonucuna ulaşılmıştır.

İncelemeye alınan son test uzunluğu koşulu olan 25 maddelik veri setleri için de madde kalan puanı regresyonları oluşturulmuş ve elde edilen istatistikler Çizelge 29’da verilmiştir.

Çizelge 29.

Yirmi Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Örneklem Büüklüğü	500				1000				3000			
	Maddeler	H _i	Çift	İhlal	Z *	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal
M1	0.26	28	2	0	0.35	21	0	0	0.30	28	0	0

(Devam ediyor.)

Çizelge 29.(devam)

Yirmi Beş Maddelik Veri Setlerinde Yer Alan Maddelerin Madde Tepki Fonksiyonlarının Monotonluğuna İlişkin Hesaplanan Değerler

Örneklem Büyüklüğü	500				1000				3000			
	Maddeler	H _i	Çift	İhlal	Z *	H _i	Çift	İhlal	Z*	H _i	Çift	İhlal
M2	0.43	28	0	0	0.41	21	0	0	0.41	28	0	0
M3	0.30	28	0	0	0.36	28	0	0	0.32	36	0	0
M4	0.42	28	0	0	0.44	28	0	0	0.44	28	0	0
M5	0.37	36	0	0	0.35	21	0	0	0.33	28	0	0
M6	0.35	28	0	0	0.38	28	0	0	0.37	36	0	0
M7	0.34	36	0	0	0.33	28	0	0	0.32	28	0	0
M8	0.36	28	1	0	0.42	28	0	0	0.41	28	0	0
M9	0.55	21	1	0	0.56	21	0	0	0.55	28	0	0
M10	0.27	36	2	0	0.34	28	0	0	0.30	28	0	0
M11	0.32	28	0	0	0.38	21	0	0	0.37	36	0	0
M12	0.22	36	1	0	0.24	21	0	0	0.20	28	0	0
M13	0.40	28	0	0	0.39	21	0	0	0.39	28	0	0
M14	0.40	28	0	0	0.42	28	0	0	0.39	28	0	0
M15	0.32	28	0	0	0.35	21	0	0	0.34	36	0	0
M16	0.39	29	1	0	0.39	28	0	0	0.40	36	0	0
M17	0.47	38	0	0	0.51	28	0	0	0.48	28	0	0
M18	0.38	28	1	0	0.40	28	0	0	0.39	21	0	0
M19	0.40	28	0	0	0.42	28	0	0	0.39	28	0	0
M20	0.30	28	3	0	0.37	28	1	0	0.34	28	0	0
M21	0.34	36	0	0	0.36	28	0	0	0.35	28	0	0
M22	0.31	28	1	0	0.37	28	1	0	0.35	28	0	0
M23	0.63	25	1	0	0.66	15	0	0	0.62	28	0	0
M24	0.32	32	2	0	0.31	28	1	0	0.33	28	0	0
M25	0.24	28	3	0	0.29	28	1	0	0.26	28	0	0

Çizelge 29 incelendiğinde, beş ve 15 maddelik test uzunluklarından elde edilen bulgulara benzer bulguların elde edildiği görülmektedir. En fazla ihlalin olduğu veri seti 500 kişilik veri setidir ve 10 madde için oluşturulan madde kalan puanı regresyon üzerinde monotonluk varsayımından ihlaller olduğu belirlenmiştir. Bu maddeler çizelgede koyu yazılarak gösterilmiştir. Ölçeklenebilirlik katsayısı sınır değer olan 0.3'ten düşük tüm maddeler için (çizelgede koyu yazılarak gösterilmiştir) ihlaller mevcutken, H_i katsayısının oldukça yüksek olduğu maddeler için de ihlallerin olduğu belirlenmiştir. Dolayısıyla bu bulgu alan yazında belirtilen H_i katsayısının yüksek olduğu durumlarda monotonluk varsayımının tamamen karşılandığının kabul edilmesi ile çelişmektedir (Junker ve Sijtsma, 2002). Belirlenen ihlallere ilişkin Z değerlerinin manidarlığı incelendiğinde ise, hiçbir

ihlalin monotonluk varsayımını ciddi düzeyde tehdit etmediği görülmekte ve veri setlerindeki ihlallerin hiçbirinin manidar olmadığı görülmektedir. Ayrıca daha önceki bulgulara benzer olarak 1000 ve 3000 kişilik veri setlerinde yer alan maddelere ilişkin ihlallerin sayısı 500 kişilik veri setinde yer alan ihlallerin sayısına göre daha azdır.

Özetle, 25 maddelik tüm veri setleri için, beş ve 15 maddelik veri setlerinde olduğu gibi, madde düzeyinden ihlaller olsa da monotonluk varsayımı karşılanmaktadır. Bu durumda oluşturulan veri setlerinin Mokken modeller kapsamında yer alan Monoton Homojenlik Modeline uyumlu olduğu sonucuna ulaşılmaktadır. Söz konusu veri setlerinin MHM'yi de kapsayan bir model olan Çift Monotonluk Modeline uyumlu olup olmadığını belirlemek için ise kısıtlayıcı bir varsayım olarak nitelendirilen (Sijtsma ve Molenaar, 2002) değişmez madde sıralaması varsayımının incelenmesine geçilmiştir.

Değişmez madde sıralaması varsayımının incelenmesi.

İncelenen veri setinde yer alan maddelerin güçlüklerine göre her yetenek düzeyinde aynı şekilde sıralanabileceğini belirten değişmez madde sıralaması varsayımı Mokken ölçekleme kapsamında yer alan Çift Monotonluk Modelini, Monoton Homojenlik Modelinden ayıran varsayımdır. Farklı test uzunluklarından ve örneklem büyüklüklerinden oluşan dokuz ayrı veri seti için bu varsayım madde kalan puanı grubu ve P matrisi yöntemi kullanılarak incelenmiştir. Öncelikli olarak madde kalan puanı yöntemine göre elde edilen sonuçlar test uzunluğu faktörüne göre sıralanarak verilmiştir.

1. Madde kalan puanı grubu yöntemi kullanılarak yapılan değişmez madde sıralaması incelemesi sonuçları

Madde çiftleri için kestirilen madde tepki fonksiyonları ile madde kalan puanı fonksiyonunun karşılaştırılmasına dayanan bu yöntem kullanılarak elde edilen beş maddelik veri setlerinde yer alan maddelere ilişkin yapılan değişmezlik incelemeleri bulguları Çizelge 30'da sunulmuştur.

Çizelge 30.

Beş Maddelik Veri Setleri İçin Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları

Örneklem Büyüklüğü	500				1000				3000			
	H _i	İhlal	Crit	Z *	H _i	İhlal	Cirt	Z*	H _i	İhlal	Crit	Z*
M1	0.27	2	79	1	0.33	0	0	0	0.28	2	85	2
M2	0.57	0	0	0	0.52	0	0	0	0.50	0	0	0
M3	0.30	2	59	0	0.33	0	0	0	0.30	2	87	2
M4	0.37	2	86	1	0.44	0	0	0	0.40	4	148	4
M5	0.38	0	0	0	0.41	0	0	0	0.37	0	0	0
Ciddi İhlal Sayısı	2				0				3			

Çizelge 30'da yer alan değerler incelendiğinde maddelerin ölçeklenebilirlik katsayıları, değişmezlikten olan ihlaller ve bu ihlallere ilişkin hesaplanan etki büyüklüğü, Crit değerleri, ile ihlallerin manidarlığını veren Z değerleri bulunmaktadır. Buna göre bulgular genel olarak incelendiğinde, 500 ve 3000 kişilik veri setlerinden benzer sonuçlar elde edildiği görülmektedir. Her iki veri setinde de 1, 3 ve 4 nolu maddeler için varsayımdan ihlaller belirlenmiş, fakat 1000 kişilik veri seti için hiçbir maddede herhangi bir ihlal bulunmamıştır. Dolayısıyla 1000 kişilik veri setinde yer alan maddelerin değişmez madde sıralaması varsayımını karşıladığı ve ÇMM kapsamında ölçeklemeye uygun olduğu sonucuna ulaşılabilir. Diğer örneklemelerin ÇMM'ye olan uygunluğunu belirlemek için ise varsayımdan olan ihlallerin Crit değerlerine bakılmalıdır. 500 kişilik veri setindeki maddelerin ihlallerine ilişkin hesaplanan Crit değerleri, incelendiğinde, 3 nolu madde için varsayımdan olan ihlallerin ihmal edilebilir düzeyde olduğu görülmekte, fakat 4 nolu madde için hesaplanan Crit değerinin 80'in üzerinde olması, bu maddenin söz konusu varsayımdan ciddi ihlalleri olduğu anlamına gelmektedir. Ayrıca 1 nolu maddenin Crit değeri sınır değere çok yakın, 79 olarak bulunmuştur ve bu durumda maddeye ilişkin ihlalin Z testi ile yapılan manidarlık incelemesi sonucu göz önünde bulundurulduğunda, bu madde için varsayımın sağlanmadığı sonucunu ortaya çıkarmaktadır. Bu bulgulara göre, beş maddelik veri setlerinden yalnızca 1000 kişilik veri seti değişmez madde sıralaması varsayımını karşılamaktadır, 500 kişilik veri setinde iki, 1000 kişilik veri setinde ise üç madde için değişmez madde sıralaması varsayımını karşılanmamıştır.

İncelenen test uzunluğu koşullarından ikincisi olan 15 maddeden oluşan veri setleri için madde kalan puanı yöntemiyle yapılan değişmez madde sıralaması bulguları Çizelge 31’te sunulmuştur.

Çizelge 31.

On Beş Maddelik Veri Setleri İçin Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları

Örneklem Büyüküğü	500				1000				3000			
	Maddeler	H _i	İhlal	Crit	Z *	H _i	İhlal	Cirt	Z*	H _i	İhlal	Crit
M1	0.24	12	88	2	0.34	2	40	1	0.28	6	74	3
M2	0.44	5	32	0	0.43	2	14	0	0.42	1	38	1
M3	0.29	10	73	1	0.33	5	48	1	0.30	7	86	5
M4	0.41	12	110	5	0.41	10	93	4	0.41	12	149	11
M5	0.37	10	56	1	0.35	2	37	1	0.33	3	41	1
M6	0.34	11	79	2	0.36	3	49	1	0.34	9	92	6
M7	0.34	14	97	3	0.33	6	59	2	0.31	5	87	4
M8	0.38	13	83	3	0.42	8	63	3	0.39	8	107	7
M9	0.61	1	4	0	0.61	1	-4	0	0.60	0	0	0
M10	0.28	41	96	3	0.32	8	51	1	0.29	7	99	6
M11	0.34	7	38	0	0.39	5	47	2	0.38	3	57	3
M12	0.20	16	132	6	0.23	9	109	6	0.18	18	191	15
M13	0.41	10	48	0	0.41	1	10	0	0.40	2	33	1
M14	0.42	6	32	0	0.45	3	39	2	0.41	1	33	1
M15	0.31	7	33	0	0.35	3	22	0	0.33	2	51	2
Ciddi İhlal Sayısı	6				2				7			

Çizelge 31’de yer alan bulgular genel olarak incelendiğinde, değişmez madde sıralaması varsayımını en iyi düzeyde karşılayan örneklemin 1000 kişilik veri seti olduğu görülmektedir. Bu örneklemden yalnızca iki madde (M4 ve M12) varsayımdan ciddi ihlaller göstermekte, altı madde ise ihmal edilebilir düzeyde ihlaller göstermektedir. Varsayıma en düşük düzeyde uyum sağlayan veri seti ise 3000 kişilik veri setidir ve bu veri setinde varsayımı ciddi düzeyde ihlal eden yedi madde yer almaktadır. Bu veri setinde yer alan farklı üç madde de söz konusu varsayımı ihmal edilebilir düzeyde ihlal etmektedir. İncelemeye alınan en küçük örneklem büyüklüğü olan 500 kişilik veri setindeki maddelerin varsayımı karşılama durumu incelendiğinde, 3000 kişilik veri setine benzer bulgular olduğu görülmektedir. Bu veri seti için ciddi ihlal gösteren madde sayısı altıdır ve bu maddelerin beşi (M4, M7, M8, M10, M12) 3000 kişilik veri setinde ihlal gösteren maddelerle aynıdır.

Elde edilen bulgular ışığında, on beş maddeden oluşan veri setleri için değişmez madde sıralaması varsayımının karşılanmadığı, dolayısıyla bu veri setleri için ÇMM'ye göre ölçekleme yapmanın olanaksız olduğu sonucuna ulaşılmış ve incelenen son test uzunluğu koşulu olan 25 maddelik veri setlerinin değişmez madde sıralaması varsayımına uyumunu incelenmiştir. Çift Monotonluk Modeline ilişkin değişmez madde sıralaması varsayımının 25 maddelik veri setlerinde karşılanma durumuna ilişkin elde edilen bulgular Çizelge 32'de verilmiştir.

Çizelge 32.

Yirmi Beş Maddelik Veri Setleri İçin Madde Kalan Puanı Grubu Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları

Örneklem Büyüklüğü	500				1000				3000			
	Maddeler	H _i	İhlal	Crit	Z *	H _i	İhlal	Cirt	Z*	H _i	İhlal	Crit
M1	0.26	19	103	5	0.35	6	52	2	0.30	11	89	6
M2	0.43	15	70	1	0.41	8	27	0	0.41	5	56	4
M3	0.30	25	83	2	0.36	8	54	1	0.32	12	79	3
M4	0.42	19	104	7	0.44	16	105	9	0.44	20	140	16
M5	0.37	19	77	3	0.35	7	52	2	0.33	7	65	3
M6	0.35	15	80	3	0.38	7	54	2	0.37	11	76	7
M7	0.34	27	87	2	0.33	9	56	2	0.32	9	89	4
M8	0.36	22	90	5	0.42	12	58	2	0.41	19	113	7
M9	0.55	7	70	3	0.56	5	76	4	0.55	3	75	3
M10	0.27	26	98	4	0.34	11	56	2	0.30	10	86	7
M11	0.32	16	72	2	0.38	8	29	0	0.37	6	56	3
M12	0.22	31	130	8	0.24	24	124	9	0.20	32	184	21
M13	0.40	21	88	2	0.39	11	61	4	0.39	10	89	9
M14	0.40	14	79	2	0.42	9	55	2	0.39	4	52	3
M15	0.32	16	57	1	0.35	9	48	2	0.34	6	71	4
M16	0.39	7	58	1	0.39	6	63	3	0.40	4	64	3
M17	0.47	7	51	2	0.51	5	77	4	0.48	8	89	8
M18	0.38	20	86	3	0.40	8	23	0	0.39	3	38	2
M19	0.40	24	104	5	0.42	12	66	5	0.39	7	84	6
M20	0.30	25	118	6	0.37	10	47	1	0.34	8	78	5
M21	0.34	16	66	1	0.36	6	46	1	0.35	7	75	4
M22	0.31	17	78	2	0.37	8	46	2	0.35	6	58	2
M23	0.63	2	-3	0	0.66	2	31	1	0.62	2	38	2
M24	0.32	12	82	3	0.31	13	104	6	0.33	9	101	8
M25	0.24	24	128	7	0.29	12	98	4	0.26	15	131	12
Ciddi İhlal Sayısı	14				4				11			

Çizelge 32’de 25 maddelik veri setleri için madde kalan puanı yöntemi kullanılarak yapılan değişmez madde sıralaması varsayımı sonucu elde edilen bulgular yer almaktadır ve bu varsayımı ciddi düzeyde ihlal eden maddeler her örneklem için koyu yazılarak gösterilmiştir. Çizelgede yer alan değerler genel olarak incelendiğinde, değişmez madde sıralaması varsayımını en iyi düzeyde karşılayan veri setinin beş ve 15 maddelik veri setlerinde olduğu gibi 1000 kişilik veri seti olduğu görülmektedir. Bu veri setindeki yalnızca dört madde (M4, M12, M24 ve M25) söz konusu varsayımı ciddi düzeyde ihlal etmekte, 17 madde ise değişmez madde sıralamasından ihmal edilebilir düzeyde ihlaller göstermektedir. Varsayımı en düşük düzeyde karşılayan veri seti ise 15 maddelik örneklemlerden farklı olarak 500 kişilik veri setidir. Bu veri setinde varsayımı ciddi düzeyde ihlal eden 14 madde yer almaktadır. Kalan 11 maddenin 10’u ihmal edilebilir düzeyde ihlaller göstermekte ve yalnızca bir madde (M23) varsayıma uyum sağlamaktadır. İncelemeye alınan son örneklem büyüklüğü olan 3000 kişilik veri setinden elde edilen sonuçlar da 500 kişilik veri setinden elde edilen bulgulara benzerdir ve bu veri seti için de 11 madde söz konusu varsayımı ciddi düzeyde ihlal etmektedir. Yine bu örneklem büyüklüğü için değişmez madde sıralamasını ihmal edilebilir düzeyde ihlal eden madde sayısı 12’dir ve yalnızca iki madde varsayıma uyumludur. Özetle, 25 maddelik veri setlerinden 500 ve 3000 kişilik veri setleri için değişmez madde sıralaması varsayımı karşılanmamıştır, 1000 kişilik veri setinde ise varsayımdan ciddi ihlallerin olduğu dört maddenin veri seti dışında tutulduğu durumda, değişmez madde sıralaması varsayımı karşılanmıştır.

Değişmez madde sıralaması varsayımı, madde kalan puanı yöntemine ek olarak P matrisi yöntemiyle de incelemeye alınmıştır. Farklı test uzunlukları ve örneklem büyüklükleri koşullarına göre oluşturulan veri setlerinin bu varsayımı karşılama durumlarına ilişkin elde edilen bulgular, test uzunluğu faktörüne göre sunulmuştur.

2. P matrisi yöntemi ile değişmez madde sıralaması varsayımının incelenmesi

Bir önceki aşamada madde kalan grubu yöntemiyle test edilen veri setlerine ilişkin değişmez madde sıralaması varsayımı bu yöntem kullanılarak da test edilmiş ve beş maddelik örneklem için elde edilen bulgular Çizelge 33’te verilmiştir.

Çizelge 33.

Beş Maddelik Veri Setleri İçin P Matrisi Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları

Örneklem Büyüküğü	500				1000				3000			
	H _i	İhlal	Crit	Z*	H _i	İhlal	Crit	Z*	H _i	İhlal	Crit	Z*
M1	0.27	1	48	1	0.33	0	0	0	0.28	0	0	0
M2	0.57	0	0	0	0.52	0	0	0	0.50	0	0	0
M3	0.30	0	0	0	0.33	0	0	0	0.30	0	0	0
M4	0.37	1	43	1	0.44	0	0	0	0.40	0	0	0
M5	0.38	0	0	0	0.41	0	0	0	0.37	0	0	0
Ciddi İhlal Sayısı	0				0				0			

Çizelge 33 incelendiğinde p matrisi yöntemiyle yapılan değişmez madde sıralaması varsayımı incelemesinde kullanılan değerler görülmektedir. Bunlar madde kalan grubuna olarak yapılan incelemeye benzer şekilde, tüm maddeler için ölçeklenebilirlik katsayısı, maddelere ilişkin oluşturulan madde tepki fonksiyonlarının varsayımı ihlal ettiği durumların sayısı, bu ihlallerin manidarlığını belirten Z testi puanları ve ihlallerin etki büyüklüğünü belirten Crit değerleridir. Elde edilen sonuçlar örneklem büyüklüklerine dayalı olarak incelendiğinde, 500 kişilik veri setinde 1 ve 4 nolu maddelerin söz konusu varsayımı ihlal ettiği görülmekte ancak bu ihlallerin etki büyüklükleri incelendiğinde ise 40 ile 80 arasında olduğu görülmektedir. Dolayısıyla söz konusu ihlallerin ihmal edilebilir düzeyde olduğu sonucuna ulaşılmaktadır. 1000 ve 3000 kişilik veri setleri için elde edilen bulgular aynıdır ve her iki veri seti için de beş maddenin de bu varsayımına uyum sağladığı görülmektedir.

İncelenen bir diğer test uzunluğu koşulu olan 15 maddelik veri setleri için de değişmez madde sıralaması varsayımı P matrisi yöntemine göre incelenmiş ve elde edilen bulgular Çizelge 34'te sunulmuştur.

Çizelge 34.

On Beş Maddelik Veri Setleri İçin P Matrisi Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları

Örneklem Büyüküğü	500				1000				3000			
	H _i	İhlal	Crit	Z *	H _i	Ihlal	Cirt	Z*	H _i	Ihlal	Crit	Z*
M1	0.24	6	58	5	0.34	0	0	0	0.28	0	0	0
M2	0.44	0	0	0	0.43	0	0	0	0.42	0	0	0
M3	0.29	4	41	3	0.33	0	0	0	0.30	1	43	1
M4	0.41	17	82	13	0.41	4	46	4	0.41	9	83	9
M5	0.37	0	0	0	0.35	0	0	0	0.33	0	0	0
M6	0.34	7	57	6	0.36	2	37	2	0.34	2	47	2
M7	0.34	1	25	1	0.33	0	0	0	0.31	0	0	0
M8	0.38	4	42	3	0.42	0	0	0	0.39	0	0	0
M9	0.61	0	0	0	0.61	0	0	0	0.60	0	0	0
M10	0.28	3	37	2	0.32	0	0	0	0.29	0	0	0
M11	0.34	1	24	1	0.39	0	0	0	0.38	0	0	0
M12	0.20	14	85	11	0.23	6	63	6	0.18	10	98	10
M13	0.41	0	0	0	0.41	0	0	0	0.40	0	0	0
M14	0.42	1	20	1	0.45	0	0	0	0.41	0	0	0
M15	0.31	2	39	2	0.35	0	0	0	0.33	0	0	0
Ciddi İhlal Sayısı	2				0				2			

Çizelge 34'te verilen bulgular örneklem büyüklüklerine göre incelendiğinde, 1000 kişilik veri setinde yer alan maddeler için hiçbir ciddi ihlalin olmadığı ve maddelerin tümünün varsayımı karşıladığı görülmektedir. 500 ve 3000 kişilik veri setlerinde ise benzer sonuçlar elde edilmiş, her iki veri setinde de 4 ve 12 nolu maddeler değişmez madde sıralaması varsayımını ciddi düzeyde ihlal etmiştir. Tüm ihlaller incelendiğinde ise, 500 kişilik veri setinde dört maddenin (M1, M3, M6 ve M8) Crit değerlerinin 40 ile 80 arasında yer aldığı görülmektedir. Dolayısıyla bu maddelerin söz konusu varsayımı ihmal edilebilir düzeyde ihlal ettiği sonucuna ulaşılmaktadır. Beş madde ise (M7, M10, M11, M14 ve M15) varsayımı düşük düzeyde ihlal etmiştir. 3000 kişilik veri setinden elde edilen sonuçlar detaylı olarak incelendiğinde ise, yalnızca iki maddenin (M3 ve M6) bu varsayımı ihmal edilebilir düzeyde ihlal ettiği görülmektedir. Özetle, on beş maddeden oluşan veri setlerinden 1000 kişilik veri seti varsayımı tamamen karşılamaktayken, 500 ve 3000 kişilik veri setlerinde yalnızca iki madde varsayımı ihlal etmekte, 13 madde ise varsayımına uyum göstermektedir.

İncelemeye alınan son test uzunluğu koşulu olan 25 maddelik veri setlerinin değişmez madde sıralaması varsayımına uyumu P matrisi yöntemiyle incelenmiş ve elde edilen sonuçlar Çizelge 35'te verilmiştir.

Çizelge 35.

Yirmi Beş Maddelik Veri Setleri İçin P Matrisi Yöntemiyle Yapılan Değişmez Madde Sıralaması İncelemesi Sonuçları

Örneklem Büyüküğü	500				1000				3000			
	H _i	İhlal	Crit	Z*	H _i	İhlal	Cirt	Z*	H _i	İhlal	Crit	Z*
M1	0.26	8	58	7	0.35	0	0	0	0.30	0	0	0
M2	0.43	3	34	3	0.41	0	0	0	0.41	0	0	0
M3	0.30	5	38	3	0.36	0	0	0	0.32	2	47	2
M4	0.42	22	74	15	0.44	8	58	8	0.44	13	83	13
M5	0.37	0	0	0	0.35	1	30	1	0.33	0	0	0
M6	0.35	7	51	6	0.38	2	35	2	0.37	2	44	2
M7	0.34	2	30	2	0.33	1	28	1	0.32	0	0	0
M8	0.36	6	50	5	0.42	1	24	1	0.41	1	39	1
M9	0.55	10	49	9	0.56	6	46	6	0.55	1	25	1
M10	0.27	10	67	9	0.34	0	0	0	0.30	0	0	0
M11	0.32	2	31	2	0.38	0	0	0	0.37	0	0	0
M12	0.22	17	77	12	0.24	9	71	9	0.20	14	97	14
M13	0.40	1	20	1	0.39	0	0	0	0.39	0	0	0
M14	0.40	3	35	3	0.42	0	0	0	0.39	0	0	0
M15	0.32	5	46	4	0.35	1	26	1	0.34	0	0	0
M16	0.39	1	21	1	0.39	0	0	0	0.40	0	0	0
M17	0.47	4	30	3	0.51	3	36	3	0.48	12	89	12
M18	0.38	4	35	3	0.40	0	0	0	0.39	0	0	0
M19	0.40	5	40	4	0.42	2	33	2	0.39	0	0	0
M20	0.30	6	46	5	0.37	0	0	0	0.34	0	0	0
M21	0.34	2	33	2	0.36	1	27	1	0.35	0	0	0
M22	0.31	1	26	1	0.37	0	0	0	0.35	0	0	0
M23	0.63	0	0	0	0.66	0	0	0	0.62	0	0	0
M24	0.32	10	60	9	0.31	8	65	8	0.33	2	65	2
M25	0.24	6	52	5	0.29	1	32	1	0.26	11	98	11
Ciddi İhlal Sayısı	0				0				4			

Çizelge 35'te verilen bulgular genel olarak incelendiğinde 500 ve 1000 kişilik veri setlerinde söz konusu varsayımı ciddi düzeyde ihlal eden herhangi bir maddenin olmadığı, 3000 kişilik veri setinde ise dört (M4, M12, M17, M25) maddenin varsayımı ciddi derecede ihlal ettiği bulunmuştur. Elde edilen değerler örneklemelere göre detaylı olarak incelenecek olursa, 500 kişilik veri setinde 11 maddenin, 1000 ve 3000 kişilik veri setlerinde ise dört

maddenin Crit değerinin 40-80 arasında olduğu belirlenmiştir. Varsayımdan ihmal edilebilir düzeyde ihlal gösteren ve çizelgede italik olarak yazılan bu maddelerin dördü, (M4, M9, M12 ve M24) 500 ve 1000 kişilik veri setlerinde ortaktır. Özetle 25 maddeden oluşan veri setleri için yalnızca 3000 kişilik veri setinde yer alan dört madde varsayıma uyum sağlamamış, diğer örneklemlerde varsayıma uyumsuzluklar olsa da ihmal edilebilir düzeyde kalmıştır.

Değişmez madde sıralaması varsayımı incelemesinde iki farklı yöntem, madde kalan puanı ve P matrisi, kullanılmış ve elde edilen sonuçlarda farklılıklar ortaya çıkmıştır. Her iki yöntemden elde edilen sonuçları birlikte incelemek için, örneklemlere ve her iki yönteme göre ciddi ihlal gösteren madde sayıları belirlenmiş ve madde sıralamalarının netliğini belirten H^T katsayıları hesaplanmıştır. Tüm bu değerler Çizelge 36'da sunulmuştur.

Çizelge 36.

Farklı Test Uzunlukları ve Örneklem Büyüklüğünden Oluşan Veri Setleri İçin Hesaplanan H^T Katsayıları

Test Uzunluğu		5			15			25		
Örneklem Büyüklüğü		500	1000	3000	500	1000	3000	500	1000	3000
Ciddi İhlal Sayısı	Madde Kalan Puanı	2	0	3	6	2	7	14	4	11
	P Matrisi	0	0	0	2	0	2	0	0	4
H^T Katsayısı		0.260	0.254	0.261	0.247	0.280	0.234	0.311	0.315	0.313

Çizelge 36'da yer alan H^T katsayıları varsayımı ihlal eden madde sayıları ile birlikte incelendiğinde, yalnızca 25 maddelik veri setlerinde H^T katsayılarının 0.3'ün üzerinde değer aldığı görülmekte, 5 ve 15 maddelik veri setlerinde ise varsayımı ihlal eden madde sayısı 25 maddelik veri setlerine göre daha düşük olmasına rağmen, H^T katsayılarının 0.3'ün altında olduğu bulunmuştur. Ayrıca Çizelge 36 incelendiğinde, yöntemlere göre elde edilen sonuçların çok farklı olduğu da görülmektedir. Madde kalan puanı yöntemi kullanılarak belirlenen uyumsuz madde sayısı P matrisi kullanılarak belirlenen uyumsuz madde sayısından çok daha fazladır ve böyle durumlarda alan yazında H^T katsayısından elde edilen bilgilerin kullanılması daha uygun görülmektedir. Bu durumun nedeni olarak da, manidar olarak belirlenen ihlallerin bazı durumlarda etkilerinin az olduğu ve manidarlık düzeyinin örnekleme göre değiştiği belirtilmektedir. Ayrıca bu durumdan kaynaklanan

yanlı kestirimler her ne kadar Crit değerleri ile ortadan kaldırılmaya çalışılsa da, Crit değerleri için yapılan herhangi bir simülasyon çalışması olmadığı için, ihlallerin etki büyüklüğü hakkında bilgi veren bu değerlerinde dikkatle yorumlanması ve farklı bulgularla birleştirilmesi gerektiği ifade edilmektedir (Molenaar ve Sijtsma, 2002; Sijtsma ve Meijer, 1992).

Çizelge 36'da yer alan bulgular bu bilgilere yorumlanmış ve değişmez madde varsayımı, bir diğer ifadeyle, kesişmeyen madde tepki fonksiyonları varsayımının yalnızca 25 maddelik veri setlerinde düşük düzeyde karşılandığı sonucuna ulaşılmıştır. Bu analizle de oluşturulan veri setlerinin POMTK kapsamında ölçeklenebilirliğine ilişkin son işlem tamamlanmıştır.

Farklı test uzunluklarından ve örneklem büyüklüklerinden oluşan veri setleri için yapılan POMTK analizlerin sonuçları özetlenecek olursa, şu bulgulara ulaşılabilir:

- POMTK analizleri kapsamında yer alan ilk koşul olan H_{ij} katsayılarının pozitif olması yalnızca iki veri setindeki (15 ve 25 maddeden oluşan 500 kişilik veri setleri) 1 ve 12 nolu madde için sağlanmamış, diğer veri setlerinde yer alan tüm maddeler için hesaplanan ve madde çiftleri arasındaki kovaryansı veren H_{ij} katsayıları pozitif olarak bulunmuştur.
- Tüm farklı madde sayısı koşullarında 1000 kişilik veri setlerindeki maddeler için hesaplanan H_i değerleri daha yüksek bulunmuş, dolayısıyla bu veri setleri için hesaplanan H değerleri de daha yüksek bulunmuştur.
- Farklı veri setlerinde 1 ve 12 nolu maddelerin H_i değerleri ortak olarak 0.3'ten düşük bulunmuştur. Dolayısıyla bu maddelerin veri seti dışında tutulması POMTK'ya uyumu artırır.
- Tüm ölçeklenebilirlik katsayıları için hesaplanan standart hata değerleri düşüktür ve madde ve örnekleme yer alan birey sayısındaki artışın standart hata değerlerini düşürdüğü sonucuna ulaşılmıştır. Daha önce de belirtildiği üzere bu bulgu alan yazında yer alan bulgular ile paralellik göstermektedir. Örnekleme yer alan birey sayısındaki artış, özellikle 200'ün üzerinde olması, ölçeklenebilirlik katsayılarının standart hata değerlerini düşürmektedir (Kuijpers, van der Ark, Croon ve Sijtsma, 2016).
- OMSİ kullanılarak yapılan tek boyutluluk incelemesinde madde sayılarına ve örneklemlere göre uyumlu madde sayısı genel olarak benzer olmakla birlikte, en fazla

uyumsuz maddenin, tüm farklı madde sayısı koşullarında, 500 kişilik veri setlerinde yer aldığı bulunmuştur.

- Tek boyutluluk varsayımını sağlamayan maddeler genel olarak benzerlik göstermekte ve 12 ile 25 nolu maddenin birden çok veri setinde kullanılan iki yonteme göre de ortak olarak uyumsuz olarak belirlenmiştir. Söz konusu maddeler araştırma kapsamında daha önce incelenen 30 maddelik veri setlerinde de OMSİ analizine göre uyumsuz olarak buldukları için, bu maddelerin Mokken ölçekleme tekniğine uyum sağlamadığı sonucuna ulaşılabilir ve ölçek dışında tutulması gerekmektedir. Elde edilen bu sonuç alanyazında yer alan bulgulara uyum sağlamaktadır. OMSİ analizi testteki madde sayısından ihmal edilebilir düzeyde etkilenmektedir (Straat, van der Ark ve Sijtsma, 2014)
- Oluşturulan tüm veri setlerinde yer alan madde tepki fonksiyonlarının monoton azalmayan niteliktedir, dolayısıyla bu varsayım da tüm veri setleri için karşılanmıştır.
- Çift Monotonluk Modelinin varsayımı olan değişmez madde sıralaması/kesişmeyen madde tepki fonksiyonları varsayımı madde kalan puanı ve P matris yöntemiyle incelenmiştir. Elde edilen sonuçlar MTF'ler arasındaki uzaklığın ölçüsünü veren H^T katsayısıyla birleştirilerek yorumlandığında yalnızca 25 maddelik veri setlerinin bu varsayımı düşük düzeyde karşıladığı belirlenmiştir
- İncelenen veri setlerinden beş ve 15 maddelik olanlar için Monoton Homojenlik Modelinin, 25 maddelik olanlar için ise Çift Monotonluk Modelinin uygun olduğu belirlenmiştir.

Farklı Örneklem Büyüklüğüne ve Test Uzunluğuna Sahip Veri Setlerinden PMTK ve POMTK Kapsamında Kestirilen Madde Parametrelerinin Karşılaştırılması

Araştırma kapsamında yanıt aranan ikinci araştırma sorusu, incelemeye alınan kitapçıkta yer alan maddelere ilişkin parametrik ve parametrik olmayan madde tepki kuramı modellerinden uygun olan modele göre kestirilen madde ayırt edicilik ve güçlük parametrelerinin örneklem büyüklüğüne ve test uzunluğuna göre farklılaşp farklılaşmadığının belirlenmesidir. Bu araştırma sorusuna yanıt aramak için öncelikli olarak farklı örneklem büyüklüklerine sahip 30 maddeden oluşan veri setleri için PMTK'ya dayalı madde analizleri yapılmış ve elde edilen parametreler farklı örneklere göre

karşılaştırılmıştır. Bu incelemenin ardından her iki yaklaşımdan kestirilen madde parametrelerine ilişkin incelemelere geçilmiştir.

500, 1000 ve 3000 kişilik veri setleri için PMTK'ya göre kestirilen madde parametreleri.

Çalışma kapsamından oluşturulan temel veri setinden seçkisiz olarak belirlenen 500, 1000 ve 3000 kişilik veri setlerinin PMTK modellerine uyumu incelendiğinde, tüm veri setleri için en iyi model veri uyumunun iki parametrelili lojistik modelde sağlandığı belirlenmiştir. Söz konusu model kullanılarak her veri seti için madde parametre kestirimleri yapılmış ve elde edilen sonuçlar Çizelge 37'de verilmiştir.

Çizelge 37.

500, 1000 ve 3000 Kişilik Veri Setleri İçin Uygun PMTK Modellerinden Kestirilen Madde Parametreleri

Örneklem Büyüküğü	500		1000		3000	
	a	b	a	b	a	b
M1	0.931	-1.152	1.341	-0.992	1.090	-1.122
M2	1.732	0.367	1.590	0.423	1.601	0.363
M3	1.214	-0.868	1.440	-0.871	1.236	-0.951
M4	<u>2.367</u>	-0.662	2.204	-0.601	2.374	-0.695
M5	1.610	-0.237	1.407	-0.256	1.289	-0.259
M6	1.490	-0.748	1.528	-0.858	1.523	-0.843
M7	1.420	-0.456	1.267	-0.661	1.259	-0.66
M8	1.716	-0.441	1.899	-0.419	1.941	-0.499
M9	2.313	0.778	2.405	0.725	2.335	0.81
M10	1.039	-0.562	1.308	-0.65	1.104	-0.77
M11	1.047	0.61	1.352	0.506	1.287	0.538
M12	0.756	<u>-1.184</u>	<u>0.787</u>	<u>-1.181</u>	<u>0.655</u>	<u>-1.295</u>
M13	1.877	-0.133	1.553	-0.031	1.653	-0.073
M14	1.605	0.377	1.687	0.397	1.549	0.334
M15	1.346	-0.333	1.372	-0.318	1.424	-0.405
M16	1.671	-0.996	1.674	-0.916	1.718	-1.002
M17	1.841	0.728	2.008	0.78	1.878	0.694
M18	1.587	0.033	1.600	0.054	1.531	0.096

(Devam ediyor.)

Çizelge 37.(devam)

500, 1000 ve 3000 Kişilik Veri Setleri İçin Uygun PMTK Modellerinden Kestirilen Madde Parametreleri

Örneklem Büyüküğü	500		1000		3000	
	a	b	a	b	a	b
M19	1.965	-0.269	2.084	-0.242	1.769	-0.264
M20	1.121	0.092	1.456	0.135	1.308	0.023
M21	1.457	-0.528	1.505	-0.503	1.460	-0.53
M22	1.229	0.115	1.486	0.096	1.375	0.048
M23	2.335	1.209	2.780	1.089	2.496	1.111
M24	0.979	1.229	0.917	1.172	0.998	1.120
M25	0.730	1.022	0.892	1.003	0.784	1.230
M26	1.215	0.285	1.382	0.24	1.229	0.213
M27	1.474	-0.330	1.736	-0.271	1.505	-0.314
M28	1.359	-0.808	1.473	-0.76	1.412	-0.844
M29	1.729	0.367	1.708	0.232	1.593	0.279
M30	1.447	-0.071	1.236	-0.122	1.265	-0.101
Ortalama	1,468	-0.085	1,569	-0.182	1,488	-0.125
(ss)	(0.431)	(0.677)	(0.424)	(0.654)	(0.422)	(0.696)

Çizelge 37’de yer alan ve testteki her madde için hesaplanan değerlerden ilki a parametresi olarak belirtilen madde ayırt edicilik parametresidir. Madde ayırt edicilik parametrelerinin her örneklem büyüğü için hesaplanan ortalamalarına bakıldığında, tüm örneklem için 1’in üzerinde değerler olduğu, en yüksek ortalamanın 1000 kişilik veri seti için hesaplandığı, fakat diğer örneklem için de birbirine yakın ve yüksek değerler hesaplandığı görülmektedir.

Kestirilen madde ayırt edicilik parametreleri örneklemere göre madde düzeyinde incelenirse, 500 kişilik veri seti için çoğu maddenin 1’in üzerinde ve üç maddenin ise 2’nin üzerinde madde ayırt edicilik değerine sahip olduğu görülmektedir. Bu parametreye göre en düşük ve en yüksek değerlerin sırasıyla 25 (0.730) ve 4 (2.367) nolu maddelerde olduğu belirlenmiştir. En düşük değere sahip 25 nolu maddeye ilişkin a parametresi 0.730 olarak belirlenmiştir ve bu değere göre söz konusu madde orta düzeyde ayırt edici olarak

nitelendirilebilir (De Ayala, 2009). Kestirilen madde ayırt edicilik değerleri 1000 kişilik veri seti için incelenecek olursa, beş madde için 2'den yüksek madde ayırt edicilik parametre değeri hesaplandığı belirlenmiş ve bu maddeler Çizelge 37'de koyu yazılarak belirtilmiştir. Söz konusu maddelerin üçü için 500 kişilik veri setinde de 2'nin üzerinde a parametresi değerleri kestirilmiştir. Bu maddeler M4, M9, M23, M17 ve M19'tur. Ortalama değerlerden de görülebileceği gibi, 1000 kişilik veri seti için a parametreleri 500 kişilik veri setine göre daha yüksek kestirilmiştir. En yüksek ve en düşük değerler de sırasıyla, M23 (2.780) ile M12 (0.787) maddelerindedir. En yüksek madde ayırt edicilik değeri 3'e oldukça yakındır ve düşük olarak belirtilen 0.787 değeri de orta düzeyde ayırt edici olarak sayılabilecek değerdir.

Çalışma kapsamında incelenen son örneklem büyüklüğü 3000 kişilik veri setidir. Bu veri setinde yer alan maddelere ilişkin kestirilen ayırt edicilik parametreleri incelendiğinde, üç maddenin 2'nin üzerinde ve çok sayıda maddenin ise 1'in üzerinde değerler aldığı belirlenmiştir. İkincinin üzerinde a parametresi değeri alan maddeler 500 kişilik veri setindeki maddelerle aynıdır ve en yüksek değer 2,496 olarak 23 nolu maddeden kestirilmiştir. En düşük madde ayırt edicilik parametresi ise 12 nolu madde için 0.655 olarak kestirilmiştir.

Çizelge 37'de yer alan ve örneklemlere göre kestirilen diğer parametre madde güçlük indeksidir. Madde güçlük parametreleri genel olarak incelendiğinde tüm örneklemlerde sıfıra yakın değerlerin çoğunlukta olduğu görülmektedir. Örneklemlere göre hesaplanan b parametresi ortalamaları da bu bulguyu desteklemektedir, tüm ortalama değerler negatif fakat sıfıra oldukça yakındır. Dolayısıyla test kapsamındaki maddelerin kolay olduğu sonucuna ulaşılabılır. Madde güçlük parametreleri oluşturulan veri setlerine göre ayrı olarak incelenecek olursa, 500 kişilik veri seti için 1'in üzerinde madde güçlük indeksi değerine sahip olan madde sayısı üçtür ve bu maddeler çizelgede koyu yazılmıştır. Bu örneklem büyüklüğü için kestirilen en düşük madde güçlük parametresi 12 nolu madde için -1.184 olarak, en yüksek madde güçlük parametresi ise 24 nolu madde için 1.229 olarak hesaplanmıştır.

Bin kişilik veri seti için kestirilen madde güçlük parametreleri incelendiğinde, 1'den yüksek b parametresine sahip olan maddelerin 500 kişilik veri setindeki maddeler ile aynı olduğu gözlenmiştir ve çizelgede koyu yazılarak gösterilmiştir. Aynı şekilde en yüksek b parametresi de 500 kişilik veri setinde olduğu gibi 24 nolu madde için 1,172 olarak

kestirilmiştir. En düşük b parametresi 12 nolu madde için -1,181 olarak kestirilmiş ve bu sonuç da 500 kişilik veri setinden hesaplanan sonuçla aynıdır. İncelenen son örneklem büyüklüğü olan 3000 kişilik veri setinden hesaplanan madde güçlük parametreleri incelendiğinde ise 1'den büyük değer alan üç madde olduğu ve bu maddelerin diğer örneklemdeki maddelerle aynı olduğu bulunmuştur. Söz konusu maddeler çizelgede koyu yazılarak gösterilmiştir. En büyük b parametresi 23 nolu madde için 1,230 olarak, en düşük b parametresi de 12 nolu madde için -1,295 olarak kestirilmiştir. Farklı örneklem büyüklüklerine göre kestirilen b parametrelerine ilişkin elde edilen bulgular özetlenecek olursa, -1 ve 1 aralığında değiştiği, her üç örnekleme de 1'in üzerinde değer alan üç madde olduğu ve bunların aynı maddeler olduğu ve madde güçlük parametresi ortalamalarının sıfıra yakın olduğu sonuçlarına ulaşılabilir.

Parametreler arasındaki ilişkiyi daha iyi incelemek için her üç örneklemden elde edilen madde parametreleri arasındaki korelasyonlar incelenmiştir. İlk analiz madde ayırt edicilik parametreleri arasındaki ilişkinin incelenmesidir. Yapılan normallik incelemesinde değişkenlerin tümü için normal dağılım sağlanmadığı için, parametreler arası ilişkilerin incelenmesinde Spearman Brown Korelasyon Katsayısı kullanılmış ve elde edilen değerler Çizelge 38'de sunulmuştur.

Çizelge 38.

Farklı Örneklem Büyüklüklerinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar

Örneklem Büyüklüğü		
Spearman Brown KK	1000	3000
500	0.880**	0.938**
1000		0.946**

**p<0.01

Çizelge 38'de 500, 1000 ve 3000 kişilik veri setlerinden kestirilen madde ayırt edicilik parametreleri arasındaki korelasyonlar verilmektedir. Çizelge incelendiğinde tüm örneklem büyüklüklerinden hesaplanan madde ayırt edicilik parametreleri arasında manidar ve yüksek bir ilişki olduğu görülmektedir. Çizelgede yer alan değerler incelendiğinde, 1000 ve 3000 kişilik veri setlerinden kestirilen değerlerin benzer olduğu, 500 kişilik veri setinden elde edilen a parametrelerinin ise 1000 kişilik veri setine göre, 3000 kişilik veri setinden elde edilen değerlere daha yakın olduğu sonucuna ulaşılmıştır. Farklı bireylerden ve farklı

örneklem büyüklüklerinden kestirilen madde parametreleri arasındaki ilişkinin yüksek ve manidar oluşu, madde değişmezliğine ilişkin elde edilmiş bir kanıt olarak kullanılabilir (Emberston ve Reise, 2000).

Çizelge 37’de yer alan ve her madde için kestirilen bir diğer parametre ise b parametresi olarak belirtilen madde güçlük parametresidir. Örneklemelere göre kestirilen madde güçlük parametreleri genel olarak incelendiğinde 1000 ve 3000 kişilik veri setlerinden elde edilen b parametrelerinin 500 kişilik veri setine göre birbirlerine daha yakın ve benzer değerler ürettiği gözlenmiştir. Örneklemelere göre elde edilen madde güçlük parametreleri arasındaki ilişkiyi daha net inceleyebilmek için kestirilen parametreler arasındaki ilişki, değişkenler normal dağılıma uyum gösterdiği için Pearson Momentler Çarpımı Korelasyon Katsayısı ile incelenmiş ve sonuçlar Çizelge 39’da verilmiştir.

Çizelge 39.

Farklı Örneklem Büyüklüklerinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Pearson MÇKK	Örneklem Büyüklüğü	
	1000	3000
500	0.994**	0.994**
1000		0.996**

**p<0.01

Çizelge 39’da verilen korelasyon katsayıları incelendiğinde tüm örneklemden kestirilen b parametreleri arasında manidar ve çok yüksek ilişkiler olduğu görülmektedir. Bu değerler incelendiğinde, tüm örneklemelerden hesaplanan b parametrelerinin mükemmel yakın ilişkilere sahip olduğu sonucuna ulaşılmıştır. Farklı örneklem büyüklüklerinden ve bireylerden oluşan veri setlerinden kestirilen madde güçlük parametreleri arasındaki ilişkinin çok yüksek ve manidar olması, PMTK kapsamında hem model veri uyumuna kanıtı olarak kullanılabilir hem de madde güçlük parametrelerinin örneklem büyüklüğünden etkilenmediği şeklinde yorumlanabilir.

PMTK modellerinden 2 PLM kullanılarak elde edilen madde parametreleri kendi içinde karşılaştırılmıştır. Yapılan incelemeler sonucunda, seçkisiz olarak oluşturulan 500, 1000 ve 3000 kişilik veri setlerinden elde edilen kestirimler ve madde parametreleri arasındaki ilişkiler şu şekilde özetlenebilir:

- Madde düzeyinde en yüksek model veri uyumu 500 ve 1000 kişilik veri setleri için sağlanmıştır. Bu örneklem için iki ve üç parametrelili modellerin uyumlu olduğu belirlenmiştir fakat iki parametrelili lojistik model kullanılarak kestirimler yapılmıştır.
- İncelemeye alınan en büyük örneklem olan 3000 kişilik veri seti için madde düzeyinde uyumlu bir parametrik model bulunamamış ancak model düzeyinde yapılan incelemelerde 2 PLM'nin uyumlu olduğu belirlenmiştir. Bu durum madde düzeyindeki uyum istatistiklerinin ki-kare testine dayalı olarak hesaplanması ve bu testin de örneklem büyüklüğüne duyarlı olmasından kaynaklanmaktadır (McDonald ve Moon, 2016; Thompson, 2000).
- Farklı örneklem büyüklüğünden oluşan veri setlerinden kestirilen parametreler arasındaki ilişkiler incelendiğinde, a parametresine göre b parametreleri arasındaki ilişkilerin daha yüksek olduğu belirlenmiştir. Bu durum da a parametresinin örneklem büyüklüğüne daha duyarlı olduğu sonucunu ortaya çıkarmaktadır.
- Tüm veri setlerinden kestirilen madde parametreleri arasında çok yüksek ve manidar bir uyum olduğu belirlenmiştir. Dolayısıyla PMTK modelleri ile yapılan madde parametre kestirimlerinin örneklem büyüklüğünden etkilenmediği sonucuna ulaşılabilir.

Yukarıda verilen bulgulardan sonra, PMTK kapsamında yer alan iki parametrelili lojistik modelden elde edilen kestirimlerin POMTK kapsamında yer alan modellerden elde edilen parametrelerin kıyaslanmasına geçilmiştir. Bu amaçla oluşturulan 500, 1000 ve 3000 kişilik veri setlerinin POMTK kapsamında incelenmesine geçilmiştir.

Farklı Örneklem Büyüklüklerinden Parametrik ve Parametrik Olmayan Modellere Göre Kestirilen Madde Parametrelerinin Karşılaştırılması

PMTK modellerinden iki parametrelili lojistik modele uyum sağlayan ve bu modele göre madde kestirimleri yapılan 500, 1000 ve 3000 kişilik veri setlerinin POMTK modellerinden Monoton Homojenlik Modeline uyum sağladığı belirlenmiştir. İki parametrelili lojistik modelde maddeler için hesaplanan a parametreleri, MHM'de maddeler için hesaplanan H_i değerlerine karşılık gelmektedir. Her iki model kapsamında elde edilen madde ayırt edicilikleri kıyaslanarak kestirimlerin benzerlik ve farklılıklarını bulmak

amaçlanmış ve bu amaçla her iki modelden elde edilen madde ayırt edicilik indeksleri arasındaki korelasyonlar incelenmiş Çizelge 40'ta verilmiştir.

Çizelge 40.

İki Parametrelili Lojistik Modelden ve Monoton Homojenlik Modelinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar

Sperman Brown KK	PMTK			POMTK	
	500A	1000A	3000A	500H	1000H
1000A	0.880**				
3000A	0.938**	0.946**			
500H	0.958**	0.845**	0.890**		
1000H	0.821**	0.945**	0.899**	0.837**	
3000H	0.867**	0.904**	0.945**	0.894**	0.942**

**p < 0.01

Çizelge 40'ta, 500, 1000 ve 3000 kişilik veri setleri için PMTK ve POMTK kapsamında hesaplanan madde ayırt edicilik parametreleri arasındaki korelasyonlar yer almaktadır. 500A, 1000A ve 3000A olarak belirtilen değerler, 500, 1000 ve 3000 kişilik veri setlerinden iki parametrelili lojistik model kapsamında kestirilen madde ayırt edicilik parametrelerini belirtirken, 500H, 1000H ve 3000H olarak belirtilen değerler de Monoton Homojenlik Modeli kapsamında hesaplanan madde ayırt edicilik parametreleridir. PMTK kapsamında yapılan kestirimlerde, farklı örneklem büyüklüklerine göre madde ayırt edicilik parametreleri arasındaki ilişkiler incelenmiş ve tüm örneklemelerden kestirilen parametreleri arasındaki ilişkinin yüksek düzeyde ve manidar olduğu bulunmuştur. Söz konusu ilişkiler daha önce incelendiği için bu aşamada öncelikli olarak MHM modelinden kestirilen H_i katsayıları arasındaki ilişkiler incelenmiş, ardından da her iki modelin madde ayırt edicilik parametreleri kestirimleri arasındaki ilişkiler yorumlanmıştır.

POMTK modeli olan MHM kapsamında 500, 1000 ve 3000 kişilik veri setlerinden hesaplanan H_i katsayıları incelendiğinde, 500 kişilik veri setinden kestirilen H_i katsayıları ile 1000 ve 3000 kişilik veri setlerinden kestirilen H_i katsayıları arasında pozitif, 0.01 düzeyinde manidar ve yüksek ilişkiler (sırasıyla 0.837 ve 0.894) olduğu belirlenmiştir. 1000 ve 3000 kişilik veri setlerinden kestirilen H_i katsayıları arasında da yüksek (0.942) ve 0.01 düzeyinde manidar ilişkiler olduğu belirlenmiştir. Bu sonuçlar gözetildiğinde, MHM kapsamında tüm örneklem büyüklükleri için kestirilen H_i katsayıları arasında yüksek düzeyde ve manidar ilişkiler olduğu sonucuna ulaşılmaktadır. Dolayısıyla, PMTK

modellerinde olduğu gibi, örneklem büyüklüklerindeki değişikliklerin POMTK kapsamında kestirilen madde ayırt edicilik parametreleri üzerinde etkisinin olmadığı yorumu yapılabilmektedir.

Çizelge 40'ta yer alan değerler parametrik ve parametrik olmayan MTK modelleri kapsamında incelendiğinde ise, genel olarak tüm örneklemelerden kestirilen madde ayırt edicilik parametreleri arasında yüksek ve manidar ilişkiler olduğu görülmektedir. 500 kişilik veri seti için hesaplanan madde ayırt edicilik parametrelerinin her iki modelde de yüksek (0.958) ve manidar ilişki gösterdiği belirlenmiştir. Beş yüz ve 1000 kişilik veri setleri arasındaki ilişki 0.845 olarak hesaplanırken, 500 ve 3000 kişilik veri setleri arasındaki ilişki 0.890 olarak hesaplanmıştır. PMTK ve POMTK farkı olmaksızın, her iki model için de (2PLM ve MHM), 500 ve 3000 kişilik veri setleri arasındaki ilişki, 500 ve 1000 kişilik veri setleri için hesaplanan ilişkiden daha yüksek bulunmuştur. İncelenen bir diğer örneklem büyüklüğü olan 1000 kişilik veri setleri için POMTK ve PMTK kapsamında kestirilen madde ayırt edicilik parametreleri arasındaki ilişkiler incelendiğinde ise, en yüksek ilişkinin 1000 kişilik veri setinden (0.945), en düşük ilişkinin 500 kişilik veri setinden (0.821) hesaplandığı görülmektedir. Üç bin kişilik veri setine olan korelasyonunun ise 0.899 olarak hesaplandığı görülmektedir. Dolayısıyla 1000 kişilik veri setine ilişkin parametrik ve parametrik olmayan MTK modellerinden kestirilen madde ayırt edicilikleri arasında yüksek ilişkiler olduğu sonucuna ulaşılmış, özellikle aynı örneklem büyüklüğü için her iki modelden hesaplanan madde ayırt edicilik parametrelerinin birbiri ile daha yüksek ilişkiler gösterdiği belirlenmiştir.

Örneklem büyüklüğü faktörü kapsamında incelenen son koşul 3000 kişilik veri setidir. Bu örneklem büyüklüğü için hesaplanan madde ayırt edicilik parametreleri 2 PLM ve MHM'ye göre incelenecek olursa, en yüksek ilişkinin 3000 kişilik veri setinden kestirilen parametreler arasında (0.945) olduğu görülmektedir. Bu örneklem ile 1000 kişilik veri setinden parametrik yaklaşımla kestirilen madde ayırt edicilik parametreleri arasındaki ilişki de yüksek 0.904, ve manidardır. MHM kapsamında 3000 kişilik veri setinden kestirilen H_i katsayıları ile 1000 kişilik veri setinden kestirilen a parametreleri arasındaki ilişki ise, 0.867 olarak bulunmuş ve bu değer diğer örneklemelerden elde edilen katsayılarla göre daha düşük olmasına rağmen yine de yüksek bir ilişkiyi göstermektedir. PMTK kapsamında 2 PLM ve POMTK kapsamında ur.MHM ile farklı örneklem büyüklüklerinden

hesaplanan madde ayırt edicilik değerleri arasındaki ilişkileri özetlemek gerekirse, şu sonuçlara ulaşılabilir;

- En yüksek ilişki 500 kişilik veri setlerinden kestirilen parametreler arasında bulunmuş ve 0.958 olarak hesaplanmıştır.
- En düşük ilişki 500 ve 1000 kişilik veri setleri arasında 0.821 olarak hesaplanmıştır.
- Tüm örneklem büyüklüklerinden PMTK ve POMTK modelleri kapsamında hesaplanan ilişkiler yüksek ve manidardır.
- Tüm örneklem büyüklükleri için en yüksek ilişki aynı örneklem büyüklüğünden hesaplanmıştır. Bir diğer ifadeyle, 500 kişilik veri seti için PMTK kapsamında kestirilen a parametreleri en yüksek ilişkiyi yine 500 kişilik veri setinden kestirilen H_i değerleri ile göstermiştir. Aynı durum 1000 ve 3000 kişilik veri setleri için de geçerlidir.
- Hesaplanan tüm ilişkiler manidar ve yüksek olduğu için, söz konusu örneklem büyüklükleri için POMTK kapsamında kestirilen H_i değerlerinin PMTK kapsamında hesaplanan a parametreleriyle eşdeğer olarak kullanılabilceği sonucuna ulaşılmıştır.

Farklı örneklem büyüklükleri için 2 PLM ve MHM kullanılarak yapılan analizlerde, maddelere ilişkin bir diğer parametre olan madde güçlük parametreleri de kestirilmiştir. İki parametrelili lojistik modelden farklı örneklem için kestirilen madde güçlük parametreleri arasındaki ilişki daha önce incelenmiş ve tüm örneklemekten kestirilen b parametrelerinin yüksek ve manidar ilişkiler gösterdiği belirlenmiştir. Bu aşamada ise 2 PLM kapsamında hesaplanan madde güçlük parametrelerinin MHM kapsamında elde edilen madde güçlük sıralaması arasındaki ilişki Spearman Brown Sıra Farkları Korelasyon Katsayısı kullanılarak incelenmiş ve sonuçlar Çizelge 41’de sunulmuştur.

Çizelge 41.

İki Parametrelili Lojistik Modelden ve Monoton Homojenlik Modelinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Sperman Brown KK	PMTK			POMTK	
	500B	1000B	3000B	500	1000
1000B	0.994**				
3000B	0.994**	0.996**			

(Devam ediyor.)

Çizelge 41. (devam)

İki Parametrelili Lojistik Modelden ve Monoton Homojenlik Modelinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Sperman Brown KK	PMTK			POMTK	
	500B	1000B	3000B	500	1000
500	-0.979**	-0.974**	-0.968**		
1000	-0.977**	-0.981**	-0.974**	0.993**	
3000	-0.977**	-0.977**	-0.976**	0.994**	0.996**

**p<0.01

Çizelge 41’de 500B, 1000B ve 3000B olarak belirtilen değerler PMTK’den elde edilen madde güçlük indeksleri iken, 500, 1000 ve 3000 olarak belirtilen değerler ise Monoton Homojen Modeli kapsamında yapılan incelemelerden elde edilen sonuçları göstermektedir. Öncelikle MHM’de yapılan kestirimler kendi içinde incelenecek olursa, tüm örneklemelerden kestirilen madde güçlük parametrelerinin sıralamasının neredeyse aynı olduğu sonucuna ulaşılmıştır. Dolayısıyla POMTK kapsamında kestirilen madde güçlük parametrelerinin örneklem büyüklüğünden etkilenmediği, birbirine çok benzer sonuçlar verdiği söylenebilir. POMTK ile PMTK’da kestirilen madde güçlük parametreleri arasındaki korelasyonlar incelendiğinde, öncelikle tüm değerlerin negatif işaretli olduğu görülmektedir. Bu ilişkinin negatif işaretli olması her iki yaklaşımdaki madde güçlük indeksinin tanımından kaynaklanmaktadır. PMTK kapsamında madde güçlük indeksi değeri büyüdükçe madde güçleşmekteyken, POMTK’da Klasik Test Kuramında olduğu gibi, madde güçlük indeksi değeri arttıkça madde kolaylaşmaktadır. Bu nedenle MHM ve 2 PLM kullanılarak yapılan kestirimler arasındaki ilişkinin negatif işaretli olması beklenen bir durumdur (Sijtsma, ve Meijer, 1992). Buna göre Çizelge 41’de yer alan değerler incelendiğinde tüm örneklem büyüklüklerinden MHM kapsamında kestirilen madde güçlük indeksi değerlerinin, parametrik yaklaşımla kestirilen değerlerle manidar ve yüksek bir ilişki gösterdiği görülmektedir. Dolayısıyla örnekleme yer alan birey sayısı fark etmeksizin, PMTK ve POMTK modellerinden kestirilen madde güçlük parametrelerinin çok yakın değerler olduğu belirlenmiş ve eşdeğer olarak kullanılacakları sonucuna ulaşılmıştır.

PMTK kapsamında 2 PLM ve POMTK kapsamında MHM ile farklı örneklem büyüklüklerinden hesaplanan madde güçlük indeksleri arasındaki ilişkiyi özetlemek gerekirse, şu sonuçlara ulaşılabilir;

- PMTK kapsamında farklı örneklem büyüklüklerinden hesaplanan tüm madde güçlük indeksleri arasında mükemmel yakın ve manidar ilişkiler vardır.
- POMTK kapsamında tüm örneklemelerden hesaplanan madde güçlük indeksleri arasında çok yüksek ve manidar ilişki vardır.
- POMTK ve PMTK kapsamında tüm örneklemelerden hesaplanan madde güçlük indeksi değerleri arasında +1.00'e yakın ve manidar bir ilişki vardır.

Bu aşamaya kadar çalışma kapsamındaki ikinci araştırma sorusunun ilk maddesi olan farklı örneklem büyüklüklerine göre PMTK ve POMTK modellerine göre yapılan madde parametre kestirimleri karşılaştırılmış ve aynı araştırma sorusunun ikinci maddesi olan farklı madde sayılarına göre PMTK ve POMTK modelleri kullanılarak yapılan madde parametre kestirimleri karşılaştırılmasına geçilmiştir.

Farklı Test Uzunluklarına Göre Oluşturulan Veri Setlerinin PMTK ve POMTK'dan Kestirilen Madde Parametrelerinin Karşılaştırılması

Çalışma kapsamında yanıt aranan ikinci araştırma sorununun ikinci alt amacı farklı madde sayıları ve örneklem büyüklükleri için parametrik ve parametrik olmayan madde tepki kuramı modellerinden kestirilen madde güçlük ve ayırt edicilik parametrelerinin karşılaştırılmasıdır. Söz konusu araştırma sorusuna yanıt aramak için bir önceki aşamada oluşturulan farklı örneklem büyüklükleri kullanılmış ve bu veri setlerinden beş, 15 ve 25 maddelik yeni veri setleri oluşturulmuştur. Oluşturulan veri setlerinden 500, 1000 ve 3000 kişilik veri setleri için beş, 15 ve 25 maddelik üçer veri seti olmak üzere toplam dokuz ayrı veri seti için PMTK ve POMTK modellerinden uyumlu olduğu belirlenen modellere kullanılarak madde parametre kestirimleri yapılmış ve elde edilen madde parametreleri örneklemelere, madde sayılarına ve kullanılan MTK yaklaşımına göre karşılaştırılarak incelenmiştir. Her iki yaklaşımdan kestirilen madde parametreleri arasındaki ilişkiler öncelikle PMTK için incelenmiş, ardından POMTK'dan kestirilen parametreler ile karşılaştırılarak analiz edilmiştir.

Farklı madde sayılarına göre PMTK'dan kestirilen madde parametrelerinin karşılaştırılması.

1.Beş maddelik veri setleri için madde parametrelerinin hesaplanması

Oluşturulan veri setleri için yapılan model veri uyumu incelemeleri sonucunda, beş maddeden oluşan veri setleri için incelenen hiçbir örneklem büyüklüğünde uyumlu madde olmadığı belirlenmiştir. Modeller arasındaki -2 loglikelihood farklarına göre manidar düzeyde uyumun iyileştiği 2 PLM kullanılarak beş maddelik veri setleri için madde güçlük ve ayırt edicilik parametreleri kestirilmiştir. Elde edilen madde parametreleri Çizelge 42'de verilmiştir.

Çizelge 42.

Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük ve Ayırt Edicilik Parametreleri

Maddeler		M1		M2		M3		M4		M5	
		a	B	a	b	a	b	a	b	a	b
Madde Parametreleri	500	0,98	-1,10	1,94	0,39	1,20	-0,85	2,05	-0,65	1,58	-0,19
Örneklem	1000	1,30	-1,00	1,67	0,43	1,32	-0,90	2,17	-0,58	1,60	-0,22
Büyüklüğü	3000	1,04	-1,13	1,86	0,37	1,19	-0,94	2,48	-0,64	1,40	-0,21

Çizelge 42 incelendiğinde beş maddeden ve farklı örneklem büyüklüklerinden oluşan setlerinden 2 PLM kullanılarak kestirilen madde güçlük ve ayırt edicilik parametreleri görülmektedir. Çizelgede yer alan değerler genel olarak incelendiğinde, madde güçlük parametresi değerlerinin madde ayırt edicilik parametresi değerlerine göre birbirine daha benzer olduğu görülmektedir. Farklı örneklem büyüklüğü koşullarına göre madde güçlük parametreleri daha az değişim göstermiştir. Madde ayırt edicilik parametreleri için ise bu değişimin düzeyi daha yüksektir. Bu bulgu çalışma kapsamında bir önceki aşamada farklı örneklem büyüklüğü koşullarına göre oluşturulan veri setlerinden kestirilen madde ayırt edicilik parametreleri arasında hesaplanan korelasyon katsayıları ile tutarlıdır. Farklı örneklem büyüklüklerinden kestirilen b parametreleri arasındaki ilişki a parametreleri arasındaki ilişkiden daha yüksek bulunmuştur ve aynı durum bu veri setleri için de geçerlidir ve a parametrelerinin örneklem büyüklüğünden daha fazla etkilenmektedir.

Kestirilen parametre değerleri genel olarak incelendikten sonra, parametreler arasındaki ilişkilerin incelenmesine geçilmiştir. İki parametrelilik lojistik model kapsamında

kestirilen madde ayırt edicilik parametreleri arasındaki ilişki, veri setleri için normallik varsayımı karşılanmadığından dolayı, Spearman Brown Korelasyon Katsayısı kullanılarak incelenmiş ve sonuçlar Çizelge 43'te verilmiştir.

Çizelge 43.

Beş Maddelik Veri Setlerinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	Örneklem Büyüklüğü	
	500	1000
1000	1.00**	1.00
3000	1.00**	1.00**

**p < 0.01

Çizelge 43'te yer alan değerler incelendiğinde, beş maddelik veri setlerinden farklı örneklem büyüklüklerine göre kestirilen madde ayırt edicilik parametrelerinin sıralamaları arasında mükemmel uyum olduğu, korelasyon değerlerinin 0.01 düzeyinde manidar olduğu görülmektedir. Bu durum her örneklem büyüklüğünden kestirilen madde ayırt edicilik parametrelerinin sıralamasının aynı olduğu anlamına gelmektedir. Dolayısıyla beş maddelik veri setlerinde farklı örneklem büyüklüklerinden kestirilen madde ayırt edicilik parametrelerinin değerleri farklılaşsa da, sıralamasının benzer olduğu sonucuna ulaşılabilir. Ancak bu sonucu madde sayısının oldukça düşük olmasıyla ilişkilendirerek dikkatli yorumlamak gerekmektedir.

İki parametrelili lojistik model kapsamında beş maddelik veri setlerinde yer alan maddeler için kestirilen bir diğer parametre madde güçlük parametresidir. Bu parametrenin de örneklem büyüklüğünden nasıl etkilendiğini belirlemek için 500, 1000 ve 3000 kişilik veri setlerinden kestirilen madde güçlük parametresi değerleri arasındaki ilişki Spearman Brown Korelasyon Katsayısı kullanılarak incelenmiş ve sonuçlar Çizelge 44'te sunulmuştur.

Çizelge 44.

Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	Örneklem Büyüklüğü	
	500	1000
1000	1.00**	
3000	1.00**	1.00**

**p < 0.01

Çizelge 44'te yer alan değerler incelendiğinde, tüm örneklem büyüklüklerinden kestirilen madde güçlük parametrelerinin madde ayırt edicilik parametrelerinde olduğu gibi mükemmel bir uyum içinde olduğu ve bu parametreye göre yapılan sıralamaların tamamen aynı olduğu görülmektedir. Özetle farklı örneklem büyüklüklerinden oluşan beş maddelik veri setlerinden kestirilen madde güçlük ve ayırt edicilik parametreleri arasında çok yüksek düzeyde ve manidar ilişkiler olduğu bulunmuş ve bir diğer test uzunluğu koşulu olan 15 maddelik veri setlerinden kestirilen parametrelerin incelenmesine geçilmiştir.

2. On beş maddelik veri setleri için madde parametrelerinin hesaplanması

Bu araştırma sorusu içinde incelemeye alınan bir diğer koşul ise madde sayısının 15 olması durumunda farklı örneklem büyüklüklerinden kestirilen madde parametrelerinin nasıl değiştiğidir. Bu amaçla yapılan model veri uyumu incelemelerinde 15 maddelik veri setleri için en yüksek düzeyde model veri uyumunun 500, 1000 ve 3000 kişilik veri setlerinde 2 PLM'de sağlandığı belirlenerek, bu model kullanılarak madde parametreleri kestirilmiş ve elde edilen madde parametre kestirimleri Çizelge 45'te sunulmuştur.

Çizelge 45.

On Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük ve Ayırt Edicilik Parametreleri

Örneklem Büyüklüğü	500		1000		3000	
	a	b	a	b	a	b
M1	0.849	-1,216	1,300	-1,006	1,072	-1,110
M2	1,751	0.401	1,587	0.443	1,645	0.389
M3	1,124	-0.890	1,335	-0.900	1,243	-0.922
M4	2,274	-0.642	2,088	-0.598	2,447	-0.653
M5	1,602	-0.203	1,369	-0.243	1,284	-0.229
M6	1,514	-0.720	1,549	-0.840	1,560	-0.805
M7	1,558	-0.410	1,348	-0.630	1,330	-0.612
M8	1,871	-0.395	2,121	-0.386	2,064	-0.452
M9	2,837	0.759	2,788	0.716	2,604	0.803
M10	1,081	-0.527	1,266	-0.652	1,128	-0.734
M11	1,125	0.600	1,314	0.530	1,271	0.567
M12	0.725	-1,209	0.803	-1,160	0.647	-1,287
M13	1,868	-0.093	1,653	-0.012	1,715	-0.037
M14	1,631	0.407	1,683	0.419	1,566	0.363
M15	1,223	-0.320	1,400	-0.301	1,371	-0.380
Ortalama	1,535	-,297	1,573	-,308	1,529	-,333

Çizelge 45 incelendiğinde 15 maddelik veri setlerinin 500, 1000 ve 3000 kişilik veri setlerinden kestirilen madde ayırt edicilik ve güçlük parametreleri görülmektedir. Farklı örneklem büyüklükleri için kestirilen madde parametreleri genel olarak incelenecek olursa, benzer sonuçların elde edildiği görülmektedir. Madde parametrelerine ilişkin hesaplanan ortalama değerlerin de birbirine yakın olduğu görülmektedir. Madde ayırt edicilik parametrelerinin ortalamasının her örneklem büyüklüğünde 1'den yüksek olduğu, 12 nolu maddenin tüm örneklemde, 1 nolu maddenin de yalnızca 500 kişilik veri setinde 1 değerinin altında olduğu görülmektedir. Madde ayırt edicilik parametreleri 1'in üzerinde olan maddeler çizelgede koyu, 2'nin üzerinde olanlar ise italik yazılarak belirtilmiştir. Madde güçlük parametreleri örneklem büyüklüklerine göre incelendiğinde ise, tüm örneklemde benzer değerler elde edildiği görülmektedir. Bu değerlere ilişkin ortalamaların negatif olduğu belirlenmiş, dolayısıyla maddelerin güçlük düzeyinin düşük olduğu sonucuna ulaşılmıştır. İncelemeye alınan 15 maddenin 11'i her üç örneklem büyüklüğünde de negatif değerler almıştır ve bu maddeler çizelgede altı çizili olarak gösterilmiştir. Her üç örneklem büyüklüğünde de madde 11 en zor madde iken, 1000 ve 3000 kişilik veri setlerinde madde 12 en kolay, 500 kişilik veri setinde ise madde 1 en zor madde olarak belirlenmiştir.

Özetle, 15 maddelik farklı örneklem büyüklüklerinden oluşan veri setleri için PMTK modellerinden kestirilen madde parametrelerinin örneklem bakımından büyük benzerlik gösterdiği sonucuna ulaşılmıştır. Söz konusu benzerliği daha detaylı incelemek için kestirimler arasındaki korelasyonlar incelenmiştir ve ilk olarak madde ayırt edicilik parametreleri arasında hesaplanan korelasyonlar Çizelge 46'da sunulmuştur.

Çizelge 46.

On Beş Maddelik Veri Setlerinden Kestirilen Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	Örneklem Büyüklüğü	
	500	1000
1000	0.950**	
3000	0.968**	0.975**

**p < 0.01

Çizelge 46 incelendiğinde 15 maddelik veri setlerinden farklı örneklem büyüklüklerine göre kestirilen madde ayırt edicilik parametreleri arasındaki korelasyon

katsayılarının yüksek ve 0.01 düzeyinde manidar olduğu görülmektedir. Farklı örneklem büyüklüklerinden kestirilen madde ayırt edicilik değerlerinin sıralamasının birbirine oldukça benzer olduğu, veri setlerine göre sıralamaların tutarlı olduğu sonucuna ulaşılmıştır. Bu bulgu beş maddelik veri setleri için hesaplanan bulgu ile tutarlıdır, söz konusu veri setlerinden kestirilen madde ayırt edicilik parametreleri arasındaki uyumun düzeyi de çok yüksektir. İki parametrelili lojistik model kapsamında 15 maddelik veri setleri için kestirilen madde güçlük parametreleri arasındaki korelasyonlar da Spearman Brown Korelasyon Katsayısı kullanılarak incelenmiş ve sonuçlar Çizelge 47’de sunulmuştur.

Çizelge 47.

On Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	Örneklem Büyüklüğü	
	500	1000
1000	0.982**	
3000	0.964**	0.971**

**p < 0.01

Çizelge 47 incelendiğinde tüm örneklem büyüklüklerinden kestirilen madde güçlük parametrelerinin madde ayırt edicilik parametrelerinde olduğu gibi yüksek bir ilişki içinde olduğu görülmektedir. Farklı örneklemelerden kestirilen tüm madde güçlük parametreleri birbirleri ile yüksek düzeyde ve manidar bir ilişki göstermektedir. Daha önce belirtilen model veri uyumu istatistikleri incelendiğinde, 3000 kişilik veri seti için madde model uyumunun sağlanamadığı görülmektedir, fakat bu örneklemde elde edilen değerlerin madde model uyumunun daha iyi sağlandığı 500 ve 1000 kişilik veri setleri ile çok yakın olduğu hem madde güçlük hem de madde ayırt edicilik parametreleri için gözlenmektedir. Dolayısıyla örneklemdeki kişi sayısının artmasının madde model uyumunu yalnızca ki-kare testi düzeyinde etkilediği, parametre kestiriminde herhangi bir etkisinin olmadığı sonucuna ulaşılabilir. Ayrıca PMTK modellerinin örneklem büyüklüğü ve madde sayısı yeterli olduğunda kestirimler yaptığını literatürde de belirtilmektedir (Baker ve Kim, 2004). Özetle farklı örneklem büyüklüklerinden oluşan 15 maddelik veri setlerinden kestirilen madde parametrelerinin, beş maddelik veri setlerinden kestirilen madde parametreleri gibi de çok yüksek bir uyum içinde olduğu sonucuna ulaşılmıştır.

3.Yirmi beş maddelik veri setleri için madde parametrelerinin hesaplanması

Test uzunluğu faktörü altında incelenen son koşul ise madde sayısının 25 olması durumunda farklı örneklem büyüklüklerinden hesaplanan madde parametrelerinin nasıl değişiklik gösterdiği. Bu amaçla oluşturulan 500, 1000 ve 3000 kişilik veri setlerinin bir, iki ve üç parametrelili lojistik modele uyumları araştırmanın ilk sorusu kapsamında incelenmiş ve 25 maddelik veri setlerinden 500 ve 1000 kişilik veri setleri için 2 PLM'nin, 3000 kişilik veri seti için ise, 1 PLM'nin uyumlu olduğu belirlenmiş ve bu modellere göre madde parametreleri kestirilmiş ve elde edilen sonuçlar Çizelge 48'de sunulmuştur.

Çizelge 48.

Yirmi Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük ve Ayırt Edicilik Parametreleri

Örneklem Büyüklüğü	500		1000		3000
Kullanılan Model	2 PLM		2 PLM		1PLM
Parametreler	a	b	a	b	b
M1	0.919	-1.133	1.314	-1.006	-0.937
M2	1.780	0.395	1.623	0.417	0.402
M3	1.178	-0.852	1.384	-0.890	-0.857
M4	2.388	-0.625	2.274	-0.593	-0.818
M5	1.655	-0.199	1.394	-0.256	-0.232
M6	1.470	-0.721	1.584	-0.844	-0.847
M7	1.461	-0.417	1.289	-0.656	-0.599
M8	1.815	-0.397	2.020	-0.408	-0.539
M9	2.511	0.786	2.605	0.702	1.023
M10	1.027	-0.535	1.301	-0.652	-0.649
M11	1.077	0.626	1.364	0.501	0.511
M12	0.776	-1.132	0.812	-1.157	-0.749
M13	1.876	-0.095	1.578	-0.030	-0.054
M14	1.594	0.413	1.681	0.398	0.365
M15	1.323	-0.302	1.404	-0.315	-0.386
M16	1.626	-0.978	1.598	-0.937	-1.072
M17	1.787	0.769	2.045	0.770	0.809
M18	1.585	0.069	1.661	0.053	0.117
M19	2.015	-0.229	2.052	-0.238	-0.262
M20	1.115	0.125	1.461	0.135	0.031

(Devam ediyor.)

Çizelge 48.(devam)

Yirmi Beş Maddelik Veri Setlerinden Kestirilen Madde Güçlük ve Ayırt Edicilik Parametreleri

Örneklem Büyüklüğü	500		1000		3000
Kullanılan Model	2 PLM		2 PLM		1PLM
Parametreler	a	b	a	b	b
M21	1.465	-0.493	1.465	-0.508	-0.515
M22	1.193	0.151	1.459	0.099	0.060
M23	2.405	1.214	2.846	1.068	1.413
M24	0.982	1.251	0.908	1.174	0.890
M25	0.746	1.031	0.884	1.003	0.811
Ortalama	1,511	-0.051	1,600	-0.087	-0.083 a =1.441

Çizelge 48’de yer alan madde parametreleri genel olarak incelenecek olursa, madde ayırt edicilik ortalamalarının her üç örnekleme de 1’den yüksek ve birbirine yakın değerler olduğu görülmektedir. Aynı durum madde güçlük parametreleri için de geçerlidir, b parametresine ilişkin hesaplanan ortalama değerler tüm örneklerde negatiftir. Dolayısıyla bu değerlere göre, tüm veri setlerinde madde güçlüklerinin düşük, madde ayırt ediciliklerinin ise yüksek olduğu sonucuna ulaşılmaktadır. Çizelgede 1’den büyük madde ayırt edicilik ve güçlük değerleri koyu yazılarak, en yüksek ve en düşük madde ayırt edicilik ve güçlük değerleri ise koyu ve italik yazılarak gösterilmiştir. Elde edilen değerler örneklere göre incelenecek olursa, 500 ve 1000 kişilik veri setlerinden kestirilen madde parametrelerinin çok benzer olduğu görülmektedir. En yüksek ayırt edicilik değerine sahip madde, 500 kişilik veri setinde 2,551 değeriyle 9 nolu madde iken, 1000 kişilik veri setinde 2,846 ile 23 nolu maddedir. 3000 kişilik veri seti için analizler 1 PLM’ye göre yapıldığı için tüm madde ayırt edicilik parametreleri 1,441 olarak hesaplanmıştır. Özetle tüm örnekler için madde ayırt edicilik parametrelerinin oldukça yüksek olduğu sonucuna ulaşılmıştır. Madde güçlük parametreleri örneklere göre incelenecek olursa, tüm örneklerde sıfıra yakın negatif değerler olduğu görülmektedir. 500 ve 1000 kişilik veri setlerinde en zor madde 24 nolu maddedir ve 500 kişilik veri setinde en kolay madde 1 nolu

madde, 1000 kişilik veri setinde ise 12 nolu maddedir. 3000 kişilik veri setinde ise en zor madde 23 nolu maddeyken, en kolay madde 16 nolu maddedir.

Farklı örneklem büyüklüklerinden oluşan veri setlerinden kestirilen madde ayırt edicilik parametreleri arasındaki ilişkiler incelenmiştir. 3000 kişilik veri seti 1 PLM'ye göre ölçeklendiği için tüm maddeler için tek bir ayırt edicilik değeri hesaplanmıştır, bu nedenle 500 ve 1000 kişilik veri setlerinden kestirilen parametreler ile 3000 kişilik veri setinden kestirilen madde ayırt edicilik parametresi arasındaki ilişki incelenmemiştir. Dier iki veri seti olan 500 ve 1000 kişilik veri setlerinden kestirilen madde ayırt edicilik parametreleri arasındaki ilişki ise Spearman Brown Korelasyon Katsayısı hesaplanarak incelenmiş ve elde edilen sonuç Çizelge 49'da sunulmuştur.

Çizelge 49.

Yirmi Beş Maddeden Oluşan 500 ve 3000 Kişilik Veri Setlerinden Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyon Katsayısı

Spearman Brown KK	Örneklem Büüklüğü	1000
		500

**p<0.01

Çizelge 49 incelendiğinde 500 ve 1000 kişilik veri setlerinden oluşan 25 maddelik veri setleri için kestirilen madde ayırt edicilik parametreleri arasındaki Spearman Brown Korelasyon Katsayısı görülmektedir. Bu katsayı incelendiğinde hesaplanan a parametreleri arasında yüksek ve manidar bir ilişki olduğu görülmektedir. Dolayısıyla küçük ve orta büyüklükteki örneklemelerden hesaplanan a parametreleri arasında yüksek düzeyde ve manidar bir ilişki olduğu sonucuna ulaşılmaktadır.

İncelemeye alınan örneklemelerden kestirilen bir diğer madde parametresi ise madde güçlük parametresidir. 500, 1000 ve 3000 kişilik veri setlerinden kestirilen b parametreleri arasındaki ilişki Spearman Brown Korelasyon Katsayısı ile hesaplanmış ve elde edilen sonuçlar Çizelge 50'de sunulmuştur.

Çizelge 50.

Yirmi Beş Maddeden Oluşan 500, 1000 ve 3000 Kişilik Veri Setlerinden Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	Örneklem Büyüklüğü	
	500	1000
1000	0.991**	
3000	0.982**	0.972**

**p<0.01

Çizelge 50 incelediğinde, farklı örneklemelerden hesaplanan madde güçlük parametreleri arasında yüksek düzeyde ve manidar ilişkiler olduğu görülmektedir. Özellikle 500 ve 1000 kişilik veri setlerinden kestirilen b parametreleri arasındaki korelasyon mükemmel düzeydedir. Aynı durum 500 ile 3000 ve 1000 ile 3000 kişilik veri setlerinden kestirilen b parametreleri için de geçerlidir. Özetle, 25 maddelik veri setleri için farklı örneklem büyüklüklerinden kestirilen b parametreleri arasındaki ilişki çok yüksek ve manidardır. Madde güçlük parametre kestirimlerine ilişkin bulguların incelenmesi ile PMTK modellerine göre kestirilen değerlerin örneklem büyüklüğünden ya da madde sayısından manidar düzeyde etkilenmediği, incelenen tüm test uzunluğu ve örneklem büyüklüğü koşullarında elde edilen madde güçlük parametreleri aralarındaki ilişkilerin yüksek olduğu belirlenmiştir.

Farklı test uzunluklarından ve örneklem büyüklüklerinden kestirilen madde parametreleri arasındaki ilişkiler özetlenecek olursa, madde güçlük ve ayırt edicilik değerleri arasında yüksek korelasyonlar olduğu belirlenmiştir. Tüm test uzunluğu koşulları için PMTK kestirimlerinin tutarlı sonuçlar verdiği, hatta model veri uyumunun çok düşük düzeyde sağlandığı beş maddelik veri setlerinden de benzer parametrelerinin elde edildiği bulunmuştur. Bu durumda literatürde belirtildiği şekilde PMTK'nın güçlü bir yaklaşım olduğunu doğrulamaktadır (Embretson ve Reise, 2000).

Farklı Test Uzunluklarından ve Örneklem Büyüklüklerinden Oluşan Veri Setleri İçin PMTK ve POMTK'dan Kestirilen Madde Parametrelerinin Karşılaştırılması

Çalışma kapsamında incelemeye alınan MTK yaklaşımlarından olan POMTK'ya dayalı madde ve yetenek parametre kestirimleri PMTK'dan farklı olarak yaklaşıma ilişkin varsayımların test edilmesi ve model veri uyumunun belirlenmesi aşamasında

gerçekleşmekte, parametre kestirimi PMTK'daki gibi ayrı bir süreç gerektirmemektedir. Araştırma kapsamında örneklem büyüklüğü ve test uzunluğu koşullarına göre oluşturulan veri setlerinin model veri uyumlarının belirlenmesi araştırmanın ilk sorusu olarak ele alınıp yanıtlandığı için, bu aşamada, POMTK için ilk araştırma sorusu içinde hesaplanan madde parametreleri kullanılarak PMTK ile kestirilen parametreler arasındaki korelasyonlar hesaplanmıştır. POMTK'daki madde güçlük parametresi için maddenin doğru yanıtlanma oranı, ayırt edicilik parametresi için ise ölçeklenebilirlik katsayısı kullanılmaktadır. İncelenen tüm veri setlerine ilişkin madde güçlük indeksleri ve ölçeklenebilirlik katsayıları, PMTK ile kestirilen madde güçlük ve ayırt edicilik parametreleri ile kıyaslanmış ve elde edilen sonuçlar bu kısımda sunulmuştur. Öncelikli olarak PMTK ile hesaplanan madde ayırt edicilik parametreleri ile POMTK ile hesaplanan H_i katsayıları arasındaki ilişki hesaplanmış ve farklı test uzunlukları temel alınarak elde edilen bulgular sunulmuştur. Beş, 15 ve 25 maddelik veri setlerinden elde edilen sonuçlar Çizelge 51, 52 ve 53'te verilmiştir. Çizelge 51.

Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar

Sperman Brown KK	PMTK			POMTK	
	500A	1000A	3000A	500H	1000H
1000A	1.00**				
3000A	1.00**	1.00**			
500H	0.900*	0.900**	0.900**		
1000H	0.800**	0.800**	0.800**	0.900**	
3000H	0.900**	0.900**	0.900**	1.00**	0.900**

**p<0.01

Çizelge 52.

On Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar

Sperman Brown KK	PMTK			POMTK	
	500A	1000A	3000A	500H	1000H
1000A	0.950**				
3000A	0.968**	0.975**			
500H	0.929*	0.883**	0.893**		
1000H	0.840**	0.885**	0.861**	0.921**	
3000H	0.893**	0.884**	0.909**	0.966**	0.957**

**p<0.01

Çizelge 53.

Yirmi Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Ayırt Edicilik Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	PMTK		POMTK	
	500A	1000A	500H	1000H
1000A	0.905**	1.00		
500H	0.937**	0.876**	1.00	
1000H	0.863**	0.966**	0.867**	1.00
3000H	0.882**	0.937**	0.915**	0.957**

**p<0.01

Çizelgelerde yer alan veri setleri isimlendirilirken öncelikli olarak veri setinde yer alan birey sayısı, 500, 1000 ve 3000 yazılmıştır. Ardından PMTK kapsamında kestirilen madde ayırt edicilik parametreleri A ile belirtilirken, POMTK modelleri ile hesaplanan madde ayırt edicilik parametreleri H olarak yazılmıştır. Buna göre, 500A, 1000A ve 3000A PMTK modellerine göre hesaplanan madde ayırt edicilik değerlerini göstermekte iken, 500H, 1000H ve 3000H ise POMTK modellerinden kestirilen madde ayırt edicilik parametrelerini belirtmektedir. Farklı test uzunlukları ve örneklem büyüklüklerinden PMTK ve POMTK'ya göre kestirilen madde ayırt edicilik parametreleri arasındaki korelasyonlar genel olarak incelendiğinde şu sonuçlara ulaşılabilir:

- Beş maddelik veri setlerinde parametrik modelden kestirilen tüm a parametreleri için hesaplanan Spearman Brown Korelasyon Katsayıları +1.00 olarak bulunmuş ve bu değer mükemmel ilişkiyi göstermektedir. Parametrik olmayan modelden kestirilen ayırt edicilik parametreleri arasındaki ilişki incelendiğinde ise, 500 ve 3000 kişilik veri setleri arasındaki ilişki +1.00 olarak belirlenmiş, 1000 kişilik veri setinden kestirilen madde ayırt edicilik parametreleri ise hem 500 hem de 3000 kişilik veri setinden elde edilen parametrelerle 0.001 düzeyinde manidar ve 0.90 düzeyinde ilişkiye sahiptir.
- Beş maddelik veri setlerinden 500 ve 3000 kişilik veri setlerinden POMTK ve PMTK modelleri ile kestirilen parametreler arasında ilişki 0.90 olarak belirlenmişken, 500 ile 1000 kişilik veri setlerinden kestirilen parametreler arasındaki ilişki 0.80 olarak bulunmuştur.

- On beş maddelik veri setleri için en yüksek ilişki 500 kişilik veri setlerinden elde edilen parametreler için hesaplanmıştır. En düşük ilişki ise POMTK kapsamında 1000 kişilik veri setinden kestirilen H_i parametreleri ile PTMK ile 500 kişilik veri setinden kestirilen a parametreleri arasında hesaplanmıştır. Genel olarak tüm ilişkiler yüksek ve manidar bulunmuş ve örneklem büyüklükleri eşit olan veri setleri arasında görece daha yüksek ilişkiler olduğu belirlenmiştir.
- Yirmi beş maddelik veri setleri için POMTK ve PMTK modelleri ile yapılan madde ayırt edicilik parametreleri arasındaki ilişkiler incelendiğinde, diğer test uzunluklarına göre korelasyon katsayılarının daha yüksek olduğu görülmektedir. En yüksek ilişki 1000 kişilik veri setlerinden kestirilen parametreler arasında iken en düşük ilişkinin ise 1000 kişilik POMTK, 500 kişilik PMTK yaklaşımı ile kestirilen madde ayırt edicilik parametreleri arasında olduğu belirlenmiştir. Ancak tüm ilişkilerin yüksek ve manidar olduğu görülmektedir.
- Özetle tüm veri setlerinden her iki yaklaşıma ve farklı örneklem büyüklüğü ve test uzunluğu koşullarına göre kestirilen madde ayırt edicilik parametreleri arasında yüksek ve manidar ilişkiler olduğu bulunmuştur. Test uzunluğunun artması kestirilen parametreler arasındaki ilişkinin düzeyini ilişkiler yükseltmiş ve en yüksek ilişkiler aynı örneklem büyüklüklerine sahip veri setlerinden kestirilen madde ayırt edicilik parametreleri için hesaplanmıştır.

Parametrik ve parametrik olmayan MTK yaklaşımı ile ölçeklenen farklı sayıda maddeden ve örneklem büyüklüklerinden oluşan veri setleri için hesaplanan bir diğer parametre olan madde güçlük parametreleri arasındaki ilişkiler de Spearman Brown Korelasyon Katsayısı kullanılarak hesaplanmış ve elde edilen sonuçlar Çizelge 54, 55 ve 56'da verilmiştir.

Çizelge 54.

Beş Maddelik Veri Setleri İçin PMTK ve POMTK Modelleri İle Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	PMTK			POMTK	
	500B	1000B	3000B	500	1000
1000B	1.00**				
3000B	1.00**	1.00**			
500	-0.975**	-0.975**	-0.975**		
1000	-1.00**	-1.00**	-1.00**	0.975**	
3000	-1.00**	-1.00**	-1.00**	0.975**	1.00**

**p<0.01

Çizelge 54'te parametrik ve parametrik olmayan MTK yaklaşımı ile kestirilen madde güçlük parametrelerinin hem kendi içinde hem de karşılıklı ilişkilerini belirten korelasyon katsayıları görülmektedir. Çizelgede yer alan veri setleri isimlendirilirken öncelikle veri setinde yer alan birey sayısına ilişkin değerler, 500, 1000 ve 3000, yazılmıştır. İncelenen madde parametresi olan madde güçlüğü ise tüm veri setlerinde B olarak gösterilmiştir. Bu açıklamalara göre veri setlerine göre kestirilen madde güçlük parametreleri arasındaki ilişkiler incelendiğinde, PMTK ve POMTK ile hesaplanan madde güçlük parametrelerinin kendi içinde mükemmel korelasyonlara sahip olduğu görülmektedir. Aynı durum farklı iki yaklaşımdan hesaplanan parametreler için de geçerlidir fakat ilişkinin yalnızca yönü negatiftir. Bu durumun nedeni ise daha önce de belirtildiği gibi, POMTK kapsamında madde güçlüğü KTK'da olduğu gibi tanımlanmasıdır. Bir diğer ifadeyle POMTK'da madde güçlük parametresi "item popularities" olarak belirtilir ve maddenin doğru cevaplanma oranına dayalı olarak hesaplanır ve bu oran maksimum değer +1.00'e yaklaştıkça madde güçlüğü azalır. PMTK'da ise durum tam tersidir, madde güçlük parametresinin sayısal olarak artışı maddenin güçleştiği anlamına gelmektedir. Dolayısıyla PMTK ve POMTK'dan kestirilen madde güçlük parametreleri arasındaki ilişkinin yönünün negatif olması beklenen bir durumdur (Sijtsma ve Junker, 1998). Özetle, beş maddelik veri setleri için iki farklı yaklaşımla kestirilen madde güçlük parametreleri arasındaki ilişki yüksek ve manidar bulunmuştur.

Çalışma kapsamında madde parametrelerine etkisi incelenen bir diğer test uzunluğu koşulu olan 15 maddelik veri setlerinden PMTK ve POMTK'ya göre kestirilen madde

güçlük parametreleri arasındaki ilişkiler hesaplanmıştır. Elde edilen değerler Çizelge 55'te verilmiştir.

Çizelge 55.

On Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	PMTK			POMTK	
	500B	1000B	3000B	500	1000
1000B	0.982**				
3000B	0.964**	0.971**			
500	-0.928*	-0.950**	-0.928**		
1000	-0.987**	-0.965**	-0.945**	0.976**	
3000	-0.974**	-0.942**	-0.929**	0.983**	0.990**

**p<0.01

Çizelge 55 incelendiğinde 15 maddeden oluşan veri setlerinden kestirilen madde güçlük parametreleri arasındaki ilişkiler görülmekte ve değerler incelendiğinde, gerek PMTK ve POMTK yaklaşımlarının kendi içinde hesaplanan, gerekse her iki yaklaşımdan ayrı ayrı hesaplanan parametreler kıyaslandığında çok yüksek ve manidar ilişkilerin olduğu görülmektedir. Tüm korelasyon katsayıları 0.90'ın üzerinde bulunmuştur ve en yüksek ilişki 500 kişilik veri setinden PMTK ile yapılan kestirimlerle 1000 kişilik veri setinden POMTK'ya dayalı olarak yapılan kestirimler arasında -0.987 olarak bulunmuştur. Diğer örneklem için hesaplanan değerlerin de oldukça yüksek olduğu göz önünde bulundurulursa, on beş maddelik veri setleri için her üç örneklem büyüklüğünde de POMTK'dan kestirilen madde güçlük parametrelerinin PMTK ile kestirilen madde güçlük parametrelerinin yerine kullanabileceği sonucuna ulaşılabilir.

Çizelge 56.

Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	PMTK			POMTK		
	500B	1000B	3000B	500	1000	3000
500B	1.00	.				
1000B	0.991**	1.00				
3000B	0.982**	0.972**	1.00			

(Devam ediyor.)

Çizelge 56.(devam)

Yirmi Beş Maddelik Veri Setleri İçin Hesaplanan Madde Güçlük Parametreleri Arasındaki Korelasyonlar

Spearman Brown KK	PMTK			POMTK		
	500B	1000B	3000B	500	1000	3000
	-0.987**	-0.975**	-0.987**	1.00		
	-0.983**	-0.980**	-0.991**	0.989**	1.00	
	-0.981**	-0.972**	-0.999**	0.989**	0.992**	1.00

**p<0.01

İncelemeye alınan son test uzunluğu koşulu olan 25 maddelik veri setleri için PMTK ve POMTK modelleri ile kestirilen madde güçlük parametreleri Çizelge 56'da verilmiştir. Elde edilen değerler incelendiğinde, her iki yaklaşımdan kestirilen parametreler arasında çok yüksek ve manidar ilişkiler olduğu görülmektedir. Tüm korelasyon katsayılarının neredeyse mükemmel yakın olduğu belirlenmiştir. İncelenen diğer test uzunluklarından elde edilen sonuçlarla paralel olarak, 25 maddelik veri setleri için de PMTK ve POMTK kapsamında kestirilen madde güçlük indekslerinin birbirine çok yakın olduğu sonucuna ulaşılmıştır. Bu bulgu ile çalışma kapsamında cevaplanması planlanan ilk araştırma sorusuna ilişkin tüm bulgular tamamlanmış ve araştırmanın ikinci sorusuna ilişkin elde edilen bulgulara geçilmiştir.

PMTK ve POMTK Kapsamında Kestirilen Yetenek Parametrelerinin Farklı Örneklem Büyüklüğü ve Test Uzunluğu Koşullarına Göre ve Maddelerin Farklı Psikometrik Niteliklerine Göre Değişiminin İncelenmesi

Bu çalışma kapsamında yanıt aranan üçüncü soru şu şekildedir: İncelenen veri setinde yer alan bireylere ilişkin PMTK modellerinden bir-iki ve üç parametrelilik lojistik modelden en iyi uyumu sağlayan model ile POMTK modellerinden uyumlu olduğu belirlenen model kullanılarak belirlenen yetenek kestirimleri,

- a. maddelerin psikometrik niteliklerine,
- b. Test uzunluğuna,
- c. Örneklem büyüklüğüne göre nasıl farklılaşmaktadır?

Yukarıda belirtilen araştırma sorusu kapsamında ilk olarak amaca uygun veri setleri oluşturulmuş, takiben PMTK ve POMTK kestirimlerine geçilmiştir. Elde edilen bulgular öncelikli olarak PMTK modellerine POMTK'ya göre sıralı olarak sunulmuştur.

Maddelerin Farklı Psikometrik Niteliklerine Göre Veri Setlerinin Oluşturulması ve Birey Yeteneklerinin Kestirilmesi

Çalışma kapsamında incelenen üçüncü soruyu yanıtlamak için daha önce oluşturulan örneklemelerden farklı ve amaca uygun örneklemelere ihtiyaç duyulmuştur. Araştırma sorusunun ilk maddesinde bireylerin yetenek kestirimlerinin maddelerin farklı psikometrik niteliklerine göre nasıl değiştiğini belirlemek için bir önceki aşamada kestirilen madde parametreleri kullanılmış ve madde güçlük ve ayırt edicilik parametrelerine ilişkin hesaplanan değerler incelenmiştir. Güçlük ve ayırt edicilik düzeyleri en yüksek ve en düşük olan maddelerin belirlenmesi için tüm maddelerin yer aldığı 30 maddelik temel veri seti ile her iki yaklaşımda da model veri uyumunun en yüksek düzeyde sağlandığı 1000 kişilik veri seti kullanılmıştır. Söz konusu veri seti kullanılarak, tüm maddeler, madde güçlük parametrelerine göre kolaydan zora ve ayırt edicilik parametrelerine göre ise düşükten yükseğe doğru sıralanmış ve bu değerler Çizelge 57'de verilmiştir.

Çizelge 57.

PMTK ve POMTK Kapsamında 1000 Kişilik Veri Setlerinden Hesaplanan Madde Güçlük ve Ayırt Edicilik Parametreleri Sıralaması

Madde Güçlük Parametresi				Madde Ayırt Edicilik Parametresi			
PMTK		POMTK		PMTK		POMTK	
-1.181	M12	M1	,73	0.787	M12	M12	0.235
-0.992	M1	M16	,73	0.892	M25	M25	0.300
-0.916	M16	M3	,71	0.917	M24	M24	0.315
-0.871	M3	M6	,71	1.236	M30	M7	0.324
-0.858	M6	M12	,69	1.267	M7	M30	0.328
-0.760	M28	M28	,69	1.308	M10	M10	0.336
-0.661	M7	M4	,67	1.341	M1	M15	0.341
-0.650	M10	M7	,65	1.352	M11	M5	0.348
-0.601	M4	M10	,65	1.372	M15	M1	0.353
-0.503	M21	M21	,62	1.382	M26	M21	0.357
-0.419	M8	M8	,61	1.407	M5	M3	0.359
-0.318	M15	M15	,57	1.440	M3	M28	0.360

(Devam ediyor.)

Çizelge 57.(devam)

PMTK ve POMTK Kapsamında 1000 Kişilik Veri Setlerinden Hesaplanan Madde Güçlük ve Ayırt Edicilik Parametreleri Sıralaması

Madde Güçlük Parametresi				Madde Ayırt Edicilik Parametresi			
PMTK		POMTK		PMTK		POMTK	
-0.271	<u>M27</u>	<u>M27</u>	,57	1.456	M20	M20	0.365
-0.256	<u>M5</u>	<u>M5</u>	,56	1.473	M28	M26	0.365
-0.242	<u>M19</u>	<u>M19</u>	,56	1.486	M22	M22	0.369
-0.122	<u>M30</u>	<u>M30</u>	,52	1.505	M21	M6	0.372
-0.031	<u>M13</u>	<u>M13</u>	,50	1.528	M6	M13	0.379
0.054	<u>M18</u>	<u>M18</u>	,48	1.553	<u>M13</u>	M11	0.380
0.096	<u>M22</u>	<u>M22</u>	,47	1.590	M2	M18	0.386
0.135	<u>M20</u>	<u>M20</u>	,46	1.600	M18	M27	0.391
0.232	<u>M29</u>	<u>M26</u>	,43	1.674	<u>M16</u>	<u>M16</u>	0.395
0.24	<u>M26</u>	<u>M29</u>	,43	1.687	M14	M8	0.401
0.397	M14	M2	,38	1.708	M29	M2	0.406
0.423	M2	M14	,38	1.736	M27	M19	0.416
0.506	<u>M11</u>	<u>M11</u>	,37	1.899	M8	M29	0.416
0.725	M9	M25	,32	2.008	M17	M14	0.424
0.780	M17	M24	,29	2.084	M19	M4	0.429
1.003	M25	M9	,27	2.204	M4	M17	0.521
1.089	M23	M17	,27	2.405	<u>M9</u>	<u>M9</u>	0.557
1.172	M24	M23	,18	2.780	<u>M23</u>	<u>M23</u>	0.669

Çizelge 57’de 1000 kişiden oluşan 30 maddelik veri setlerinin PMTK ve POMTK ile kestirilen madde güçlük ve ayırt edicilik parametreleri görülmektedir. Hesaplanan madde parametreleri her iki yaklaşım gözetilerek incelenmiş ve sıralamaları aynı olan maddeler altı çizili olarak gösterilmiştir. Çizelgede yer alan değerler incelendiğinde her iki yaklaşımda sıralamaların tamamen aynı olmadığı fakat benzer olduğu görülmektedir. PMTK’ya göre negatif güçlük indeksine sahip olan ilk 17 maddenin, POMTK’ya göre hesaplanan madde güçlük indeksleri kolaydan zora doğru sıralandığında ilk 17’de yer aldığı belirlenmiştir. Aynı durum madde ayırt edicilik parametreleri için de geçerlidir ve POMTK’da H_i katsayısı olarak adlandırılan madde ayırt edicilik değerleri incelendiğinde, 0.4’ün üzerinde H değerine sahip olan dokuz maddenin sekizinin PMTK kapsamında da en yüksek a değerlerine sahip olduğu çizelgede görülmektedir. Her iki yaklaşımdan kestirilen parametreler gözetilerek, maddeler, güçlük düzeyleri bakımından kolay ve zor, ayırt

edicilikleri bakımından ise düşük ve yüksek olarak dört ayrı madde seti şeklinde gruplandırılmıştır. Madde setleri oluşturulurken, maddelerin psikometrik niteliklerinin yetenek kestirimleri üzerinde olan etkisini öne çıkarmak için madde sayılarının eşit olmasına dikkat edilmiştir. Bu amaçla güçlük düzeyi en düşük ve yüksek olan on beş madde, ayırt edicilik düzeyi en düşük ve yüksek olan on beşer madde belirlenerek veri setleri oluşturulmuştur.

Madde güçlük indekslerine göre yapılacak incelemeler için, maddeler belirlenirken yalnızca güçlük düzeyleri dikkate alınmış, ayırt edicilik parametrelerine ilişkin herhangi bir koşul aranmamıştır. Bu durumun nedeni ise tüm maddelerin ayırt edicilik değerlerinin kabul edilebilir düzeyde ve daha yüksek değerler olmasıdır. Madde güçlük parametresi düşük olan maddelerden oluşturulan veri seti için düşük güçlük düzeyinden yüksek güçlük düzeyine göre sıralanmış maddelerin ilk 15'i alınmıştır. Yüksek madde güçlük düzeyine sahip veri oluşturulurken ise sıralamada son 15'te yer alan maddeler alınmıştır. Söz konusu maddelerin sıralamalarında bazı değişiklikler olsa da, bu maddeler her iki veri setinde de aynı grupta yer alan maddelerdir. Ayırt edicilik parametreleri düşük olan maddelerden oluşturulan veri setleri için de aynı yöntem izlenmiştir, fakat diğer veri setlerinden farklı olarak ayırt edicilikleri sırasıyla PMTK ve POMTK yaklaşımlarına göre yüksek olarak kestirilen 11 ve 21 nolu maddeler alt ayırt edicilik grubuna dâhil edilmiştir. Bu durumun nedeni ise bu maddelerin her iki yaklaşımda birbirinden farklı değerler almasıdır. PMTK'ya göre 1.503 madde ayırt edicilik indeksine sahip 21 nolu madde PMTK'da 0.357 madde ayırt edicilik değerine sahiptir ve alt grupta yer almaktadır. Aynı durum 11 nolu madde için de geçerlidir, bu maddenin ayırt edicilik değeri POMTK'da yüksek iken, PMTK'da görece olarak düşüktür. Kestirilen madde ayırt edicilik indekslerinin yaklaşımlar arasındaki tutarsızlığından dolayı her iki madde de düşük ayırt edicilik indeksine sahip maddelerin yer aldığı veri setine dâhil edilmiş ve PMTK ve POMTK'ya göre yetenek kestirimlerine geçilmiştir.

En Düşük ve En Yüksek Madde Ayırt Edicilik Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden Kestirilen Yetenek Parametreleri

Madde ayırt edicilik parametreleri temel alınarak düşük ve yüksek ayırt ediciliğe sahip maddelerden oluşan iki veri setine ilişkin PMTK kapsamında yetenek kestirimlerine

geçilmeden önce, oluşturulan veri setlerinin PMTK modellerine uyumu incelenmiştir. Söz konusu veri setleri daha önce iki parametreli lojistik modele uyumlu olduğu belirlenen 1000 kişilik 30 maddeden oluşan veri seti temel alınarak oluşturulmuş olsa da, madde sayısının azalmasının model veri uyumunu değiştirebileceği göz önünde bulundurularak, veri setleri için model veri uyumu incelemesi yapılmıştır.

Yapılan model veri uyumu incelemelerinde düşük ayırt edicilik değerine sahip maddelerden oluşan veri seti için -2 loglikelihood değerleri arasındaki fark ve uyumlu madde sayısı incelendiğinde, sekiz uyumlu madde ile iki parametrelili lojistik modelin diğer modellere göre veri setine daha iyi uyum sağladığı belirlenmiştir. Yüksek ayırt edicilik değerine sahip maddelerden oluşan veri setine ilişkin yapılan model veri uyumu incelemelerinde ise bu veri seti için model veri uyumun diğer veri setine kıyasla daha yüksek düzeyde sağlandığı belirlenmiştir ve on iki uyumlu madde sayısı ile bir parametrelili lojistik model kullanılarak yetenek kestirimleri yapılmıştır. Birey yetenek kestirimleri R programındaki “Latent Trait Model (ltm)” paketi ile Bayesian yaklaşım kullanılarak, EAP yöntemine göre yapılmıştır.

POMTK kapsamında yapılan incelemeler için ise ilk olarak veri setlerinin POMTK kapsamında ölçeklenmeye uygunluğu incelenmiş ve maddeler, madde çiftleri ve ölçeğin tümü için H katsayıları hesaplanmıştır. Düşük madde ayırt edicilik parametresine sahip maddelerden oluşan veri seti için yapılan POMTK incelemesinde, madde çiftleri için hesaplanan H_{ij} katsayılarının tümünün pozitif olduğu belirlenmiştir. Maddeler için hesaplanan H katsayıları incelendiğinde altı maddeye ilişkin kestirilen H_i değerlerinin (M7, M10, M12, M15, M25 ve M30) sınır değer olan 0.30’un altında olduğu belirlenmiştir. H_i katsayısının madde ayırt edicilik parametresi olarak yorumlandığı göz önünde bulundurulduğunda, maddelerin H_i değerlerinin düşük olması beklenen bir durumdur. Ölçek için hesaplanan H katsayısı ise 0.298 olarak bulunmuştur ve veri setini ölçeklemeye sınır düzeyde uyum sağladığı belirlenmiştir. Ayrıca madde tepki fonksiyonlarının monoton azalmayan nitelikte olduğu belirlenmiş ve OMSİ ile incelenen tek boyutluluk varsayımının ise bir madde (M7) hariç karşılandığı belirlenmiştir. Söz konusu madde test dışında bırakıldığında tüm H katsayılarının sınır değerinin üzerinde değerler aldığı bulunmuş ve bu madde POMTK modelleri ile yapılan kestirimlerde veri seti dışında tutulmuştur. Ayrıca veri seti için değişmez madde sıralaması varsayımının karşılanmadığı ve bundan dolayı veri

seti için Monoton Homojenlik Modelinin uygun olduğu belirlenmiş ve bireylerin yetenekleri toplam puana göre sıralama ölçeği düzeyinde hesaplanmıştır.

İncelemeye alınan ikinci veri seti olan yüksek ayırt edicilik parametrelerine sahip maddelerden oluşan veri seti için POMTK analizleri yapılmış ve analizlere H katsayılarının hesaplanması ile başlanmıştır. Veri setinin yapısı gereği madde ayırt edicilikleri yüksek olan maddelerden oluştuğu için tüm H katsayılarının yüksek olduğu bulunmuştur. Madde çiftleri için hesaplanan H_{ij} katsayılarının tümünün pozitif olduğu, maddeler için hesaplanan H_i katsayılarının 0.406 ile 0.599 arasında değiştiği ve ölçeğin tümü için hesaplanan H katsayısının ise 0.475 olarak hesaplandığı bulunmuştur. Monotonluk ve tek boyutluluk varsayımlarının da veri setleri için karşılandığı belirlenmiştir. Değişmez madde sıralaması varsayımı için yapılan üç ayrı analiz sonucunda bu varsayımın veri seti için ihmal edilebilir ihlaller ile karşılandığı sonucuna ulaşılmıştır. 0.309 olarak hesaplanan H^T katsayısına göre ise, bu varsayım veri setinde düşük düzeyde karşılanmaktadır. Yapılan incelemelere göre, veri setinin Çift Monotonluk Modeline uyum sağladığı sonucuna varılmıştır. Dolayısıyla bireylerin yetenekleri ÇMM'ye uygun olarak eşit aralıklı ölçek düzeyinde hesaplanmıştır.

PMTK ve POMTK'ya uyumlu olan modeller belirlendikten sonra, her iki veri seti için de iki yaklaşıma dayalı olarak yetenek kestirimleri yapılmıştır. Her veri seti için hesaplanan birey yeteneklerine ilişkin betimsel istatistikler Çizelge 58'de verilmektedir.

Çizelge 58.

En Düşük ve En Yüksek Madde Ayırt Edicilik Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden Kestirilen Birey Yeteneklerine İlişkin Betimsel İstatistikler

Veri Setleri	N	En Düşük Puan	En Yüksek Puan	Ortanca	Ortalama	S. Sapma	Çarpıklık	Basıklık
PMTKendusuk	1000	-2.034	1.741	-0.086	.710	.717	0.217	1.102
PMTKenyukse	1000	-1.670	1.789	-0.032	1.126	.730	0.263	1.283
POMTKendusuk	1000	0	14	7	7.69	3.492	0.030	-0.940
POMTKenyukse	1000	0	15	7	7.29	4.356	0.168	-1.069

Çizelge 58'de yer alan değerler incelendiğinde, PMTKenyukse ve PMTKendusuk olarak belirtilen veri setleri, PMTK ile en düşük ve en yüksek madde ayırt edicilik değerlerine sahip maddelerden oluşan veri setlerinden kestirilen birey yeteneklerini ifade etmekteyken, POMTKendusuk ve POMTKenyukse olarak adlandırılan veri setleri ise POMTK kapsamında madde ayırt edicilik düzeyi en düşük ve en yüksek olan maddelerden

oluşan veri setlerinden kestirilen birey yeteneklerini belirtmektedir. Farklı veri setlerinden uygun modellere göre kestirilen yeteneklere ilişkin betimsel istatistikler incelendiğinde, POMTK kapsamında kestirilen yeteneklerde hiçbir maddeye doğru yanıt veremeyen ve tüm maddeleri doğru yanıtlayan bireylerin olduğu görülmektedir. Yine POMTK kapsamında en düşük ve yüksek ayırt edicilik değerine sahip maddelerden oluşan veri setlerinden kestirilen yeteneklere ilişkin hesaplanan ortalamaların birbirine yakın olduğu görülmektedir. Dağılımların çarpıklık ve basıklık değerleri incelendiğinde ise, her iki dağılımın da hafif sola çarpık ve basık olduğu fakat normallikten fazla sapma göstermedikleri görülmektedir. PMTK modelleri ile kestirilen yetenekler incelendiğinde ise, madde ayırt edicilik değerlerinin düşük olduğu veri setinden kestirilen yeteneklerde diğer veri setine göre daha düşük yetenek kestirimlerin yer aldığı görülmekte, fakat her iki veri setinden kestirilen en yüksek yetenekler benzerlik göstermektedir. Aritmetik ortalamalar incelendiğinde, yüksek madde ayırt edicilik parametresine sahip veri seri setinden kestirilen yeteneklerin ortalamalarının, düşük ayırt edicilik veri setine göre daha yüksek olduğu bulunmuştur. Ayrıca her iki veri setinden kestirilen yeteneklerin pozitif çarpıklık değerine sahip olduğu ve sivri bir dağılım gösterdiği görülmektedir. Basıklık değerlerine göre normal dağılım varsayımının ihlal edildiği görülmektedir.

Gerek PMTK ile yapılan analizlerde basıklık değerinin -1, 1 aralığı dışında kalması ile normal dağılım sağlanamaması, gerekse POMTK ile yapılan analizler ile düşük ayırt edicilik değerlerine sahip maddelerden oluşan veri setinin MHM'ye uyum gösterdiğinin belirlenmesi ve sıralama düzeyinde bilgi vermesi nedeniyle kestirilen yetenekler arasındaki korelasyon Kendall Tau Korelasyon Katsayısı kullanılarak hesaplanmış ve elde edilen sonuçlar Çizelge 59'da verilmiştir.

Çizelge 59.

En Yüksek ve En Düşük Madde Ayırt Edicilik Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenekler Arasındaki İlişkiler

Kendall Tau KK	PMTK		POMTK
	Endusuk	enyukse	endusuk
PMTKenyukse	0.849**		
POMTKendusuk	0.991**	0.840**	
POMTKenyukse	0.849**	1.000**	0.840**

**p<0.01

Çizelge 59’da verilen değerlerden PMTK ve POMTK’ya göre kestirilen yetenek parametreleri arasındaki ilişki incelenecek olursa, düşük madde ayırt edicilik parametrelerine sahip maddelerden oluşan veri setinden her iki yaklaşıma göre kestirilen yetenekler arasında mükemmel ve manidar bir ilişki olduğu görülmektedir. Aynı durum en yüksek madde ayırt edicilik parametrelerine sahip veri setlerinden kestirilen yetenekler için de geçerlidir. İki farklı veri setinden kestirilen yetenek parametreleri incelendiğinde ise, PMTK kapsamında yapılan kestirimlerde en düşük ve en yüksek ayırt edicilik değerlerine sahip veri setlerinden kestirilen parametreler arasında yüksek ve manidar 0.849, bir ilişki olduğu görülmektedir. POMTK’ya göre yapılan kestirimler incelendiğinde ise, en düşük ve yüksek ayırt edicilik değerlerine sahip maddelerden oluşan veri setlerinden kestirilen yetenek parametreleri arasındaki korelasyon 0.840 olarak belirlenmiştir. Ayrıca PMTK ile POMTK’dan kestirilen en düşük ve en yüksek ayırt ediciliğe sahip maddelerden oluşan veri setleri arasındaki ilişki de 0.840 ve 0.849 olarak bulunmuştur. Özetle, her iki yaklaşımdan aynı veri seti kullanılarak kestirilen yetenekler arasındaki ilişkinin, aynı yaklaşımla farklı veri seti kullanılarak yapılan yetenek kestirimleri arasındaki ilişkiden daha yüksek olduğu bulunmuştur. Bir diğer ifadeyle PMTK ile en düşük ayırt edicilik değerine sahip veri setinden kestirilen yetenek parametreleri ile POMTK ile aynı veri setinden kestirilen yetenekler arasındaki ilişki, PMTK ile en yüksek ayırt ediciliğe sahip maddelerden oluşan veri setinden kestirilen yetenek parametreleri arasındaki ilişkiden daha yüksektir.

Bu sonuçlara göre madde ayırt edicilik parametrelerindeki farklılıkların hem PMTK’ya hem de POMTK’ya göre yapılan yetenek kestirimlerini etkilediği ve yetenekler arasındaki korelasyonu düşürdüğü, fakat aynı veri setinden kestirilen yeteneklerin, madde ayırt edicilik değerleri gözetilmeksizin her iki yaklaşımda da benzer olduğu bulunmuştur. Madde ayırt edicilik parametresine göre yapılan yetenek kestirimlerinin ardından, madde güçlük parametresinin düşük ve yüksek olduğu maddelerden oluşturulan veri setlerinden yetenek parametrelerinin kestirilmesine geçilmiştir.

En Düşük ve En Yüksek Madde Güçlük Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden Kestirilen Yetenek Parametreleri

Madde güçlük parametreleri temel alınarak oluşturulan veri setleri kullanılarak yetenek kestirimlerine geçilmeden önce, veri setlerinin PMTK ve POMTK’ya olan uyumu

incelenmiştir. Madde güçlük parametreleri yüksek olan maddelerden oluşturulan veri setinin PMTK modellerinden en iyi uyumu yedi madde ile üç parametrelili lojistik model için sağladığı belirlenmiştir. Madde güçlük düzeyi düşük olan maddelerden, bir diğer ifadeyle kolay maddelerden oluşan veri seti ise en iyi uyumu sekiz uyumlu madde ile bir parametrelili lojistik model için sağlamıştır. Dolayısıyla en kolay ve zor maddelerden oluşan veri setleri için PMTK kapsamında bir ve üç parametrelili lojistik model kullanılarak yetenek kestirimleri yapılmıştır. Birey yetenek kestirimleri R programındaki “Latent Trait Model (ltm)” paketi ile Bayesian yaklaşım kullanılarak, EAP yöntemine göre yapılmıştır.

POMTK kapsamında yapılan incelemeler için ise ilk olarak veri setlerinin POMTK kapsamında ölçeklemeye olan uygunluğu incelenmiş ve maddeler, madde çiftleri ve ölçeğin tümü için H katsayıları hesaplanmıştır. Düşük madde güçlük parametresine sahip maddelerden oluşan veri seti için yapılan POMTK incelemesinde, madde çiftleri için hesaplanan H_{ij} katsayılarının tümünün pozitif olduğu belirlenmiştir. Maddeler için hesaplanan H katsayıları incelendiğinde beş maddeye ilişkin kestirilen H_i değerlerinin (M1, M3, M7, M10 ve M12) sınır değer olan 0.30’un altında olduğu belirlenmiştir. Ölçek için hesaplanan H katsayısı ise 0.307 olarak bulunmuştur ve veri setini ölçeklemeye uyumlu olduğu sonucuna ulaşılmıştır. Ayrıca madde tepki fonksiyonlarının monoton azalmayan nitelikte olduğu belirlenmiş ve OMSİ ile incelenen tek boyutluluk varsayımının ise bir madde (M12) hariç karşılandığı belirlenmiştir. Söz konusu madde test dışında bırakıldığında maddelere ilişkin H katsayılarında büyük değişimler olmadığı belirlenmiş ve 0.3 katsayısının katı bir değer olarak nitelendirilmesinden dolayı maddenin veri setinde bırakılmasına karar verilmiştir. Ayrıca veri seti için değişmez madde sıralaması varsayımı madde kalan puanı grupları ve P matrisleri oluşturularak incelenmiş ve veri setinin bu varsayımı karşılamadığı belirlenmiştir. Özetle bu veri seti için Monoton Homojenlik Modelinin uygun olduğu belirlenmiş ve bireylerin yetenekleri toplam puana göre sıralama ölçeği düzeyinde hesaplanmıştır.

İncelemeye alınan ikinci veri seti olan yüksek güçlük parametrelerine sahip maddelerden oluşan veri seti için POMTK analizleri yapılmış ve analizlere H katsayılarının hesaplanması ile başlanmıştır. İncelenen diğer veri seti olan madde güçlük düzeylerinin düşük olduğu veri setine göre, hesaplanan H_i katsayılarının daha yüksek olduğu bulunmuştur. Yalnızca iki maddenin (M24 ve M25) 0.30’tan düşük H_i değerine sahip

olduğu belirlenmiştir. Ayrıca madde çiftleri için hesaplanan H_{ij} katsayılarının tümünün pozitif olduğu, maddeler için hesaplanan H_i katsayılarının, 0.283 ile 0.611 arasında değiştiği ve ölçeğin tümü için hesaplanan H katsayısının ise 0.373 olarak hesaplandığı bulunmuştur. Monotonluk ve tek boyutluluk varsayımlarının da veri setleri için karşılandığı belirlenmiştir. Değişmez madde sıralaması varsayımı için yapılan madde kalan puanı grupları ve P matrisi analizleri sonucunda, üç maddenin bu varsayımı ciddi düzeyde ihlal ettiği ve bu varsayımın veri seti için karşılanmadığı sonucuna ulaşılmıştır. Ayrıca, madde değişmezliğinin düzeyini belirten H^T katsayısı, 0.134 olarak bulunmuştur, dolayısıyla bu veri seti için değişmez madde sıralaması bir diğer ifadeyle kesişmeyen madde tepki fonksiyonları varsayımının karşılanmadığı belirlenmiş ve birey yetenek kestirimleri Monoton Homojenlik Modeline uyumlu olarak sıralama ölçeği düzeyinde hesaplanmıştır.

PMTK ve POMTK'ya uyumlu olan modeller belirlendikten sonra, her iki veri seti için de iki yaklaşıma dayalı olarak yetenek kestirimleri yapılmıştır. Her veri seti için hesaplanan birey yeteneklerine ilişkin betimsel istatistikler Çizelge 60'da verilmiştir.

Çizelge 60.

En Düşük ve En Yüksek Madde Güçlük Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden Kestirilen Birey Yeteneklerine İlişkin Betimsel İstatistikler

Veri Setleri	N	En Düşük Puan	En Yüksek Puan	Ortanca	Aritmetik Ortalama	S.Sapma	Çarpıklık	Basıklık
PMTKenzor	1000	-1.095	1.884	-,108	,088	,770	0.416	-0.438
PMTKenkolay	1000	-2.170	1.316	,100	-,041	,817	-0.018	-0.936
POMTKenkolay	1000	0	15	10	9,73	3,962	-0.372	-1,004
POMTKenzor	1000	0	15	5	5,74	4,177	0.629	-0.631

Çizelge 60'da veri setleri olarak belirtilen, PMTKenkolay ve PMTKenzor, PMTK ile en kolay ve en zor maddelerden oluşan veri setlerinden kestirilen yetenekleri POMTKenkolay ve POMTKenzor ise POMTK ile en kolay ve en zor maddelerden oluşan veri setlerinden kestirilen yetenekleri belirtmektedir. Farklı veri setlerine göre kestirilen yeteneklerin betimsel istatistikleri incelendiğinde, POMTK ile en kolay maddelerden oluşan veri setinden kestirilen yetenek parametrelerinin ortalamasının en zor maddelerden oluşan veri setinden kestirilen yeteneklerin ortalamasına göre daha yüksek olduğu

görülmektedir ve bu bulgu beklenen bir bulgudur. POMTK kapsamında yapılan yetenek kestirimlerinin maddelerden bağımsız olma gibi bir avantajı olmadığı için, bireylerin güçlük düzeyi düzeyi düşük olan maddelere doğru yanıt verme olasılığı yüksek olan maddelere göre daha yüksektir.

PMTK ile yapılan kestirimlerde ise bu durumun tersi gözlenmekte, en zor maddelerden oluşan veri setinden kestirilen yetenek parametrelerinin ortalaması, en kolay maddelerden oluşan veri setinden kestirilen yetenek parametrelerinin ortalamasından daha yüksektir. Benzer şekilde PMTK kapsamında en kolay maddelerden kestirilen yeteneklerin en düşük değeri (-2.170) aynı yaklaşımla en zor maddelerden kestirilen yeteneklerin minimum değerinden çok daha düşüktür (-1.095). Bu durumun nedeni, yetenek kestirimleri yapılırken farklı modellerin kullanılması ve PMTK kapsamında model veri uyumunun sağlandığı durumlarda, birey yeteneklerinin maddelerden bağımsız olarak kestirilmesidir (Embretson ve Reise, 2000). Ayrıca hesaplanan standart sapmalarının tüm veri setleri ve yaklaşımlar için yüksek olduğu görülmektedir. Çarpıklık değerleri incelendiğinde ise, tüm değerlerin -1 ve 1 arasında olduğu görülmektedir, dolayısıyla normal dağılımdan büyük sapmaların olmadığı sonucuna varılabilir. Her iki yaklaşımla en kolay veri setlerinden kestirilen yeteneklerin hafif sağa çarpık, en zor maddelerden kestirilen yeteneklerin ise sola çarpık dağılımlar gösterdiği belirlenmiştir. Basıklık değerlerine göre ise tüm veri setlerinin negatif basıklık değerlerine sahip olduğu ve dağılımların sivri olduğu belirlenmiştir.

Yetenek kestirimlerine ilişkin hesaplanan betimsel istatistiklerin incelenmesinden sonra, kestirilen parametreler arasındaki ilişkinin incelenmesine geçilmiştir. Yapılan model veri uyumu incelemelerinde en yüksek ve en düşük madde güçlük parametresine sahip olan maddelerden oluşan veri setlerinin POMTK modellerinden Monoton Homojenlik Modeline uyum sağladığı belirlenmiştir. Söz konusu model yetenek kestirimlerini sıralama düzeyinde verdiği için POMTK ile yapılan kestirimler ile PMTK kullanılarak yapılan yetenek kestirimleri arasındaki ilişki Kendall Tau Korelasyon Katsayısı kullanılarak hesaplanmış ve elde edilen sonuçlar Çizelge 61'de verilmiştir.

Çizelge 61.

En Yüksek ve En Düşük Madde Güçlük Parametrelerine Sahip Maddelerden Oluşan Veri Setlerinden POMTK ve PMTK'ya Göre Kestirilen Yetenekler Arasındaki İlişkiler

Kendall Tau KK	POMTK		PMTK
	En kolay	En zor	En zor
POMTKenzor	0.837**		
PMTKenzor	0.829**	0.975**	
PMTKenkolay	0.100**	0.837**	0.829**

**p<0.01

Çizelge 61'de yer alan değerlerden öncelikli olarak her yaklaşımın kendi içindeki ilişkilere ilişkin hesaplanan katsayılar incelenecek olursa, POMTK kullanılarak en kolay ve en zor maddelerden kestirilen yetenekler arasında yüksek ve manidar bir ilişki olduğu görülmektedir. Aynı durum PMTK için de geçerlidir ve her iki yaklaşımda da en zor ve kolay maddelerden kestirilen yetenekler arasındaki ilişkinin düzeyleri de birbirine yakındır. Bu değerler POMTK için 0.837, PMTK için 0.829 olarak bulunmuştur. Dolayısıyla PMTK ve POMTK için, zor ve kolay maddelerden kestirilen yetenekler arasında manidar ve yüksek bir ilişki olduğu sonucuna ulaşılmaktadır. Yetenekler arasındaki ilişkiler PMTK ve POMTK yaklaşımlarına göre incelenecek olursa, POMTK kapsamında en kolay maddelerden oluşan veri setinden kestirilen yetenekler ile aynı veri setinden PMTK ile kestirilen yetenekler arasında mükemmel ve manidar bir ilişki olduğu görülmektedir. Benzer durum en zor maddelerden oluşan veri setlerinden kestirilen yetenek parametreleri için de geçerlidir. POMTK ve PMTK'dan en zor maddeler kullanılarak yapılan yetenek kestirimleri arasında mükemmele yakın 0.975 ve manidar bir ilişki olduğu belirlenmiştir. Her iki yaklaşıma göre farklı veri setlerinden kestirilen yetenek parametrelerinin ilişkileri incelendiğinde ise, yine yüksek ve manidar ilişkiler (0.829 ve 0.837) olduğu görülmektedir. POMTK ve PMTK'ya göre farklı güçlük düzeyindeki veri setlerinden kestirilen yetenekler arasındaki ilişkiler özetlendiğinde şu sonuçlara ulaşılabilir:

- Veri setinde yer alan maddelerin güçlük düzeyi fark etmeksizin, her iki yaklaşımdan kestirilen yetenek parametreleri arasında yüksek ve manidar ilişki bulunmaktadır.
- Aynı veri setinden her iki yaklaşımla kestirilen yetenekler arasındaki ilişkilerin mükemmele yakın ve manidar olduğu belirlenmiştir.

POMTK kapsamında yapılan analizlerde Monoton Homojenlik Modeli kullanıldığı için, elde edilen yetenek kestirimleri sıralama düzeyinde bilgi vermektedir. Bu modelden

hesaplanan yetenek kestirimlerinin PMTK modellerinden kestirilen yetenekler ile yüksek düzeyde uyum sağladığı göz önünde bulundurulduğunda, yapılacak ölçme uygulamalarında bireylerin yetenek düzeyleri bakımından sıralanması yeterli ise, POMTK kapsamında ölçekleme yapılabileceği sonucuna ulaşılmaktadır. Bu bulgu alan yazınla da tutarlılık göstermektedir (Meijer, Tenderio ve Wanders, 2015).

Farklı Test Uzunluklarına ve Örneklem Büyüklüklerine Göre Veri Setlerinin Oluşturulması ve Birey Yeteneklerinin Kestirilmesi

Çalışma kapsamında yanıt aranan üçüncü sorunun ikinci maddesi farklı örneklem büyüklükleri ve örneklem büyüklüklerine göre oluşturulan veri setlerinden kestirilen yetenek parametreleri arasındaki ilişkilerin incelenmesidir. Bu amaçla çalışma kapsamında oluşturulan temel veri setinden farklı örneklem büyüklüğü faktörü içinde yer alan üç koşul için 500, 1000 ve 3000 kişilik veri setleri oluşturulmuştur. Farklı madde sayısı faktörü içinde de üç koşul incelemeye alınmış ve bu koşullar beş, 15 ve 25 maddelik test uzunluğu koşullarıdır ve bu maddelerin temel veri setinden seçkisiz olarak oluşturulmuştur. Her iki faktörden incelemeye alınan koşullar çaprazlanmış ve 500, 1000 ve 3000 kişiden oluşan dokuz ayrı veri seti oluşturulmuştur. Oluşturulan veri setlerinden PMTK ve POMTK modellerine göre yetenek kestirimleri yapılmadan önce her iki yaklaşıma göre model veri uyumu incelemesi yapılmış ve en iyi uyumun sağlandığı modeller belirlendikten sonra, yetenek parametrelerine ilişkin kestirimler yapılmış ve elde edilen bulgular PMTK ve POMTK yaklaşımlarına göre sıralı olarak sunulmuştur.

Farklı Test Uzunluklarına ve Örneklem Büyüklüklerine Göre PMTK ve POMTK'ya Dayalı Olarak Yapılan Yetenek Kestirimleri

Farklı test uzunluğu koşulları ve örneklem büyüklüğü koşullarına göre oluşturulan dokuz veri seti için öncelikli olarak PMTK'ya göre yetenek kestirimleri yapılmış ve ilk olarak veri setleri için model veri uyumu incelenmiştir. Her veri setinin bir, iki ve üç parametrelili lojistik model olan uyumu incelenmiş ve elde edilen sonuçlar Çizelge 62'de sunulmuştur.

Çizelge 62.

Farklı Örneklem Büyüklüğü ve Madde Sayılarından Oluşan Veri Setlerinin PMTK Modellerine Uyumunun İncelenmesi

Veri Seti	Madde Sayısı	Örneklem Büyüklüğü	Uyumlu Model	Uyumlu Madde Sayısı
1	5	500	1 PLM	Yok
2	5	1000	1 PLM	Yok
3	5	3000	1 PLM	Yok
4	15	500	2 PLM	8
5	15	1000	1 PLM	6
6	15	3000	1 PLM	1
7	25	500	2 PLM	20
8	25	1000	2 PLM	19
9	25	3000	1 PLM	4

Çizelge 62’de oluşturulan veri setlerinin PMTK modellerine olan uyumlarını belirten madde sayıları verilmiştir. Çizelgede yer alan değerler incelendiğinde beş madden oluşan üç veri seti için hiçbir PMTK modelinin uyumlu olmadığı görülmektedir. Bu veri setleri için bir, iki ve üç PLM uyumlar incelenmiş ancak madde düzeyindeki p değerlerinin tüm veri setlerinde 0.05’in altında olduğu bulunmuştur. Ölçek için hesaplanan loglikelihood değerleri arasındaki farklar incelendiğinde ise, her üç veri seti için de iki parametrelili lojistik modelin daha iyi uyum gösterdiği belirlenmiştir. Madde sayılarına göre oluşturulan veri setlerinin model veri uyumları incelendiğinde ise, 15 ve 25 maddeden oluşan veri setlerinde, en yüksek uyumun 500 ve 1000 kişilik veri setlerinden oluşan verilerden elde edildiği ve 3000 kişilik veri setlerindeki uyumlu madde sayısının daha az olduğu görülmektedir. Bu bulgu çalışma kapsamında yanıt aranan ilk araştırma sorusunda elde edilen bulgu ile tutarlıdır. Her iki örnekte de bir ve iki parametrelili lojistik modele göre kestirimler yapılmış ve kestirilen yetenekler arasındaki korelasyonlar hesaplanmıştır.

Farklı sayıda maddeden ve büyüklükteki örneklemlerden oluşan veri setleri için POMTK’ya dayalı analizler yapılarak, veri setlerinin POMTK modelleri kapsamında ölçeklemeye uygunluğu incelenmiştir. Analizlere beş maddelik veri setlerine ilişkin ölçeklenebilirlik katsayılarının hesaplanması ile başlanmış ve her üç örneklem büyüklüğü için de maddeler arası kovaryansa dayalı olarak hesaplanan H_{ij} değerlerinin pozitif olduğu belirlenmiştir. Madde düzeyinde bilgi veren H_i katsayıları incelendiğinde ise, 500 ve 1000 kişilik veri setlerinde yalnızca 1 nolu maddenin 0.3’ten düşük bir değer aldığı (0.279)

bulunmuştur. Diğer dört maddenin H_i katsayıları ise sınır değer olan 0.3'ten büyük olarak bulunmuştur. Ölçek düzeyinde bilgi veren H katsayısı incelendiğinde ise, beş maddelik 500, 1000 ve 3000 kişiden oluşan veri setleri için sırasıyla 0.359, 0.395 ve 0.364 olarak hesaplanmıştır. Dolayısıyla tüm veri setlerinin Mokken ölçeklemeye düşük düzeyde uyumlu olduğu, ancak 1000 kişilik veri setinin diğer örneklere göre daha iyi uyum sağladığı belirlenmiştir. Ayrıca madde tepki fonksiyonlarının monoton azalmayan nitelikte olduğu belirlenmiş ve OMSİ ile incelenen tek boyutluluk varsayımının tüm maddeler için karşılandığı belirlenmiştir. Son olarak veri setleri için kesişmeyen madde tepki fonksiyonları varsayımı madde kalan puanı grubu ve p matrisi yöntemleri ile incelenmiştir. 1000 kişilik veri seti her iki yöntemde söz konusu varsayımı sağlarken, 500 ve 3000 kişilik veri setleri için bu varsayımın karşılanmadığı belirlenmiştir. Dolayısıyla 500 ve 3000 kişilik veri setleri için POMTK kapsamında yer alan Monoton Homojenlik Modelinin, 1000 kişilik veri setinin ise, Çift Monotonluk Modeline uygun olduğu belirlenmiştir.

İncelemeye alınan bir diğer test uzunluğu koşulu olan 15 maddeden ve 500, 1000 ve 3000 kişiden oluşan veri setlerinin POMTK kapsamında model veri uyumunun incelenmesine geçilmiştir. Model veri uyumunun incelenmesine H katsayılarının hesaplanması ile başlanmış ve hesaplanan H_{ij} değerlerinin incelendiğinde 500 kişilik veri setinde 1 ve 12 nolu maddeler için hesaplanan değer negatif olduğu, diğer tüm maddeler ve örneklemler için H_{ij} katsayılarının pozitif olduğu belirlenmiştir. Ölçek düzeyinde bilgi veren H katsayıları incelendiğinde ise, 500 ve 3000 kişilik veri setlerinde 0.352 olarak, 1000 kişilik veri setinde ise 0.376 olarak bulunmuştur. Dolayısıyla veri setlerinin ölçeklemeye zayıf düzeyde uyum sağladığı sonucuna ulaşılmış ve madde tepki fonksiyonlarının monotonluğu incelenmiştir. Üç veri setinde yer alan tüm maddelerin madde tepki fonksiyonlarının monoton azalmayan nitelikte olduğu belirlenmiştir. Veri setlerinin tek boyutluluğu OMSİ ile incelenmiş, 500 kişilik veride 1, 10 ve 12 nolu maddelerin, 1000 ve 3000 kişilik veri setlerinde ise 12 nolu maddenin tek boyutluluğu ihlal ettiği bulunmuştur. Sınır değer olarak alınan 0.3 değeri, 0.2'ye çekilerek analizler tekrarlanmış ve 1000 kişilik veri setindeki tüm maddelerin varsayımı karşıladığı fakat 500 ve 3000 kişilik veri setlerinde 12 nolu maddenin varsayımdan olan ihlalinin devam ettiği belirlenmiştir. Dolayısıyla bu madde söz konusu veri setleri dışında bırakılarak analizler tekrar edilmiş ve bu kez tüm veri setleri için bu varsayımın karşılandığı belirlenmiştir. Bir

diğer varsayım olan, kesişmeyen madde tepki fonksiyonu varsayımı daha önce belirtilen iki yönteme göre incelenmiş ve bu varsayımın hiçbir veri seti için karşılanmadığı belirlenmiştir. Bu nedenle üç veri seti için de MHM'nin uygun olduğu sonucuna ulaşılmıştır.

Çalışma kapsamında incelenen test uzunluğuna ilişkin son koşul ise 25 maddeden oluşan veri setleridir. Beş yüz, 1000 ve 3000 kişiden oluşan 25 maddelik veri setlerinin POMTK'ya olan uyumları incelenirken, öncelikle madde çiftleri için ölçeklenebilirlik katsayıları hesaplanmıştır. Yalnızca 500 maddelik veri setinde 1 ve 12 nolu maddeler için hesaplanan H_{ij} katsayılarının negatif olduğu bulunmuştur ve bu sonuç diğer veri setlerinden elde edilen sonuçlarla örtüşmektedir. Diğer örneklem büyüklükleri için hesaplanan tüm H_{ij} katsayılarının pozitif olduğu bulunmuştur. Madde düzeyinde bilgi veren H_i katsayıları incelendiğinde ise, 500 kişilik veri setinde yedi maddenin (M1, M3, M10, M122, M12, M20 ve M25), 1000 kişilik veri seti iki maddenin (M12 ve M25) ve 3000 kişilik veri setinde ise dört maddenin (M1, M10, M12 ve M25) 0.3'ten düşük değerlere sahip olduğu belirlenmiştir. Ölçek düzeyinde bilgi veren H katsayıları incelendiğinde ise, 500 kişilik veri setinde 0.328, 1000 kişilik veri setinde 0.379 ve 3000 kişilik veri setinde ise 0.367 değerinin elde edildiği belirlenmiştir. Bu sonuçlara göre, POMTK kapsamında ölçeklemeye en uygun veri setinin 1000 kişilik veri seti olduğu ve her üç veri setinin de POMTK yaklaşımına zayıf düzeyde uyum gösterdiği bulunmuştur.

Veri setlerinin tek boyutluluk varsayımına uyumu OMSİ yöntemi ile incelenmiş ve 0.3 değeri sınır olarak alındığında 500 kişilik veri setinde dört maddenin (M1, M10, M20 ve M25), 1000 ve 3000 kişilik veri setlerinde ise iki maddenin (M12 ve M25) tek boyutluluğu ihlal ettiği belirlenmiştir. Ardından sınır değer 0.2 olarak değişimlenmiş ve analizler tekrarlanmıştır. Tekrarlanan analiz sonuçlarına göre, her üç veri setinde de 12 nolu madde uyumsuz olarak belirlenmiş ve bu madde tüm veri setlerinden çıkarılmıştır. Tek boyutluluk varsayımından sonra, monotonluk varsayımı incelemeye alınmış ve tüm veri setleri için bu varsayımın karşılandığı belirlenmiştir. Son olarak veri setlerinde yer alan maddelerin kesişmeyen madde tepki fonksiyonlarına sahip olup olmadığı incelenmiş ve iki ayrı yönteme göre elde edilen sonuçlara göre, tüm veri setlerinde söz konusu varsayımdan ciddi ihlallerin olduğu belirlenmiştir. Dolayısıyla 25 maddeden oluşan tüm veri setlerinin MHM'ye göre ölçeklenebileceği belirlenmiş ve hem PMTK hem de POMTK kapsamında

uygun olan modellere göre yetenek kestirimleri yapılmış ve kestirimler arasındaki korelasyonlar incelenmiştir.

İncelemeye alınan veri setlerinde PMTK ve POMTK yaklaşımlarına dayalı olarak yetenek kestirimleri yapılmış ve her iki yaklaşım kullanılarak elde edilen sonuçlar test uzunluğu koşullarına göre sunulmuş ve sunulmuş verilmiştir.

Test Uzunluğu Faktörüne Göre Yetenek Kestirimleri Arasındaki İlişkiler

Test uzunluğu ve örneklem büyüklüğü faktörlerine göre oluşturulan veri setlerinden kestirilen yetenek parametreleri arasındaki ilişkiler incelenmiş ve elde edilen bulgular öncelikli olarak test uzunluğu faktörüne göre sunulmuş, ardından örneklem büyüklüğü faktörüne göre incelenmiştir. Bu başlık altında test uzunluğu koşulları olan beş, 15 ve 25 maddelik veri setlerinden kestirilen yetenek parametreleri arasındaki ilişkiler bulunmaktadır.

Beş maddeden ve farklı örneklem büyüklüklerinden oluşan veri setlerinden kestirilen yetenek parametreleri arasındaki ilişkiler.

En düşük test uzunluğu koşulu olan beş maddelik veri setlerinden kestirilen yetenek parametrelerinin dağılımları incelenmiş ve normal dağılıma uygun olmadığı belirlenmiştir. Dolayısıyla her iki yaklaşımdan kestirilen yetenek parametreleri arasındaki ilişkiler Kendal'ın Tau Korelasyon Katsayısı kullanılarak hesaplanmıştır ve elde edilen değerler Çizelge 63'te sunulmuştur.

Çizelge 63.

Beş Maddeden ve Farklı Örneklem Büyüklüklerinden Oluşan Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler

Kendal'ın Tau KK	PMTK			
	Örneklem	500	1000	3000
POMTK	500	-0.040		
	1000		-0.032	
	3000			0.931**

**p<0.01

Çizelge 63 incelendiğinde, beş maddelik 500, 1000 ve 3000 kişiden oluşan veri setlerinden PMTK ve POMTK'ya göre kestirilen yetenekler arasındaki ilişkiler görülmektedir. Farklı örneklem büyüklüklerine göre, PMTK ve POMTK'dan yapılan kestirimler incelendiğinde, 500 ve 1000 kişilik veri setlerinden kestirilen yetenekler arasında manidar

bir ilişkinin olmadığı bulunmuş ve ilişki düzeyi de sıfıra oldukça yakın olarak hesaplanmıştır. Dolayısıyla her iki veri setinden kestirilen yetenekler arasında herhangi bir ilişkinin olmadığı sonucuna ulaşılmıştır. İncelenen son örneklem büyüklüğü koşulu olan 3000 kişilik veri setlerinden kestirilen yetenekler arasında ise çok yüksek ve manidar bir ilişki olduğu bulunmuştur. Dolayısıyla testin beş maddeden oluşması durumunda yalnızca örneklem büyüklüğünün 3000 olduğu durumda her iki yaklaşımdan kestirilen parametreler arasında anlamlı ve yüksek ilişki olduğu belirlenmiştir.

Beş maddelik veri setlerinden her iki yaklaşıma göre kestirilen yeteneklerin ardından 15 maddelik veri setlerine ilişkin yapılan yetenek kestirimlerine geçilmiştir.

On beş maddeden ve farklı örneklem büyüklüklerinden oluşan veri setlerinden kestirilen parametreler arasındaki ilişkiler

Çalışma kapsamında incelemeye alınan bir diğer test uzunluğu koşulu 15'tir. On beş maddeden oluşan veri setlerinden PMTK ve POMTK modelleri ile kestirilen yetenekler incelenmiş ve dağılımların normallik varsayımını karşıladığı ancak veri setlerinin POMTK kapsamında MHM ile ölçeklendiği dolayısıyla sıralama ölçeği düzeyinde bilgi verdiği göz önünde bulundurularak yetenekler arasındaki ilişkiler parametrik olmayan korelasyon katsayılarının biri olan Kendal'in Tau Katsayısı ile incelenmiştir. Elde edilen sonuçlar Çizelge 64'te verilmiştir.

Çizelge 64.

On Beş Maddeden ve Farklı Örneklem Büyüklüklerinden Oluşan Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler

Kendal'in Tau KK	PMTK			
	Örneklem	500	1000	3000
POMTK	500	-0.011		
	1000		-0.036	
	3000			1.00**

**p<0.01

Çizelge 64'te yer alan korelasyon katsayıları incelendiğinde, bir önceki test uzunluğu koşulu olan beş maddelik veri setlerinden elde edilen değerlere çok benzer olduğu görülmektedir. Beş maddelik veri setlerinde olduğu gibi, on beş maddelik veri setlerinden PMTK ve POMTK kapsamında 500 ve 1000 kişilik veri setlerinden kestirilen yetenekler arasında herhangi bir ilişkinin olmadığı belirlenmiştir. Her iki yaklaşımdan yalnızca 3000 kişilik veri setlerinden kestirilen yetenekler arasında mükemmel düzeyde manidar bir ilişki

olduğu bulunmuştur. Dolayısıyla 15 maddelik test uzunluğu koşulunda da yalnızca örneklem büyüklüğünün 3000 olduğu durumda her iki yaklaşımdan kestirilen yetenek parametreleri arasında manidar ve yüksek ilişkiler ve 15 maddelik test uzunluğunun her iki yaklaşımda da yetenek kestirimine uygun olmadığı sonucuna ulaşılmıştır.

Yirmi beş maddeden ve farklı örneklem büyüklüklerinden oluşan veri setlerinden kestirilen parametreler arasındaki ilişkiler

Test uzunluğu faktörü kapsamında incelemeye alınan son koşul olan 25 maddeden oluşan 500, 1000 ve 3000 kişilik veri setlerinden PMTK ve POMTK'ya dayalı olarak bireylere ilişkin yetenekler parametreleri kestirilmiştir. Hesaplanan yetenek parametreleri arasındaki ilişkiler her iki yaklaşıma göre incelenmiştir. Söz konusu veri setleri POMTK kapsamında MHM'ye göre ölçeklendiği için elde edilen yetenek parametreleri sıralama ölçeği düzeyindedir, bu nedenle yetenek parametreleri arasındaki ilişkiler Kendal'ın Tau Katsayısı hesaplanarak incelenmiştir. İncelemeler sonucu elde edilen değerler Çizelge 65'te verilmektedir.

Çizelge 65.

Yirmi Beş Maddeden ve Farklı Örneklem Büyüklüklerinden Oluşan Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler

Kendal'm Tau KK	PMTK			
	Örneklem	500	1000	3000
POMTK	500	0.939**		
	1000		0.954**	
	3000			1.00**

**p<0.01

Çizelge 65'te verilen 25 maddelik veri setlerine ilişkin kestirilen korelasyonlar incelendiğinde, incelemeye alınan iki yaklaşımdan kestirilen parametreler arasındaki ilişkilerin tümünün yüksek ve manidar olduğu görülmektedir. Bu durum daha önce incelenen test uzunluğu koşulları olan beş ve 15 maddelik veri setlerinden elde edilen bulgularla tamamen terstir. Hem beş hem de 15 maddelik veri setlerinden PMTK ve POMTK modellerine dayalı olarak kestirilen yetenekler arasında, 3000 kişilik veri seti hariç, sıfıra yakın ve manidar olmayan ilişkiler gözlenmiştir. Madde sayısının 25'e çıkması ile bu durumun tamamen tersi ortaya çıkmış ve tüm örneklerde iki yaklaşıma göre kestirilen yetenekler arasında +1.00'e çok yakın ve manidar ilişkiler gözlenmiştir. Dolayısıyla, madde sayısının artış göstermesi ile tüm örneklerde PMTK ve POMTK

yaklaşımlarından kestirilen yetenek parametreler tutarlı hale gelmiştir. Bu bulgulardan şu sonuçlara ulaşılmaktadır:

- Testte yer alan madde sayısının beş ve 15 olduğu durumlarda ise, PMPK kapsamında kestirilen yetenekler ile POMTK kestirimleri ile tutarlı sonuçlar veren tek veri seti 3000 kişiden oluşan veri setidir. Dolayısıyla PMPK'ya dayalı kestirimlerde madde sayısının düşük olduğu durumlarda, örneklem büyüklüğünün 1000'den yüksek olması kestirimleri tutarlı hale getirmektedir.
- Madde sayısı her iki yaklaşım için de yetenek kestirimleri için oldukça önemli bir faktördür. Hem madde sayısının hem de örneklem büyüklüğünün az olduğu durumlarda iki yaklaşımdan kestirilen yetenekler arasında herhangi bir ilişki olmadığı gözlenmiştir.

Test uzunluğu faktörü altında yer alan koşullara ilişkin yapılan incelemeleri takiben, oluşturulan veri setlerinden kestirilen yetenek parametreleri örneklem büyüklüğü koşuluna göre analiz edilmiş ve elde edilen bulgular sunulmuştur.

Örneklem Büyüklüğü Faktörüne Göre Yetenek Kestirimleri Arasındaki İlişkiler

Daha önceki aşamada test uzunluğu faktörüne göre sunulan bulgular, bu aşamada örneklem büyüklüğü faktörüne göre düzenlenerek sunulmuş, böylelikle örneklem büyüklüğü faktörünün yetenek kestirilerine etkisini ön plana çıkarmak amaçlanmıştır. En küçük örneklem büyüklüğü olan 500 kişilik veri setlerinden başlanarak, sırasıyla 1000 ve 3000 kişilik veri setlerinden PMPK ve POMTK'ya göre kestirilen yetenekler arasındaki ilişkiler incelenmiş ve elde edilen bulgular sunulmuştur.

500 kişilik veri setinden ve farklı madde sayılarından oluşan veri setlerinden kestirilen yetenek parametreleri arasındaki ilişkiler

Farklı madde sayılarından oluşan 500 kişilik veri setlerinden elde edilen yetenek parametrelerinin arasındaki ilişkiler Kendal'ın Tau Katsayısı kullanılarak hesaplanmış ve elde edilen değerler Çizelge 66'da verilmiştir.

Çizelge 66.

Farklı Madde Sayılarının Oluşan Beş Yüz Kişilik Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler

Kendal'in Tau KK	PMTK			POMTK		
	Madde Sayısı	5	15	25	5	15
PMTK	15	0.705**				
	25	0.009	-0.005			
POMTK	5	-0.040	-0.053	0.660**		
	15	-0.005	-0.011	0.820**	0.731**	
	25	0.005	-0.005	0.939**	0.678**	0.857**

**p<0.01

Çizelge 66'da beş, 15 ve 25 maddeden oluşan 500 kişilik veri setleri için hesaplanan katsayılar kullanılan yaklaşımlara göre incelenecek olursa, PMTK kapsamında beş ve 15 maddelik veri setlerinden kestirilen yetenekler arasında orta düzeyde (0.705) ve manidar bir ilişki olduğu belirlenmiştir. Ancak 25 maddeden oluşan veri seti için durum böyle değildir. Söz konusu veri setinden kestirilen yetenek parametreleri ile beş ve 15 maddelik veri setlerinden kestirilen yetenekler arasındaki ilişkinin sıfırdan farklı olmadığı görülmüştür. Dolayısıyla PMTK kapsamında, 500 kişilik veri setlerinde beş ve 15 maddenin birey yetenek kestirimlerini benzer düzeyde etkilediği, ancak 25 maddelik veri setinden kestirilen yeteneklerin, bu veri setlerinden kestirilen yeteneklerden farklı olduğu sonucuna ulaşılmıştır. POMTK kapsamında yapılan analizlerde ise, üç farklı madde sayısı koşulu için de manidar korelasyonlar olduğu belirlenmiştir. Hesaplanan katsayılar bakıldığında, en yüksek korelasyonun 15 ile 25 (0.857) maddelik veri seti arasında, en düşük korelasyonun ise beş ile 25 (0.678) maddelik veri seti arasında bulunduğu görülmektedir. Beş ve on beş maddelik veri setleri arasındaki korelasyon (0.731) ise, beş ve 25 maddelik veri seti arasındaki korelasyondan (0.678) yüksek bulunmuştur. Dolayısıyla POMTK kapsamında yapılan analizlerde, madde sayısının artış göstermesinin kestirimler arasındaki ilişkiyi arttırdığı yorumuna gidilebilir. Özellikle 15 ve 25 maddelik veri setlerinden kestirilen birey yetenekleri birbirlerine çok yakındır.

İki ayrı yaklaşım olarak incelenen PMTK ve POMTK'dan kestirilen yetenek parametreleri arasındaki ilişkiler incelendiğinde ise, PMTK kapsamında 25 maddelik veri setinden kestirilen yeteneklerin POMTK kapsamındaki her üç veri setiyle manidar ilişkiler gösterdiği, diğer test uzunluğu koşulları olan beş ve 15 maddelik veri setlerinden kestirilen

yeteneklerin ise, POMTK ile kestirilen yeteneklerle ilişkisiz olduğu belirlenmiştir. PMTK ile POMK'ya dayalı kestirimler arasında gözlenen en yüksek ilişki, 25 maddelik veri setlerinden kestirilen yetenekler arasındadır. Aynı veri setinin POMTK'ya göre 15 maddelik veri setinden kestirilen yetenekler ile ilişkisi de yüksek, (0.820) ancak beş maddelik veri setinden kestirilen yetenekler ile ilişkisi düşük (0.660) olarak bulunmuştur. Beş yüz kişilik örneklem büyüklüğü koşulundan kestirilen yetenekler arasındaki ilişkileri özetlemek gerekirse, şu sonuçlara ulaşılabilir:

- PMTK'ya göre yalnızca beş ve 15 maddelik veri setlerinden kestirilen yetenekler arasında orta düzeyde ve manidar korelasyonlar olduğu belirlenmiştir. Dolayısıyla 500 kişilik veri seti için madde sayısının beş ve 15 olması kestirimler için yetersizdir.
- POMTK'ya göre kestirilen yetenekler arasındaki korelasyon katsayısı madde sayısının artışıyla beraber artış göstermektedir.
- Her iki yaklaşımdan kestirilen yetenekler arasındaki en yüksek ilişki 25 maddelik veri seti için hesaplanmıştır. Dolayısıyla 500 kişilik veri setinde madde sayısı 25 olduğunda her iki yaklaşımın da benzer yetenek kestirimlerinde bulunduğu sonucuna ulaşılabilir.

1000 kişilik veri setinden ve farklı madde sayılarından oluşan veri setlerinden kestirilen parametreler arasındaki ilişkiler

Örneklem büyüklüğü faktörü koşulu kapsamında incelenen ikinci koşul olan 1000 kişilik veri setlerinden PMTK ve POMTK kullanılarak yetenek kestirimleri yapılmıştır. Farklı madde sayılarından oluşan 1000 kişilik veri setlerinden elde edilen yetenek parametrelerinin arasındaki ilişkiler Kendal'ın Tau Katsayısı kullanılarak hesaplanmış ve elde edilen değerler Çizelge 67'de verilmiştir.

Çizelge 67.

Farklı Madde Sayılarından Oluşan Bin Kişilik Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler

Kendal'ın Tau KK	PMTK			POMTK		
	Madde Sayısı	5	15	25	5	15
PMTK	15	0.706**				
	25	0.036	-0.022			

(Devam ediyor.)

Çizelge 67. (devam)

Farklı Madde Sayılarından Oluşan Bin Kişilik Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler

Kendal'in Tau KK	PMTK			POMTK	
POMTK	5	-0.032	-0.032	0.691**	
	15	-0.048	-0.036	0.845**	0.744**
	25	-0.040	-0.025	0.954**	0.705** 0.875**

**p<0.01

Çizelge 67'de yer alan 1000 kişilik veri setlerinden PMTK ve POMTK modelleri ile kestirilen yetenekler arasındaki ilişkileri belirten korelasyon katsayıları, iki yaklaşıma göre ayrı ayrı incelenecek olursa, PMTK modelleri ile beş ve 15 maddelik veri setlerinden kestirilen yetenekler arasında orta düzeyde ve manidar bir ilişki olduğu görülmektedir. Otuz maddeden oluşan veri setinden kestirilen yetenek parametreleri ile PMTK kapsamında incelenen diğer iki veri setinden kestirilen yetenekler arasında herhangi bir korelasyonun olmadığı belirlenmiştir. POMTK kapsamında yer alan MHM ile kestirilen yetenekler arasındaki ilişkiler incelendiğinde ise beş ve 15 ve beş ile 25 maddelik veri setleri arasında manidar ve orta düzeyde (0.744; 0.705) ilişkiler olduğu belirlenmiştir. Diğer test uzunluğu koşulu olan 25 maddelik veri seti ile 15 maddelik veri setinden kestirilen yetenekler arasında ise, yüksek ve manidar bir ilişki olduğu (0.875) bulunmuştur. Her iki yaklaşımdan kestirilen yetenekler arasındaki ilişkiler incelendiğinde ise, bir diğer örneklem büyüklüğü koşulu olan 500 maddelik veri setlerinden elde edilen sonuçlara benzer sonuçlar bulunduğu görülmektedir. PMTK 'ya göre yalnızca 25 maddeden oluşan veri setinden kestirilen yetenekler ile POMTK kapsamında tüm veri setlerinden kestirilen yetenekler arasında manidar ve orta ve yüksek düzeyde ilişkiler bulunmuştur. İncelenen veri setleri arasındaki en yüksek ilişki her iki yaklaşıma göre de 25 maddelik veri setlerinden kestirilen yetenekler parametreleri arasında bulunmuştur. PMTK ile 25 maddelik veri setinden kestirilen yetenekler ile beş maddelik veri setinden kestirilen yetenekler arasında orta düzeyde bir ilişki olduğu belirlenmiştir. Elde edilen bulgular özetlenecek olursa, şu sonuçlara ulaşılabilir:

- PMTK'ya göre yalnızca beş ve 15 maddelik veri setlerinden kestirilen yetenekler arasında orta düzeyde ve manidar korelasyonlar olduğu belirlenmiştir. Dolayısıyla

1000 kişilik veri seti için madde sayısının beş ve 15 olması kestirimler için yetersiz olduğu sonucuna ulaşılmıştır.

- POMTK'ya göre kestirilen yetenekler arasındaki ilişkiler, PMTK'dan kestirilen parametreler arasındaki ilişkilere göre daha tutarlıdır ve madde sayısının artışıyla beraber yükselmiştir; 15 ve 25 maddelik veri setlerinden kestirilen yetenekler arasındaki ilişki orta düzeyde korelasyonu gösterirken, beş ile 15 ve 25 maddelik veri setleri arasında hesaplanan korelasyon katsayısı yüksek ilişkiyi göstermektedir.
- Her iki yaklaşımdan kestirilen yetenekler arasındaki en yüksek ilişki 25 maddelik veri seti için hesaplanmıştır. Dolayısıyla 1000 kişilik veri setinde madde sayısı 25 olduğunda her iki yaklaşımın da benzer sonuçlar verdiği sonucuna ulaşılabilir.

3000 kişilik veri setinden ve farklı madde sayılarından oluşan veri setlerinden kestirilen parametreler arasındaki ilişkiler

Örneklem büyüklüğü faktörü koşulu kapsamında incelenen son koşul 3000 kişilik veri setidir. Beş, 15 ve 25 maddeden oluşan 3000 kişilik üç veri seti için PMTK ve POMTK modelleri kullanılarak yetenek kestirimleri yapılmıştır. Söz konusu veri setlerinden kestirilen yetenek parametreleri arasındaki ilişkiler Kendal'ın Tau Katsayısı kullanılarak hesaplanmış ve elde edilen değerler Çizelge 68'de verilmiştir.

Çizelge 68.

Farklı Madde Sayılarının Oluşan Üç Bin Kişilik Veri Setlerinden PMTK ve POMTK'ya Göre Kestirilen Yetenek Parametreleri Arasındaki İlişkiler

Kendal'in Tau KK	PMTK			POMTK		
	Madde Sayısı	5	15	25	5	15
PMTK	15	0.702**				
	25	0.666**	0.875**			
	5	0.931**	0.736**	0.693**		
POMTK	15	0.702**	10.00**	0.875**	0.736**	
	25	0.666**	0.875**	1.00**	0.693**	0.875**
	5					

**p<0.01

Çizelge 68'de yer alan 3000 kişilik veri setlerinden PMTK ve POMTK modelleri ile kestirilen yetenekler hesaplanan korelasyon katsayıları iki yaklaşıma göre ayrı ayrı incelenecek olursa, PMTK modelleri ile kestirilen yetenekler arasında daha önceki örneklem büyüklüğü koşulu olan 500 ve 1000 kişilik veri setlerinden yapılan kestirimlerden farklı olarak orta düzeyde ve yüksek ilişkiler olduğu belirlenmiştir. Özellikle 15 ve 25

maddelik veri setlerinden kestirilen yetenekler arasında 0.875 düzeyinde ve manidar bir ilişki bulunmuştur. Diğer test uzunluğu koşulu olan beş maddelik veri setinden kestirilen yetenekler ile 15 ve 25 maddelik testlerden kestirilen yetenekler arasında orta düzeyde ve manidar ilişkiler olduğu bulunmuştur. Bu bulgulara göre, PMTK kestirimlerinde, örneklem büyüklüğünün artış göstermesi test uzunluğuna göre daha yüksek oranda kestirimlerin tutarlılığını artırdığı yorumuna gidilebilir.

POMTK kapsamından kestirilen yetenekler arasındaki ilişkiler incelendiğinde, diğer örneklem büyüklüklerinden elde edilen ilişkilere benzer sonuçların elde edildiği görülmektedir. Madde sayılarının yakın olmasının parametreler arasındaki ilişkinin düzeyini artırdığı görülmüştür. Bu duruma örnek olarak 15 ve 25 maddelik veri setleri arasındaki ilişkinin (0.875) beş ve 25 maddelik veri setlerinden kestirilen yetenekler arasındaki ilişki (0.693) daha yüksek olduğu gösterilebilir.

Son olarak farklı yaklaşımlardan kestirilen yetenek parametreleri arasındaki ilişkiler incelendiğinde ise, 3000 kişilik örneklem büyüklüğü için hesaplanan korelasyon katsayılarının daha önceki örneklem büyüklüklerinden kestirilen katsayılardan daha yüksek olduğu bulunmuştur. Özellikle test uzunluğunun eşit olduğu veri setlerinden kestirilen ilişkilerin mükemmel yakın ve manidar olduğu bulunmuştur. Dolayısıyla POMTK'dan kestirilen yetenekler arasındaki ilişkiye benzer şekilde, madde sayısının yakınlığı yetenek kestirimleri arasındaki ilişkiyi artırmıştır. En düşük korelasyon katsayısı ise PMTK ile beş maddelik veri setinden kestirilen yetenekler ile, POMTK ile 25 maddeden kestirilen yetenekler arasında belirlenmiştir, ancak bu ilişkinin düzeyi de 0.666 olarak hesaplanmış ve orta düzeye yakın bir ilişkiyi göstermektedir. Genel olarak örneklem büyüklüğüne göre kestirilen yetenekler arasında belirlenen ilişkiler özetlenecek olursa, şu sonuçlara ulaşılabilir:

- Örneklem büyüklüğünün 3000 olduğu durumda, madde sayısının çok az olduğu durumlarda dahi, PMTK kapsamında kestirilen yetenekler arasındaki ilişkiler manidar ve yüksek bulunmuştur. Dolayısıyla örneklem büyüklüğündeki artışın PMTK modellerine dayalı olarak yapılan yetenek kestirimlerinde önemli olduğu sonucuna ulaşılabilir.
- POMTK'ya dayalı kestirimler tüm örneklem büyüklüklerinde benzer kalmış, yalnızca, madde sayısının artmasıyla artış göstermiştir. Ayrıca her örneklem

büyükliğünde, beş ve 15 maddelik verilerinden kestirilen ilişki, beş ve 25 maddelik veri setlerinden kestirilen ilişkidenden daha yüksek bulunmuştur.

- Örneklem büyüklüğünün 500 ve 1000 olduğu durumlarda, PMTK ile yalnızca 25 maddelik veri setinden kestirilen yetenekler ile POMTK ile kestirilen yetenekler arasında yüksek ve manidar ilişkiler olduğu belirlenmiştir. Ancak örneklem büyüklüğünün 3000 olduğu durumda, beş maddelik veri setlerinin dahi POMTK ile kestirilen yeteneklerle yüksek ve manidar ilişkileri olduğu belirlenmiştir. Buradan hareketle, PMTK ile yapılacak kestirimlerde, tutarlı sonuçlara ulaşmak için ya örneklemin büyük olması ya da madde sayısının 25 ve üzeri olması gerektiği sonucuna ulaşılabilir. Yetenek kestirimlerinin farklı faktör ve koşullara göre ilişkisinin incelenmesinden elde edilen bulgular bu şekildedir ve araştırma kapsamında yanıt aranan bir diğer soru olan test güvenilirliklerinin incelenmesine geçilmiştir.

Çalışmada Yer Alan Faktörlere Göre Oluşturulan Veri Setlerinin Güvenirliklerinin PMTK ve POMTK Yaklaşımlarına Göre İncelenmesi

Çalışma kapsamında yanıt aranan son soru olan üçüncü araştırma sorusu şu şekildedir: Çalışma kapsamında incelenen faktörlere ilişkin oluşturulan veri setlerinin ölçeklendiği yaklaşıma göre güvenilirlik düzeyleri nasıldır? Bu araştırma sorusuna yanıt oluşturmak için , araştırmada PMTK ve POMTK'ya dayalı olarak madde ve yetenek parametreleri kestirimleri yapılan tüm veri setlerinin güvenilirlik hesaplamaları yapılmış ve elde edilen sonuçlar PMTK ve POMTK yaklaşımlarına göre yorumlanmıştır. PMTK yaklaşımına göre güvenilirlik kestirimleri için test bilgi fonksiyonları ve marjinal güvenilirlik katsayıları hesaplanmıştır. POMTK'ya göre ise, test düzeyinde güvenilirlik bilgisi veren dört ayrı katsayı hesaplanmıştır. Elde edilen bulgular öncelikle PMTK yaklaşımı için sunulmuş, ardından POMTK için hesaplanan güvenilirlik katsayıları verilmiştir.

PMTK'ya Dayalı Olarak Yapılan Güvenirlik İncelemeleri

Test uzunluğu, örneklem büyüklüğü ve maddelerin psikometrik nitelikleri faktörüne göre oluşturulan veri setlerinin PMTK'ya göre güvenilirlik incelemeleri için, her veri seti için test bilgi fonksiyonlarının oluşturuluş ve marjinal güvenilirlik katsayıları hesaplanmıştır.

Oluşturulan test bilgi fonksiyonları incelenen faktörlere göre gruplandırılarak çalışmanın ekler bölümünde verilmiştir (EK-1). Test bilgi fonksiyonlarının verdikleri maksimum bilgi düzeyleri ve bu düzeye karşılık gelen yetenek düzeyleri incelenen her faktör ve koşullarına göre sıralı olarak sunulmuştur. Elde edilen bulgular ilk olarak örneklem büyüklüğü faktörüne ve bu faktöre ilişkin koşullara göre oluşturulan veri setleri için sunulmuş, takiben test uzunluğu faktörüne ilişkin oluşturulan veri setleri için yapılan güvenilirlik incelemeleri verilmiştir.

Örneklem Büyüklüğü Faktörüne Göre Oluşturulan Veri Setlerine İlişkin Yapılan Güvenirlik İncelemeleri

Çalışmada örneklem büyüklüğü faktörü altında üç koşul incelemeye alınmıştır. Bunlar örneklem büyüklüğünün 500, 1000 ve 3000 olduğu durumlardır. Ayrıca bu faktöre ilişkin koşullar, test uzunluğu faktörü ile çaprazlanarak incelemeye alınmış ve oluşturulan 12 veri seti için kestirilen test bilgi fonksiyonları ve marjinal güvenilirlik katsayıları örneklem büyüklüğü koşullarına göre sıralı olarak sunulmuştur.

1. 500 kişilik veri setleri için kestirilen test bilgi fonksiyonları ve marjinal güvenilirlik katsayıları

Bu örneklem büyüklüğü koşulu altında, beş, 15, 25 ve 30 maddeden oluşan 500 kişilik veri setleri incelemeye alınmıştır. Model veri uyumu sonuçları hatırlanacak olursa, bu veri setlerinden yalnızca beş maddelik veri seti için 1 PLM'nin uyumlu olduğu, diğer tüm veri setleri için ise 2 PLM'nin uyumlu olduğu belirlenmiştir. Bu nedenle bu aşamada da test bilgi fonksiyonları 1 ve 2 PLM'ye göre oluşturulmuştur. Çalışmada incelemeye alınan en küçük örneklem büyüklüğü olan 500 kişilik veri setleri için oluşturulan test bilgi fonksiyonlarından elde edilen maksimum bilgi düzeyleri ve bu bilgi düzeyinin denk geldiği yetenek düzeyi ve hesaplanan marjinal güvenilirlik katsayıları Çizelge 69'da verilmiştir.

Çizelge 69.

500 Kişiden Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Veri setleri	Madde sayısı	Uyumlu PMTK Modeli	Maks.Bilgi	Yetenek Düzeyi	Marjinal Güvenirlik Katsayıları
1	5	1 PLM	2	0	0.643
2	15	2PLM	6	0	0.841

(Devam ediyor.)

Çizelge 69. (devam)

500 Kişiden Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Veri setleri	Madde sayısı	Uyumlu PMTK Modeli	Maks.Bilgi	Yetenek Düzeyi	Marjinal Güvenirlik Katsayıları
3	25	2 PLM	10	0-1	0.894
4	30	2 PLM	14	0- (-1)	0.906

Çizelge 69 incelendiğinde, madde sayısının en fazla olduğu dört nolu veri setinden en yüksek düzeyde bilgi elde edildiği görülmektedir. İkinci en yüksek bilginin elde edildiği veri seti ise 25 maddeden oluşan veri setidir. Bu veri seti için oluşturulan test bilgi fonksiyonu incelendiğinde, en yüksek bilginin 0 ile 1 yetenek düzeyi aralığında olduğu ancak 2 yetenek düzeyine yakın noktalarda da yüksek düzeyde bilgi verdiği görülmektedir. En düşük bilgi ise, madde sayısının en az olduğu beş maddelik veri setinden elde edilmiştir. Benzer bulgular veri setleri için hesaplanan marjinal güvenirlik indeksleri için de geçerlidir. En yüksek marjinal güvenirlik katsayısı 30 maddelik veri seti için hesaplanmışken, en düşük katsayı ise beş maddelik veri seti için sınır değer olan 0.70'in altında hesaplanmıştır. Bu bulgu beklenen bir bulgudur, testteki madde sayısının artışı testin verdiği bilgi düzeyini, dolayısıyla testin güvenirliliğini artırmaktadır (Embretson ve Reise, 2000) Çizelgede de belirtildiği üzere, tüm testler orta yetenek düzeyi olan, 0 ve etrafındaki yetenek düzeylerinde maksimum düzeyde bilgi vermiştir. Dolayısıyla söz konusu testlerin orta yetenek düzeyinde bulunan bireyleri ölçmede daha kullanışlı olduğu sonucuna gidilebilir.

Elde edilen test bilgi fonksiyonları incelendiğinde (EK-1), ulaşılan bir diğer bulgu ise, bu örneklem büyüklüğü koşulunda incelenen tüm veri setlerinden kestirilen test bilgi fonksiyonlarının, alt ve üst uç yetenek düzeylerinde sifıra yakın bilgi verdikleridir. Ayrıca madde sayısının artışının testten elde edilen bilgi düzeyine ek olarak, yetenek ranjını da genişlettiği bulunan bir başka sonuçtur. 25 ve 30 maddelik veri setlerine ilişkin kestirilen test bilgi fonksiyonları diğer test bilgi fonksiyonlarından farklı olarak daha geniş yetenek aralığında daha yüksek düzeyde bilgi vermektedir. Bu koşul altında incelenen tüm testlerin yalnızca orta yetenek düzeyleri için yüksek düzeyde bilgi sağladığı belirlenmiştir ve bir

diğer örneklem büyüklüğü olan 1000 kişilik veri setleri için yapılan güvenilirlik incelemelerine geçilmiştir.

2. 1000 kişilik veri setleri için kestirilen test bilgi fonksiyonları ve marjinal güvenilirlik katsayıları

Örneklem büyüklüğü faktörü altında incelenen bir diğer koşul 1000 kişilik veri setidir ve bu faktör altında incelenen dört veri setinin de PMTK kapsamında 2 PLM'ye uyum sağladığı çalışmanın önceki aşamalarında belirlenmiştir. Her veri setine ilişkin 2 PLM'ye göre test bilgi fonksiyonları oluşturulmuş ve her testte ilişkin kestirilen test bilgi fonksiyonların verdiği maksimum bilgi düzeyi ile yetenek düzeyi ve marjinal güvenilirlik katsayıları Çizelge 70'te sunulmuştur.

Çizelge 70.

1000 Kişiden Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Veri setleri	Uyumlu PMTK Modeli	Madde sayısı	Maks.Bilgi	Yetenek Düzeyi	Marjinal Güvenirlik Katsayıları
1	1 PLM	5	2.5	0- (-1)	0.612
2	1 PLM	15	7	0- (-1)	0.836
3	2 PLM	25	12	0- (-1)	0.908
4	2 PLM	30	15	0- (-1)	0.911

Çizelge 70'te yer alan bulgular incelendiğinde, bir önceki örneklem büyüklüğü koşulu olan 500 kişilik veri setlerinden elde edilen bulgularla benzerlik gösterdiği belirlenmiş fakat test bilgi fonksiyonlarına göre elde edilen maksimum bilgi düzeyinin 500 kişilik veri setlerine göre bu veri setlerinde daha yüksek olduğu bulunmuştur. Ancak elde edilen maksimum bilgi düzeyindeki yükselmenin düşük düzeyde olduğu göz önünde bulundurulduğunda, örneklemdaki birey sayısının 500'den 1000'e çıkmasının test bilgi fonksiyonlarının verdiği bilgi düzeyinde önemli bir etkisi olmadığı sonucuna ulaşılabilir. Ayrıca yine 500 kişilik veri setlerinde olduğu gibi, madde sayısının artışı ile testten elde edilen bilgi düzeyinin de artış gösterdiği belirlenmiştir. Söz konusu durum marjinal güvenilirlik katsayıları ile de desteklenmektedir. En yüksek marjinal güvenilirlik katsayısı yine 30 maddelik veri setinden kestrilmişken, en düşük katsayısı ise beş maddelik veri setinden hesaplanmıştır. Bu durumda beş maddelik veri setlerinin PMTK kapsamında çok düşük düzeyde bilgi verdiği ve kullanımının uygun olmadığı sonucuna ulaşılabilir.

Kestirilen test bilgi fonksiyonlarının şekli incelendiğinde (EK-1), tüm veri setleri için en yüksek bilginin 0 ile -1 yetenek düzeyinde olduğu belirlenmiştir bu durumun nedeni ise, testte yer alan maddelerin güçlük düzeylerinin genel olarak orta ve düşük düzeyde olmasıdır. Özellikle üst yetenek düzeyleri için hiçbir veri setinin bilgi vermediği oluşturulan fonksiyonlarda görülmektedir. Yetenek ranjına göre en geniş aralığa hitap eden veri seti, 25 maddelik veri setidir. Bu veri seti daha geniş yetenek aralığında diğer veri setlerine göre daha yüksek düzeyde bilgi vermektedir. Madde sayısı düşük olan beş ve 15 maddelik testlere ilişkin elde edilen test bilgi fonksiyonlarının ise daha sivri olduğu görülmektedir. Özetle, testte yer alan madde sayısının artışı testin verdiği bilgi düzeyini artırmakta iken, örneklem büyüklüğünün 500'den 1000'e çıkmasının testin verdiği bilgi düzeyi üzerinde önemli bir etkisinin olmadığı sonucuna ulaşılabılır.

Orta büyüklükte olarak nitelendirilebilecek 1000 kişilik veri setlerinden oluşan veri setleri için test bilgi fonksiyonlarının incelenmesinin ardından, geniş örneklem büyüklüklerini temsil etmesi amacıyla çalışmaya dahil edilen 3000 kişilik veri setlerinden kestirilen test bilgi fonksiyonlarının incelenmesine geçilmiştir.

3. 3000 kişilik veri setleri için kestirilen test bilgi fonksiyonları

Örneklem büyüklüğü faktörü altında incelenen son koşul 3000 kişilik veri setidir ve bu örneklem büyüklüğüne göre farklı test uzunluklarından oluşan veri setlerinin PMTK kapsamında yapılan model veri uyumu incelemelerinde 1 ve 2 PLM'ye uyum sağladığı belirlenmiştir. Söz konusu veri setlerine ilişkin belirtilen PMTK modelleri kullanılarak test bilgi fonksiyonları oluşturulmuş ve her testte ilişkin kestirilen test bilgi fonksiyonların verdiği maksimum bilgi düzeyi ile yetenek düzeyi ve marjinal güvenirlilik katsayıları Çizelge 71'de sunulmuştur.

Çizelge 71.

3000 Kişiden Oluşan Veri Setleri İçin Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlilik Katsayıları

Veri setleri	Uyumlu PMTK modeli	Madde sayısı	Maks. Bilgi	Yetenek Düzeyi	Marjinal Güvenirlilik Katsayıları
1	1 PLM	5	2.5	0- (-1)	0.574
2	1 PLM	15	7	0- (-1)	0.827
3	1 PLM	25	12	0- (-1)	0.883
4	2 PLM	30	12	0	0.908

Çizelge 71 incelendiğinde, 3000 kişiden oluşan dört veri setinden beş, 15 ve 25 maddelik veri setlerinin 1 PLM'ye, 30 maddelik veri setinin ise 2 PLM'ye uyum sağladığı görülmektedir. Oluşturulan test bilgi fonksiyonlarına dayalı olarak elde edilen bulgular incelendiğinde, diğer örneklem büyüklüğü koşulu olan 1000 kişilik veri setlerinden elde edilen bulgulara oldukça benzer olduğu bulunmuştur. Önceki bulgularla paralel olarak madde sayısının artması ile testten elde edilen bilgi düzeyi ve hesaplanan marjinal güvenirlik katsayıları artmıştır. Her veri seti için kestirilen test bilgi fonksiyonları incelendiğinde, daha önceki örneklem büyüklüğü koşullarında incelenen test bilgi fonksiyonlarına benzer oldukları görülmektedir (EK-1). Yalnızca 25 maddeden oluşan veri seti, önceki örneklem büyüklüğü koşullarından elde edilen test bilgi fonksiyonlarından biraz daha farklı formda bir test bilgi fonksiyonuna sahiptir. İncelenen 500 ve 1000 kişilik veri setlerinde iki yetenek düzeyinde maksimuma yakın düzeyde bilgi veren 25 maddelik veri seti, 3000 kişilik veri setinde normal dağılım eğrisine daha yakın bir şekle sahip olmuş ve tek bir noktada verdiği bilgi düzeyi maksimuma ulaşmıştır. Bu durumun nedeni olarak örnekleme yer alan birey sayısının artması gösterilebilir. Ancak örnekleme yer alan birey sayısı diğer koşullara göre oldukça yüksek olmasına rağmen, testlerden elde edilen bilgi düzeyinin diğer koşullara oldukça benzer, hatta 30 maddelik veri setlerinde, bu örneklem büyüklüğü koşulu için daha düşük olduğu bulunmuştur. Dolayısıyla bu bulgulara göre, örneklem büyüklüğünün testten elde edilen bilgi miktarı üzerinde çok etkili olmadığı sonucuna ulaşılabilir.

Örneklem büyüklüğü faktörü altında incelenen üç koşul için kestirilen tüm test bilgi fonksiyonlarından elde edilen bulgular incelendiğinde, testin verdiği bilgi düzeyinin madde sayısı ile doğrudan ilişkili olduğu ve örneklem büyüklüğü faktörünün testin verdiği bilgi düzeyi için madde sayısı kadar önemli olmadığı sonucuna ulaşılmıştır. Ayrıca incelenen tüm örneklem büyüklüğü koşullarından elde edilen bulgulara göre, beş maddelik veri setlerinin PMTK'ya göre ölçeklenmesinin uygun olmadığı, çünkü güvenirlik düzeyinin oldukça düşük olduğu belirlenmiştir. Örneklem büyüklüğüne göre değişmeksizin, 25 ile 30 maddelik veri setlerinden elde edilen maksimum bilgi düzeyi arasında oldukça düşük farklar olduğu, fakat 15 ile 25 maddeden elde edile maksimum bilgi düzeyleri arasındaki farkın ise oldukça yüksek olduğu bulunmuştur. Dolayısıyla 25 ile 30 madde arasında testin verdiği bilgi düzeyini etkileyecek büyüklükte fark olmadığı, ancak 15 ile 25 madde

arasındaki farkın testin verdiği maksimum bilgi düzeyi açısından büyük olduğu sonucuna ulaşılabilir. Ayrıca 15 ile 25 maddelik veri setleri için hesaplanan marjinal güvenilirlik indeksleri arasındaki farklar da 25 ve 30 maddelik veri setleri için hesaplanan indeksler arasındaki farklardan yüksektir. Madde sayısındaki artışın güvenliliği arttığı bu bulgu ile desteklenmekte ancak bu durum doğrudan maddelerin nitelikleri ile ilişkili olduğu da göz önünde bulundurulmalıdır (De Ayala, 2009; Embretson ve Reise, 2000). Elde edilen bulgulara göre beş ve 15 madde testin verdiği bilgi bakımından yetersiz madde sayısı olarak kabul edilebilir.

Örneklem büyüklüğü faktörü altında yer alan koşulların test puanlarının güvenilirliğine etkisinin incelenmesini takiben, çalışmada madde ve yetenek parametrelerine etkisi incelenen bir diğer faktör olan test uzunluğu altında yer alan koşullara göre yapılan güvenilirlik incelemelerine geçilmiştir.

Test Uzunluğu Faktörüne Göre Oluşturulan Testlere İlişkin Yapılan Güvenirlik İncelemeleri

Çalışma kapsamında incelemeye alınan test uzunluğu faktörü altında üç koşul ele alınmıştır. Bunlar test uzunluğunu beş, 15 ve 25 olduğu koşullardır. Ayrıca bu faktöre ilişkin koşullar, örneklem büyüklüğü faktörüne ait koşullar ile çaprazlanarak incelenmiş ve böylece bu faktöre göre oluşturulan dokuz ayrı veri seti için incelemeler yapılmıştır. Elde edilen test bilgi fonksiyonları ve marjinal güvenilirlik katsayıları test uzunluğu koşullarına göre sunulmuştur.

1. Beş maddelik veri setleri için kestirilen test bilgi fonksiyonları

Test uzunluğu faktörü altında çok kısa testleri temsil etmesi nedeniyle oluşturulan beş maddelik veri setleri üç ayrı örneklem büyüklüğünden oluşmaktadır. Beş maddeden ve 500, 1000 ve 3000 kişiden oluşan veri setlerinin PMTK modelleri kapsamında 1 PLM'ye uyumlu olduğu belirlenmiştir ve bu modele göre test bilgi fonksiyonları oluşturulmuş ve marjinal güvenilirlik katsayıları hesaplanmıştır. Kestirilen test bilgi fonksiyonlarından elde edilen maksimum bilgi düzeyleri ve bu bilgi düzeyinin karşılık geldiği yetenek düzeyleri ile marjial güvenilirlik katsayıları Çizelge 72'de verilmiştir.

Çizelge 72.

Beş Maddeden Oluşan Veri Setlerinden Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Veri setleri	Örneklem Büyüklüğü	Uyumlu PMTK modeli	Maks.Bilgi	Yetenek Düzeyi	Marjinal Güvenirlik Katsayıları
1	500	1 PLM	2	0- (-1)	0.643
2	1000	1 PLM	2.5	-1- (-2)	0.611
3	3000	1 PLM	2.5	-1- (-2)	0.573

Çizelge 72’de yer alan değerler incelendiğinde, her üç örneklem büyüklüğünde de testlerden elde edilen maksimum bilgi düzeylerinin ve marjinal güvenirlik katsayılarının çok düşük olduğu bulunmuştur. Ayrıca üç veri seti de maksimum bilgiyi negatif yetenek düzeylerinde vermektedir, bu durumun nedeni ise testte yer alan maddelerin güçlük düzeyinin ortalamasının altında olmasıdır. Örneklem büyüklüklerine göre veri setlerinden elde edilen bilgi düzeyleri incelendiğinde, bu test uzunluğu koşulu için örnekleme yer alan birey sayısının artması ile testin verdiği bilgi düzeyi arasında bir ilişki olmadığı, aksine en yüksek marjinal güvenirlik katsayısının 500 kişiden oluşan veri setinden hesaplandığı çizelgede görülmektedir. Özetle beş maddelik veri setlerinin PMTK kapsamında ölçeklenmeye uygun olmadığı, gerek güvenirlik düzeyinin altında kaldığı sonucuna ulaşılmıştır. Bu bulgulardan sonra bir diğer test uzunluğu koşulu olan 15 maddelik veri setleri için yapılan güvenirlik incelemelerine geçilmiştir.

2. On beş maddelik veri setleri için kestirilen test bilgi fonksiyonları ve marjinal güvenirlik katsayıları

On beş maddeden ve farklı örneklem büyüklüklerinden oluşan üç veri setinin, PMTK kapsamında bir ve iki parametrelili lojistik modele uyumlu olduğu bulunmuş ve söz konusu modeller kullanılarak veri setleri için test bilgi fonksiyonları kestirilmiş ve marjinal güvenirlik katsayıları hesaplanmıştır. Elde edilen test bilgi fonksiyonlarının verdikleri maksimum bilgi düzeyi ve buna karşılık gelen yetenek düzeyleri ile hesaplanan marjinal güvenirlik indeksleri Çizelge 73’te sunulmuştur.

Çizelge 73.

On Beş Maddeden Oluşan Veri Setlerinden Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Veri setleri	Örneklem Büyüklüğü	Uyumlu PMTK modeli	Maks.Bilgi	Yetenek Düzeyi	Marjinal Güvenirlik Katsayıları
1	500	2 PLM	6	0	0.841
2	1000	1 PLM	7	-1- (-2)	0.837
3	3000	1 PLM	6.5	0 - (-1)	0.827

Çizelge 73'te yer alan değerler incelenecek olursa, en düşük bilginin 500 kişilik veri setinden elde edildiği, en yüksek bilginin ise, 1000 kişilik veri setinden elde edildiği görülmektedir. Aynı durum marjinal güvenirlik katsayıları için de geçerli değildir, en düşük katsayı 3000 kişilik en yüksek katsayı ise 500 kişilik veri setinden hesaplanmıştır. Bu durumun nedeni 3000 kişilik veri setinin 500 kişilik veri setine göre yetenek düzeyi daha heterojen bireylerden oluşuyor olmasıdır. Grup yetenek düzeyi bakımından heterojenleştikçe elde edilen güvenirlik katsayılarında da düşüşler ortaya çıkmaktadır (Crocker ve Algina, 1986). Ayrıca test bilgi fonksiyonlarından elde edilen maksimum bilgi düzeyleri incelendiğinde bir önceki test uzunluğu koşulu olan beş maddelik veri setlerine göre elde edilen bilgi düzeyinin çok yüksek olduğu belirlenmiştir. Dolayısıyla madde sayısındaki artış beklenildiği üzere testten elde edilen maksimum bilgi düzeyini artırmıştır. Daha önce incelenen test bilgi fonksiyonlarında olduğu gibi, yine en yüksek bilginin karşılık geldiği yetenek düzeyi sıfır ve negatif yetenek düzeyleridir. Dolayısıyla testlerde yer alan maddelerin güçlük düzeylerinin düşük olduğu ve yeteneği ortalamanın altında olan bireyleri daha az hata ile ölçtüğü yorumuna ulaşılabilir.

Elde edilen test bilgi fonksiyonları incelendiğinde (EK-1) ulaşılan bir başka bilgi ise, tüm veri setlerinde uç yetenek düzeylerine ilişkin testin sağladığı bilgi düzeyinin çok düşük olduğudur. Veri setlerine göre inceleme yapılacak olursa, 1000 ve 3000 kişilik veri setleri için kestirilen test bilgi fonksiyonlarının benzerlik gösterdiği görülmektedir. 500 kişiden oluşan veri seti için kestirilen test bilgi fonksiyonu ise pozitif yetenek düzeylerinde diğer test bilgi fonksiyonlarına göre daha fazla bilgi vermektedir. Bir diğer ifadeyle, bu veri seti için kestirilen test bilgi fonksiyonu daha yavaş bir alçalma göstermekte, dolayısıyla daha geniş yetenek ranjında yüksek düzeyde bilgi sağlamaktadır. Ayrıca Çizelge 78'de de görüldüğü üzere, 1000 kişilik veri seti için kestirilen test bilgi fonksiyonunun sağladığı

bilgi az da olsa diğer veri setlerinden yüksektir ve bu durumun nedeni model veri uyumunun bu veri seti için daha yüksek düzeyde sağlanıyor olmasıdır (Baker ve Kim, 2004).

On beş maddelik test uzunluğu koşulu için elde edilen güvenilirlik kestirimlerinin incelenmesinin ardından son test uzunluğu koşulu olan 25 maddelik veri setlerine ilişkin kestirilen test bilgi fonksiyonlarının incelenmesine geçilmiştir.

3. Yirmi beş maddelik veri setleri için kestirilen test bilgi fonksiyonları ve marjinal güvenilirlik katsayıları

Test uzunluğu faktörü altında incelenen son koşul olan 25 maddeden oluşan veri setlerine ilişkin yapılan PMTK analizlerinde, 500 ve 1000 kişiden oluşan veri setlerinin 2 PLM'ye, 3000 kişilik veri setinin ise 1 PLM'ye uyum sağladığı belirlenmiştir. Söz konusu modeller kapsamında üç veri seti için test bilgi fonksiyonları kestirilmiş ve marjinal güvenilirlik katsayıları hesaplanmıştır. Kestirilen test bilgi fonksiyonlarının verdiği en yüksek bilgi düzeyi ve bu bilgi düzeyinin karşılık geldiği yetenek düzeyleri ile marjinal güvenilirlik katsayıları Çizelge 74'te verilmiştir.

Çizelge 74.

Yirmi Beş Maddeden Oluşan Veri Setlerinden Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Veri setleri	Örneklem Büyüklüğü	Uyumlu PMTK Modeli	Maks.Bilgi	Yetenek Düzeyi	Marjinal Güvenirlik Katsayıları
1	500	2 PLM	9	0- (-1)	0.894
2	1000	2 PLM	12	0- (-1)	0.908
3	3000	1 PLM	10	0 - (-1)	0.883

Çizelge 74'te verilen değerler incelendiğinde, daha önce incelenen test uzunluğu koşullarına benzer sonuçların elde edildiği görülmektedir. En fazla bilgiyi veren test 1000 kişiden oluşan veri seti iken, en az bilgi yine 500 kişiden oluşan veri seti için sağlanmıştır. Bu test uzunluğu koşulları için elde edilen en düşük bilgi düzeyi dahi, bir önceki test uzunluğu koşulu olan 15 maddelik tüm veri setlerinden kestirilen maksimum bilgi düzeylerinden yüksektir. Dolayısıyla örneklem büyüklüğü faktörü altında çalışılan veri setlerinde ve alanyazında da belirtildiği üzere, testte yer alan madde sayısının artması ile testin verdiği bilgi düzeyinin artış gözlenmektedir (Hambleton ve Swaminathan, 1985). Hesaplanan marjinal güvenilirlik katsayıları incelendiğinde ise, yine en yüksek katsayının

1000 kişiden oluşan veri seti için hesaplandığı ve 15 maddelik testlere göre hesaplanan katsayıların daha yüksek olduğu görülmektedir. En yüksek güvenilirlik indeksinin 15 maddelik veri setinden kestirilmesinin nedeni olarak ise model veri uyumunun en iyi sağlandığı veri seti olması gösterilebilir.

Yirmi beş maddelik test uzunluğu koşulu kapsamında incelenen tüm veri setlerinden kestirilen test bilgi fonksiyonları incelendiğinde, özellikle 500 ve 1000 kişilik veri setlerine ilişkin kestirilen test bilgi fonksiyonlarının oldukça benzer olduğu görülmektedir (EK-1). Söz konusu iki fonksiyon, 3000 kişilik veri setinden elde edilen fonksiyona göre daha geniş yetenek aralığında yüksek bilgi vermektedir. 3000 kişilik veri seti için kestirilen test bilgi fonksiyonu ise, diğerlerine göre normal dağılım eğrisine daha yakın bir formdadır ve verdiği maksimum bilgi düzeyi 500 kişilik veri setinden elde edilen maksimum bilgi düzeyi ile aynıdır. Ayrıca çalışma boyunca elde edilen tüm test bilgi fonksiyonlarında olduğu gibi, bu test uzunluğu koşulu için oluşturulan test bilgi fonksiyonları da uç yetenek düzeyine bilgi vermemektedir.

Test uzunluğu ve örneklem büyüklüğü faktörü altında yer alan koşullara ilişkin yapılan güvenilirlik incelemelerinden elde edilen bulgulara göre şu genel sonuçlara ulaşılmıştır:

- Testin verdiği bilgi düzeyi örneklem büyüklüğünden ziyade, test uzunluğundan etkilenmiştir.
- Model veri uyumunun daha iyi sağlandığı 1000 kişilik verilerden elde edilen bilginin ve hesaplanan marjinal güvenilirlik katsayılarının tüm madde sayısı koşullarında en yüksek olduğu görülmüştür.
- Beş maddelik veri setlerinin hiçbir örneklem büyüklüğünde yeterli güvenilirlik düzeyini sağlamadığı belirlenmiştir.
- Test uzunluğu faktörü içinde incelenen 25 ve 30 maddelik veri setleri için test bilgi fonksiyonlarına göre hesaplanan maksimum bilgi düzeylerinin birbirine çok yakın olduğu bulunmuştur.
- Farklı test uzunluğundan ve örneklem büyüklüklerinden oluşan veri setlerinin tümünün en fazla bilgi verdiği yetenek düzeylerinin sıfır ve negatif değerler etrafında olduğu belirlenmiştir. Bu durum da testte yer alan maddelerin uç yetenek düzeylerini ölçmede yetersiz kaldığını göstermektedir. Testte yer alan maddelerin

çoğunluğu, önceki araştırma sorularında incelendiği üzere orta güçlük düzeyindedir, dolayısıyla en fazla bilgiyi de orta yetenek düzeylerinde vermiştir.

Test uzunluğu ve örneklem büyüklüğü faktörlerine göre test bilgi fonksiyonlarının incelenmesinin ardından, son faktör olan maddelerin farklı psikometrik niteliklerine göre oluşturulan veri setlerine ilişkin yapılan güvenilirlik incelenmelerine geçilmiştir.

Maddelerin Farklı Psikometrik Nitelikleri Faktörüne Göre Oluşturulan Testlere İlişkin Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Araştırma kapsamında incelenen son faktör maddelerin güçlük ve ayırt edicilik parametrelerinin değiştiği durumları kapsamaktadır. İncelemeye alınan ana veri seti temel alınarak madde güçlük ve ayırt edicilik parametrelerinin yüksek ve düşük olduğu dört ayrı veri seti oluşturulmuştur. Söz konusu veri setlerinin tümü 1000 kişiden oluşmaktadır. Bu durumun nedeni çalışmanın daha önceki bölümlerinde de açıklandığı gibi en yüksek model veri uyumunun bu örneklem için sağlanmış olmasıdır. Uyumlu oldukları belirlenen PMTK modellerine göre oluşturulan veri setleri için test bilgi fonksiyonları kestirilmiş ve marjinal güvenilirlik katsayıları hesaplanmıştır. Test bilgi fonksiyonlarının verdiği maksimum bilgi düzeyi ve bu bilgi düzeyinin denk geldiği yetenek düzeyleri ile hesaplanan marjinal güvenilirlik katsayıları Çizelge 75’te verilmiştir.

Çizelge 75.

Madde Güçlük ve Ayırt Edicilik Parametrelerine Göre Değişen Maddelerden Oluşan Veri Setlerinden Kestirilen Test Bilgi Fonksiyonları ve Marjinal Güvenirlik Katsayıları

Veri setleri	Uyumlu PMTK modeli	Maksimum Bilgi	Yetenek Düzeyi	Marjinal Güvenirlik Katsayıları
A-yüksek	2 PLM	10	0	0.857
A-düşük	2 PLM	5	0	0.764
B-yüksek	3 PLM	15	1-2	0.719
B-düşük	1 PLM	6	0- (-1)	0.802

Çizelge 75’te yer alan veri setlerine ilişkin isimlendirmeyi hatırlatmak gerekirse; “A- en yüksek” madde ayırt edicilik parametresi yüksek olan maddelerden oluşan veri setini temsil etmekteyken, “A- en düşük”, madde ayırt edicilik değerleri düşük olan maddelerden oluşan veri setini ifade etmektedir. Aynı şekilde “B-en yüksek” olarak isimlendirilen veri seti, madde güçlük parametresi yüksek olan maddelerden oluşan veri setini belirtmekteyken, “B- en düşük” olarak belirtilen veri seti ise madde güçlük değerleri

düşük olan maddelerden oluşan veri setini temsil etmektedir. Veri setlerinin farklı PMTK modellerine göre ölçeklendiği çizelgede yer alan bilgilerdendir.

Test bilgi fonksiyonlarından elde edilen bilgiler veri setlerine göre incelendiğinde, madde ayırt edicilik parametrelerinin değiştiği veri setlerinden, a parametrelerinin yüksek olduğu veri setinden elde edilen bilgi düzeyinin, a parametresinin düşük olduğu veri setine göre daha yüksek olduğu görülmektedir. Madde ayırt edicilik indekslerinin artması ile testten elde edilen bilgi düzeyinin artış göstermesi beklenen bir durumdur ve alan yazında da bu şekilde belirtilmektedir (Emberson ve Reise, 2000; Hambleton ve Swaminathan, 1985). Dolayısıyla bu bulgu alan yazınla uyumludur. Ayrıca iki veri setinin de maksimum bilgiyi verdiği yetenek düzeyinin de 0 ve civarında olduğu belirlenmiştir.

Madde ayırt edicilik değerlerinin yüksek ve düşük olduğu gruplardan kestirilen test bilgi fonksiyonları incelendiğinde, her ikisinin de maksimum bilgiyi sıfır yetenek düzeyi etrafında verdiği ve uç yetenek düzeyleri için sağladıkları bilginin 0'a yakın olduğu görülmektedir (EK-1). Ancak a parametresi yüksek olan maddelerden oluşan teste ilişkin hesaplanan test bilgi fonksiyonunun düşük maddelerden oluşan teste göre daha sivri olduğu, yani verdiği maksimum bilgi düzeyinin daha yüksek olduğu belirlenmiştir. Bu bulgudan hareketle, alan yazında yer alan testte yer alan maddelerin ayırt edicilik düzeyleri ile testin verdiği bilgi düzeyi arasında pozitif korelasyon olduğu durumu doğrular niteliktedir (Baker ve Kim, 2004; Emberson ve Reise, 2000; Hambleton ve Swainathan, 1985).

Çizelge 75'te incelenen diğer veri setleri madde güçlük parametrelerinin yüksek ve düşük olduğu veri setleridir. Madde güçlük parametrelerinin yüksek olduğu veri setinin vermiş olduğu bilgi düzeyi çalışma kapsamında incelenen tüm veri setlerinden elde edilen maksimum bilgi düzeyinden, madde sayısının daha yüksek olduğu veri setleri de dâhil olmak üzere, daha yüksektir. Bu veri setinde yer alan maddelerin ayırt edicilik düzeylerinin de orta ve iyi düzeyde olduğu göz önünden bulundurulurken, madde güçlük indekslerinin yüksek olması, madde sayısının artmasına ya da madde ayırt edicilik indekslerinin yükselmesi göre, testin verdiği bilgi düzeyini daha fazla etkilediği yorumuna ulaşılabilir. Ayrıca bu veri seti incelenen tüm veri setleri içinde pozitif yetenek düzeyinde maksimum bilgi veren tek veri setidir ve bu durumun nedeni maddelerin daha üst yetenek düzeyindeki bireylere daha uygun olmasıdır.

İncelenen son veri ise, madde güçlük indekslerinin düşük olduğu veri setidir ve bu veri setine ilişkin kestirilen test bilgi fonksiyonu incelendiğinde, verdiği bilgi düzeyinin b parametrelerinin yüksek olduğu veri setine göre oldukça düşük olduğu görülmektedir. Bu fonksiyon en fazla bilgiyi 0 ile -2 arasında vermekte ve özellikle pozitif yetenek düzeylerinde, 2' den itibaren, çok düşük düzeyde bilgi vermektedir. Dolayısıyla madde güçlük düzeyinin artmasının testin verdiği maksimum bilgi düzeyini arttırdığı, düşmesinin ise testin verdiği maksimum bilgi düzeyin azalttığı sonucuna ulaşılabılır.

Ayırt edicilik ve güçlük düzeyleri değişen veri setleri için hesaplanan marjinal güvenilirlik katsayıları incelendiğinde ise, beklenildiği üzere ayırt edicilik değeri yüksek olan veri setinden hesaplanan güvenilirlik katsayısının diğer tüm veri setlerinden hesaplanan katsayılardan daha yüksek olduğu görülmektedir. Bu veri setini güçlük düzeyi düşük olan maddelerden oluşan veri seti takip etmektedir. Ayırt edicilik ve güçlük düzeyi düşük olan veri setleri için hesaplanan marjinal güvenilirlik katsayılarının ise, birbirine yakın ve düşük olduğu çizelgede görülmektedir. Hesaplanan en düşükmarjinal güvenilirlik katsayısı ise, test bilgi fonksiyonlarından elde edilen bulgulara zıt bir durum olarak, madde güçlüğüyle yüksek olduğu veri setinden kestirilmiştir. Bu bulgudan hareketle, her ne kadar güvenliliğin testin verdiği bilgi düzeyiyle ilişkisi olsa da, elde edilen maksimum bilgi düzeyinin yüksek olması, testin güvenliliğinin de yüksek olacağı anlamına gelmediği ve güvenliliğin testin verdiği bilgi düzeyinden farklı olarak değişebildiği sonucuna ulaşılmaktadır.

Maddelerin farklı psikometrik niteliklerine göre oluşturulan testlerden elde edilen test bilgi fonksiyonları incelendiğinde genel olarak şu sonuçlara ulaşılmıştır:

- Madde ayırt edicilik parametre değerleri yüksek olan maddelerden oluşan testlerden elde edilen maksimum bilgi düzeyi madde ayırt edicilik parametreleri nispeten düşük olan veri setlerinden elde edilen maksimum bilgi düzeyinden daha yüksek olduğu bulunmuştur.
- Madde güçlük parametreleri yüksek olan maddelerden oluşan testlerden elde edilen bilgi düzeyi diğer veri setlerine göre daha yüksek bulunmuştur.
- İncelenen veri setlerinde yer alan maddelerin ayırt edicilik parametreleri yüksek olsa dahi, testten elde edilen maksimum bilgi düzeyi yetenek ölçeğine orta düzeylere denk gelmektedir. Yalnızca madde güçlüğüünün artması, maksimum bilgi düzeyinin pozitif yetenek düzeylerinde olmasını sağlamıştır.

PMTK modellerine göre oluşturulan ve incelenen test bilgi fonksiyonlarından sonra, raporun bu kısmında PMTK yaklaşımı ile karşılaştırılan POMTK yaklaşımına göre ölçeklenen veri setlerine ilişkin yapılan güvenilirlik incelemeleri ve elde edilen bulgular sunulmuştur.

POMTK'ya Dayalı Olarak Yapılan Güvenirlik İncelemeleri

Araştırmada madde ve yetenek parametrelerine etkisi incelenen faktörlere göre oluşturulan ve POMTK kapsamında ölçeklenen 16 veri setine ilişkin POMTK'ya dayalı olarak güvenilirlik incelemeleri yapılmıştır. Her veri seti için alanyazında önerilen dört ayrı güvenilirlik katsayısı hesaplanmıştır, bunlar: Cronbach Alfa, Lambda2, MS ve LCRC katsayılarıdır. Hesaplanan katsayılardan diğer katsayılara göre güvenirlği daha yansız olarak hesapladığı alan yazında belirlenen MS katsayısı temel alınarak yorumlar yapılmıştır (Sijtsma ve Molenaar, 2002; Van Schuur, 2011). Elde edilen değerler Çizelge 76'da verilmiştir.

Çizelge 76.

Farklı Faktörlere Göre Oluşturulan Veri Setlerinden POMTK Yaklaşımına Göre Kestirilen Güvenirlik Katsayıları

Veri setleri	MS	Alpha	Lambda 2	LCRC	Ortalama
30madde500	0.910	0.909	0.910	0.929	0.914
30madde1000	0.918	0.916	0.917	0.931	0.920
30madde3000	0.911	0.910	0.911	0.922	0.913
5madde500	0.626	0.626	0.632	0.698	0.645
5madde1000	0.662	0.652	0.655	0.676	0.661
5madde3000	0.645	0.639	0.643	0.670	0.649
15madde500	0.815	0.810	0.816	0.850	0.822
15madde1000	0.842	0.839	0.841	0.867	0.847
15madde3000	0.839	0.839	0.838	0.857	0.843
25madde500	0.884	0.881	0.884	0.907	0.889
25madde1000	0.901	0.898	0.900	0.916	0.903
25madde3000	0.840	0.841	0.843	0.866	0.847
A-endüşük	0.805	0.803	0.805	0.832	0.811
A-enyüksek	0.880	0.878	0.879	0.900	0.884
B-enkolay	0.842	0.841	0.843	0.866	0.848
B-enzor	0.864	0.861	0.863	0.886	0.868

Çizelge 76 incelendiğinde yer alan veri setlerinin isimlendirilmesinde çalışmada incelemeye alınan faktörler temel alınmış, madde sayısı ve örneklem büyüklüklerine göre isimlendirmeler yapılmıştır. Çizelgenin sonunda yer alan A-en düşük, en yüksek ve B-en düşük, en yüksek olarak belirtilen veri setleri ise, maddelerin güçlük ve ayırt edicilik parametrelerine göre oluşturulan veri setlerini temsil etmektedir. A-en düşük ve en yüksek, PMTK modellerine göre a parametresi olarak adlandırılan madde ayırt edicilik parametresinin en düşük ve en yüksek olduğu veri setlerini belirtirken, B-en düşük ve en yüksek, yine PMTK kapsamında madde güçlük parametresi olarak belirtilen madde güçlük parametrelerine göre yüksek ve düşük olan maddelerden oluşan veri setini ifade etmektedir. Veri setlerinin tanıtılmasından sonra, elde edilen değerlerin yorumlanmasına geçilebilir.

Çizelge 76'da her veri setine ilişkin hesaplanan dört ayrı güvenilirlik kestirimi görülmektedir. Elde edilen değerler genel olarak incelenecek olursa, dört katsayının da her veri setinde benzer çıktığı görülmektedir. Hesaplanan katsayılar içinde, tüm veri setleri için en yüksek değerleri veren katsayı LRCR katsayısıdır. Diğer üç katsayının ise birbirine yakın değerler aldığı görülmektedir. Elde edilen katsayılar veri setleri düzeyinde incelendiğinde ise, en yüksek güvenilirliğe sahip olan veri setinin 30 maddeden oluşan 1000 kişilik veri seti olduğu bulunmuştur. Ayrıca beklenildiği üzere en yüksek güvenilirlik değerleri 30 maddelik veri setleri için kestirilmişken, en düşük güvenilirlik katsayıları beş maddeden oluşan veri setleri için hesaplanmıştır ve bu değerler sınır olarak kabul edilen 0.70'in altında bulunmuştur. Bu bulgu Şengül Avşar (2015)'in çalışmasında çok kategorili veriler için elde edilen bulgu ile tutarlıdır ve madde sayısındaki artış, PMTK'da olduğu gibi, POMTK kapsamında da testin güvenilirlik düzeyini artırmıştır.

Diğer test uzunluğu koşulu olan 25 ve 15 maddelik veri setleri için hesaplanan güvenilirlik katsayıları incelendiğinde, her iki koşuldan kestirilen katsayıların birbirine benzer olduğu, fakat yine madde sayısının artmasıyla 25 maddelik veri setlerinde güvenilirliğin daha yüksek olduğu görülmektedir. Kestirilen katsayılar örneklem büyüklüğü faktörüne göre incelenecek olursa, tüm madde sayısı koşulları için 1000 kişilik veri setlerinden hesaplanan güvenilirlik katsayılarının 500 ve 3000 kişilik veri setlerinden kestirilen veri setlerine göre daha yüksek olduğu bulunmuştur. Bu durum model veri uyumu incelemelerinden elde edilen bulgularla paralellik göstermektedir. POMTK'ya yönelik olarak yapılan model veri uyumu incelemelerinde en iyi uyumun 1000 kişilik veri

setlerinden sağlandığı bulunmuştur ve bu veri setlerine yönelik güvenilirliğin de daha yüksek olduğu bu aşamada belirlenmiştir. Dolayısıyla POMTK'ya dayalı analizlerde veri setinin 1000 kişiden oluşması ideal durum olarak kabul edilebilir.

Çalışma kapsamında incelenen bir başka faktör olan maddelerin psikometrik özelliklerine göre oluşturulan veri setlerine ilişkin kestirilen güvenilirlik katsayıları incelendiğinde ise, madde ayırt edicilik indeksleri yüksek olan maddelerden oluşan veri setinin güvenilirliğinin, madde ayırt edicilik parametreleri düşük olan maddelerden oluşan veri setine göre daha yüksek olduğu bulunmuştur. Madde güçlük parametresine göre düşük ve yüksek maddelerden oluşturulan veri setlerinden kestirilen güvenilirlik katsayılarının ayırt edicilik parametrelerine göre oluşturulmuş veri setlerinden kestirilen güvenilirlik katsayılarından daha yüksek olduğu bulunmuştur. Değerler birbirine yakın olsa da, madde güçlüğü yüksek olan maddelerden oluşturulan veri setlerinden kestirilen güvenilirlik katsayıları, kolay maddelerden oluşan veri setinden kestirilen katsayılara göre daha yüksektir. POMTK ile kestirilen güvenilirlik katsayılarını genel olarak özetlemek gerekirse şu sonuçlara ulaşılabılır:

- Madde sayısının artması güvenilirliği artırmıştır.
- En yüksek güvenilirlik katsayıları 30 maddelik veri setlerinden kestirilmiştir.
- En düşük güvenilirlik katsayıları ise beş maddelik veri setlerinden kestirilmiştir.
- Tüm koşullarda en yüksek güvenilirlik katsayıları 1000 kişilik veri setlerinden kestirilmiştir.
- Madde ayırt edicilik ve güçlük parametrelerinin değişmesi güvenilirlik katsayılarında büyük değişimlere neden olmamıştır.
- Hesaplanan güvenilirlik katsayıları genel olarak birbirine benzer değerler üretse de, en yüksek değer LCRC'den elde edilmiştir. Lambda 2'de diğer katsayılara göre daha yüksektir. Cronbach Alfa ve MS'den kestirilen değerler birbirine daha yakındır. MS'nin daha yansız kestirimler yaptığı POMTK alan yazınında kanıtlanmış olduğu için, LCRC ve Lambda 2 kestirimlerinin gerçek güvenilirliğin üzerinde değerler ortaya çıkardığı sonucuna ulaşılmıştır.

Yukarıda özetlenen bulgulardan yola çıkarak, POMTK ile yüksek düzeyde güvenilirliğin sağlanması için, madde sayısının 15'in üzerinde, örneklem büyüklüğünün ise 1000'e yakın olması gerektiği sonucuna ulaşılmıştır. Beş maddelik veri setlerinin POMTK

kapsamında tutarlı sonuçlar vermediđi, 15 maddenin tüm örneklerde orta düzeyde (0.80'lerde) güvenilirliğe sahip olduđu, ancak 25 madde ve 1000 kişiden oluşan veri setinin güvenilirliğinin 0.90'ın üzerinde olduđu belirlenmiştir. Ayrıca madde sayısı 30 olduğunda, 500 kişilik veri setinde dahi, güvenilirliđin yüksek olduđu bulunmuştur.



BÖLÜM 5

SONUÇ VE ÖNERİLER

Bu bölümde, araştırmadan elde edilen sonuçlar ve bu sonuçlara dayalı olarak sunulan önerilere yer verilmiştir.

Sonuçlar

Bu çalışmada, örneklem büyüklüğü, test uzunluğu ve maddelerin güçlük ve ayırt edicilik parametrelerine göre farklı koşullara sahip veri setlerinden parametrik ve parametrik olmayan madde tepki kuramı uyumlu modeller belirlenerek kestirilen madde ve yetenek parametrelerinin karşılaştırılması amaçlanmıştır. Bu amaç doğrultusunda yapılan analizlerden elde edilen sonuçlar aşağıda maddeler halinde özetlenmiştir.

1. Örneklem büyüklüğü faktörüne ilişkin koşullara göre oluşturulan 500, 1000 ve 3000 kişilik veri setlerinden uyumlu PMTK modellerine göre kestirilen madde parametreleri arasında yüksek ve manidar ilişkiler bulunmuştur. Madde güçlük parametreleri arasındaki ilişkiler madde ayırt edicilik parametreleri arasında hesaplanan ilişkilerden daha yüksek bulunmuştur ve bu bulguya dayanarak madde ayırt edicilik parametrelerinin örneklem büyüklüğünden daha fazla etkilendiği yorumuna gidilmiştir. Ayrıca PMTK modelleri ile kestirilen parametreler arasındaki ilişkinin örneklem büyüklüğü ile orantılı bir şekilde arttığı sonucuna ulaşılmıştır.
2. POMTK kapsamında farklı örneklem büyüklüğü koşullarına göre incelenen veri setlerinden kestirilen madde parametreleri arasındaki uyumun yüksek düzeyde ve manidar olduğu sonucuna ulaşılmıştır. Bu bulgudan yola çıkarak POMTK yaklaşımının küçük örneklemelerde bile etkili madde parametre kestirimleri sağladığı sonucunavarılmıştır.

Örneklem büyüklüğüne göre PMTK ve POMTK'dan kestirilen madde parametreleri arasındaki ilişkiler incelendiğinde ise, tüm örneklemelerden kestirilen madde ayırt edicilik parametreleri arasında yüksek ve manidar ilişkiler olduğu bulunmuştur. Madde güçlük parametreleri arasındaki ilişkilerin ise ayırt edicilik parametreleri arasındaki ilişkilere göre daha yüksek olduğu bulunmuştur.

3. Çalışma kapsamında madde parametreleri kestirimlerine etkisi incelenen bir diğer faktör test uzunluğudur ve bu faktörün koşulları ile örneklem büyüklüğünün koşulları çaprazlanarak incelenmeye alınmıştır. Test uzunluğu faktörü kapsamında beş, 15 ve 25 maddelik test uzunlukları olarak üç ayrı koşul incelenmiş ve her iki faktöre ilişkin koşulların birlikte ele alınması ile dokuz veri seti oluşturulmuştur. PMTK kapsamında yapılan model veri uyumu incelemesinde, beş maddelik veri setleri için madde düzeyinde yapılan model veri uyumu incelemesinde uyumlu madde olmadığı belirlenmiştir. On beş maddelik veri setlerinde madde düzeyinde model veri uyumunun iyileştiği ve madde düzeyinde model veri uyumunun en yüksek düzeyde sağlandığı veri setlerinin 25 maddelik veri setleri olduğu belirlenmiştir.
4. Test uzunluğu koşulları, örneklem büyüklüğü koşulları ile çaprazlanarak oluşturulan veri setleri için POMTK analizleri yapılmış ve veri setleri için değişmez madde sıralaması varsayımının karşılanmadığı, dolayısıyla bu veri setlerinin POTMK kapsamında yer alan ilk model olan Monoton Homojenlik Modeline uyumlu olduğu belirlenmiştir.
5. Farklı test uzunluğu ve örneklem büyüklüğü koşullarından oluşan veri setlerinden PMTK ve POMTK'ya göre kestirilen madde parametreleri arasındaki ilişkilerin yüksek ve manidar olduğu bulunmuştur. İlişkilerin düzeyleri örneklemelere göre incelendiğinde, en yüksek ilişkilerin tüm test uzunluğu koşullarında, 500 ve 1000 kişilik veri setlerinde olduğu belirlenmiş ve madde ayırt edicilik parametreleri arasındaki ilişkilerin madde güçlük parametrelerine oranla daha yüksek olduğu belirlenmiştir. Tüm veri setlerinde yüksek ve manidar ilişkiler gözlenmiş, madde

sayısı arttıkça ilişkilerin düzeyi yükselmiş ve en yüksek ilişkiler aynı örneklem büyüklükleri için hesaplanmıştır.

6. Araştırmada yanıt aranan üçüncü soru kapsamında değişen madde psikometrik niteliklerinin birey yetenek kestirimlerine etkisi incelenmiştir. Veri setleri oluşturulurken araştırmanın ilk sorusunda kestirilen madde parametreleri kullanılmış ve her iki kuramdan elde edilen parametre değerleri sıralanarak incelenmiş ve sıralamaların benzer olduğu belirlenmiştir. Her iki yaklaşımda da ortak olarak yer alan madde güçlükleri yüksek ve düşük, madde ayırt edicilikleri yüksek ve düşük olan maddeler bir araya getirilerek dört ayrı veri seti oluşturulmuş ve bu veri setlerine uyum sağlayan PMTK ve POMTK modelleri ile birey yetenekleri kestirilmiştir. Madde güçlükleri ve ayırt edicilikleri yüksek ve düşük olan veri setlerinden kestirilen yetenekler arasındaki ilişkiler incelendiğinde, kestirilen yetenekler arasında manidar ve yüksek düzeyde ilişkiler olduğu belirlenmiştir. Dolayısıyla yapılacak ölçme uygulamalarında bireylerin yetenek düzeyleri bakımından sıralanması yeterli ise, POMTK kapsamında ölçekleme yapılabileceği sonucuna ulaşılmıştır.
7. Çalışma kapsamında farklı örneklem büyüklüğü ve test uzunluğu koşullarından oluşan veri setlerinden PMTK ve POMTK modellerine göre kestirilen yetenek parametrelerinin karşılaştırılması amaçlanmıştır. Beş, 15 ve 25 maddelik 500,1000 ve 3000 kişilik veri setleri için yetenek kestirimleri uygun PMTK ve POMTK modellerine göre yapılmıştır. Beş maddelik farklı örneklem büyüklüklerinden oluşan veri setleri için PMTK kapsamında kestirimler arasındaki ilişkilerin sıfıra yakın değerler aldığı ve manidar olmadığı bulunmuştur. PMTK ve POMTK'dan yapılan kestirimler incelendiğinde ise, POMTK ile tüm veri setlerinden kestirilen yetenekler ile PMTK kapsamında yalnızca 3000 kişilik veri setinden kestirilen yetenekler arasında yüksek ve manidar ilişkiler olduğu belirlenmiştir. On beş maddelik veri setlerinden kestirilen parametreler arasındaki ilişkiler incelendiğinde ise, PMTK ile kestirilen yetenekler parametreleri arasında manidar bir ilişki bulunmazken, POMTK ile kestirilen tüm yetenekler arasında manidar ve yüksek

ilişkiler olduğu belirlenmiştir. PMTK ve POMTK'ya göre yapılan yetenek kestirimlerinden yalnızca PMTK'ya dayalı 3000 kişilik veri setinden kestirilen parametreler ile POMTK'ya dayalı tüm örneklemelerden kestirilen yetenek parametreleri arasında yüksek ve manidar ilişkiler olduğu belirlenmiştir. Yirmi beş maddelik veri setlerinden kestirilen yetenekler incelendiğinde ise, PMTK modelleri ile kestirilen yetenekler arasında manidar ve yüksek ilişkiler olduğu bulunmuştur. Özetle tüm örneklem büyüklükleri ve madde sayıları için POMTK'dan tutarlı yetenek kestirimleri elde edilirken, PMTK ile yalnızca madde sayısının 25 ya da örneklem büyüklüğünün 3000 olduğu durumlarda tutarlı yetenek kestirimleri elde edilmiştir.

8. Çalışma kapsamında cevap aranan son soruya ilişkin oluşturulan tüm veri setleri için güvenilirlik incelemesi yapılmıştır. PMTK için güvenilirlik kestirimleri test bilgi fonksiyonları ile incelenirken, POMTK kapsamında alan yazında önerilen katsayılar kullanılmıştır. PMTK kapsamında yapılan güvenilirlik incelemelerinde beş maddelik veri setlerinden elde edilen test bilgi miktarının oldukça düşük olduğu belirlenmiş ancak test uzunluğunun 15'ten 25'e çıkması testin verdiği maksimum bilgi düzeyini büyük oradan arttırmaktayken, 25'ten 30'a çıkması testin verdiği bilgi düzeyini düşük düzeyde değiştirdiği bulunmuştur. Madde sayısının artması PMTK kapsamında güvenilirliği artırırken, örneklem büyüklüğündeki artışın güvenilirlik üzerinde çok etkili olmadığı belirlenmiştir. POMTK ile analiz edilen verilerde en yüksek güvenilirlik 1000 kişilik veri setlerinden elde edilmiş ve madde sayısının artışının PMTK'da olduğu gibi güvenilirlik katsayılarında olumlu yönde değişime neden olduğu belirlenmiştir.

Öneriler

1. Araştırmada elde edilen bulgulara dayalı olarak, testte yer alan madde sayısının 25'ten az ya da veri setinin geniş olmadığı durumlarda, bireylere ilişkin yetenek kestirimleri sıralama amaçlı kullanılacaksa, POMTK modelleri tercih edilebilir.

2. Arařtırmacılar tarafından yalnızca maddelerin ölçeklenmesi isteniyorsa, PMTK modelleriyle incelenen tüm örneklem büyüklüğü koşullarında ve test uzunluğu koşullarında PMTK modelleriyle yüksek ve manidar ilişkiler gösterdiği çalışma kapsamında belirlenen POMTK modelleri kullanılabilir.
3. Psikolojik ölçmeler için deęişen madde fonksiyonu belirleme, gelişimi aşamalı olarak takip etme gibi birçok önemli işleve sahip olan deęişmez madde sıralaması varsayımı PMTK yaklaşımı ile yapılan arařtırmalarda incelenebilir. Böylece maddelerin her yetenek düzeyi için nasıl sıralandığı belirlenebilir.
4. Çalışma kapsamında PMTK yaklaşımı içinde yer alan bir, iki ve üç parameteli lojistik modellerle POMTK kapsamında yer alan monoton homojenlik ve çift monotonluk modellerinden kestirilen parametreler kıyaslanmıştır. Farklı çalışmalarda çift monotonluk modeli ile benzerlik gösteren Rasch model incelemeye alınabilir.
5. Dünya çapında olduğu gibi, Türkiye’de de özellikle saęlık bilimlerinde ve psikolojide kullanılan ölçekler için POMTK ile ölçeklemeler yapılabilir. Bu sayede söz konusu psikolojik özelliklerin doğasına ilişkin PMTK modellerinden elde edilemeyen ölçülen psikolojik özellięi tanımlar nitelikte ayrıntılı bilgiler elde edilebilir.
6. Çalışma kapsamında incelenen maddelerin psikometrik nitelikleri faktörü altında madde güçlük ve ayırt edicilik parametreleri temel alınarak kestirimler yapılmıştır. Farklı çalışmalarda şans parametresinin PMTK ve POMTK kestirimlerine etkisi incelenebilir.
7. Çalışma kapsamında yalnızca ikili puanlanan maddeler için PMTK ve POMTK modelleri ile yapılan madde ve yetenek parametre kestirimleri kıyaslanmıştır. Arařtırmacılar tarafından karma testler için söz konusu yaklaşımlardan elde edilen parametreler kıyaslanabilir.

8. PMTK kapsamında model veri uyumunun zayıf kaldığı durumlarda POMTK yaklaşımındaki tek boyutluluk ve monotonluk varsayımlarını test eden yöntemler kullanılarak veri yapısı hakkında açıklayıcı ve doğrulayıcı faktör analizinden elde edilen bilgilerden daha detaylı bilgiler elde edilebilir.



KAYNAKÇA

- Avşar, Şengül. A. (2015). *Çok kategorili puanlanan maddelerin psikometrik özelliklerinin farklı test koşullarında parametrik olmayan Madde Tepki Kuramı modellerine göre incelenmesi*. Yayınlanmamış doktora tezi. Ankara Üniversitesi Eğitim Bilimleri Fakültesi.
- Baker, F. B. and Kim. S.H. (2004). *Item response theory: Parameter estimation techniques* (2nd Edition). New York: Marcel Dekker.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. Lord, F.M., and Novick, M.R. (Ed.). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Büyüköztürk,Ş., Çokluk, Ö. ve Köklü, N. (2012). *Sosyal Bilimler İçin İstatistik*. Ankara, Pegem Akademi.
- Christensen, K. and Kreiner, S. (2010). Monte Carlo tests of the Rasch model based on scalability coefficients. *British Journal of Mathematical and Statistical Psychology*, 63, 101-111.
- Crocker, L., and Algina, J. (1986). *Introduction classical and modern test theory*. New York: Harcourt Brace Javonovich College Publishers.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 6, 297-334.
- Croon, M.A. (1991). Investigating Mokken scalability of dichotomous items by means of ordinal latent class analysis. *British Journal of Mathematical and Statistical Psychology*. 44 (2).315-331.
- De Ayala, R.J. (2009). *The Theory and Practice of Item Response Theory*. New York: The Guilford Press.
- DeMars, C. (2010). *Item response theory*. Oxford University Press.
- Douglas, J. (1997). Joint consistency of nonparametric item characteristic curve and ability estimation. *Psychometrika*, 62, 7-28.

- Douglas, J. and Cohen, A. (2001). Nonparametric item response function estimation for assessing parametric model fit. *Applied Psychological Measurement*, 25 (3), 234-243.
- Dyehouse, M. A. (2009). A comparison of model-data fit for parametric and nonparametric item response theory models using ordinal level ratings. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database.
- Embretson, S. E. and Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Feldt, L. S. & Qualls, A. L. (1996). Bias in coefficient alpha arising from heterogeneity of test content. *Applied Measurement in Education*, 9 (3), 277-286.
- Field, A. (2009). *Discovering statistics using SPSS (3rd edition)*. Dubai: Sage.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.
- Hambleton, R. K. and Swaminathan, H. (1985). *Item response theory. Principles and applications*. Boston: Kluwer.
- Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hair, J., Anderson, R., Babin, B., Black, W., and Tahtam R. (2006). *Multivariate data analysis*. New Jersey, Prentice Hall, Inc.
- Jöreskog, K. G., and Sörbom, D. (2001). *Lisrel 8: User's reference guide*. Chicago, IL: Scientific Software International, Inc.
- Junker, B. (2000). Some topics in nonparametric and parametric IRT. with some thoughts about the future. Carnegie Mellon University: Department of Statistics.
- Junker, B. W. and Sijtsma, K. (2001). Nonparametric item response theory in action: an overview of the special issue. *Applied Psychological Measurement*, 25, 211-220.
- Kelecioğlu, H. (2001). Örtük özellikler teorisindeki b ve a parametreleri ile klasik test teorisindeki p ve r istatistikleri arasındaki ilişki. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 20, 104-110.
- Kline, R.B. (2000). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Karasar, N. (2010). Bilimsel Arastırma Yöntemi-Kavramlar, İlkeler, Teknikler. 21. Baskı, Ankara: Nobel. Yayınevi.

- Koğar, H. (2014). *Madde tepki kuramının farklı uygulamalarından elde edilen parametrelerin ve model uyumlarının örneklem büyüklüğü ve test uzunluğu açısından karşılaştırılması*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü.
- Kujipers, R. E., Van der Ark, L. A. , Croon, M. A. and Sijtsma, K. (2016). Bias in point estimates and standard errors of Mokken's scalability coefficients. *Applied Psychological Measurement*. 40 (5), 331-345.
- Lee, Y. S. (2007). A comparison of methods for nonparametric estimation of item characteristic curves for binary items. *Applied Psychological Measurement*. 31. 121-134.
- Liang, T. (2010). An assessment of the nonparametric approach for evaluating the fit of item response models. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database.
- Ligtvoet, R., Van der Ark, L., The Marvelde, J.M. and Sijtsma, K. (2010). Investigating invariant item ordering for polythomously scored items. *Educational and Psychological Measurement*, 70 (4), 578-595.
- Loevinger, J. (1947). A systematic approach to the construction and evaluation of tests of ability. *Psychological Monographs*. 61 (A).
- Loevinger, J. (1948). The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychological Bulletin*. 45 (6). 507-529.
- Lord, F. M. (1970). Item characteristic curves estimated without knowledge of their mathematical form—A confrontation of Birnbaum's logistic model. *Psychometrika*, 55 (1), 43-50.
- Lord, F. (1980). *Applications of item response theory to practical testing problems*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Lord, F. M., and Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison- Wesley.
- Meijer, R. R., Sijtsma, K and Smidt, N. G. (1990). Theoretical and empirical comparison of the Mokken and the Rasch approach to IRT. *Applied Psychological Measurement*. 14. 283–298.
- Meijer, R. R., Sijtsma, K., and Molenaar, I. W. (1995). Reliability estimation for single dichotomous items based on Mokken's IRT model. *Applied Psychological Measurement*, 19 (4), 323-335.
- Meijer, R.R., and Baneke, J.J. (2004). Analyzing psychopathology items: a case for nonparametric item response theory modeling. *Psychological Methods*, 9 (3), 354-368.

- Meijer, R. R., and Egberink, I. J. L. (2012). *Investigating Invariant Item Ordering in Personality and Clinical Scales: Some Empirical Findings and a Discussion. Educational and Psychological Measurement*, 72 (4), 589-607. DOI: [10.1177/0013164411429344](https://doi.org/10.1177/0013164411429344)
- Meijer, R. R., Egberink, I.J.L., Emons, W.H.M. and Sijtsma, K. (2008). Detection and validation of unscalable item score patterns using item response theory: an illustration with Harter's self perception profile for children. *Journal of Personality Assessment*, 90 (3), 227-238.
- Meijer, R., Tendeiro, J., and Wanders, R. (2014). The use of nonparametric item response theory to explore data quality. S. P. Reise & D. Revicki (Ed.). *Handbook of item response theory modeling: Applications to typical performance assessment*. In S. P. Reise & D. Revicki (Eds.). *Handbook of item response theory modeling: Applications to typical performance assessment*. (s. 85-110). New York: Routledge.
- Mislevy, R. J. & Stocking, M. L. (1989). A consumer's guide to LOJIST and BILOG. *Applied Psychological Measurement*. 13. 57-75.
- Molenaar, I. W. (1982). Mokken scaling revised. *Kwantitatieve Methoden*. 3. 145-164.
- Molenaar, I.W. ve Sijtsma. K. (1984). Internal consistency and reliability in Mokken's nonparametric item response model. *Tijdschrift voor onderwijsresearch*. 9 (5). 257-268.
- Molenaar, I. W. (1991). A weighted Loevinger H coefficient extending Mokken scaling to multicategory items. *Kwantitatieve Methoden*. 37. 97- 117.
- Molenaar. I. W. (1997). Nonparametric models for polytomous responses. In W.J. Man der Linden and R.K. Hambleton (Ed.). *Handbook of modern item response theory* (s. 369-380). New York: Springer.
- Molenaar, I. W. (2001). Thirty years of nonparametric item response theory. *Applied Psychological Measurement*, 25 (3), 295-299.
- Mokken, R. J. (1971). *A theory and procedure of scale analysis with applications in political research*. New York. Berlin: Walter de Gruyter. Mouton.
- Mokken. R. J. and Lewis. C. (1982). A nonparametric approach to the analysis of dichotomous item responses. *Applied Psychological Measurement*, 6 (A), 417-430.
- Mokken, R. J. (1997). Nonparametric models for dichotomous responses. In W. J. Man der Linden and R. K. Hambleton (Ed.). *Handbook of Modern Item Response Theory* (s. 351-367). New York: Springer.
- Muncer, S.J. and Speak, B. (2016). Mokken scale analysis and confirmatory factor analysis of the health of the nation outcome scales. *Personality and Individual Differences*, 94, 272-276.

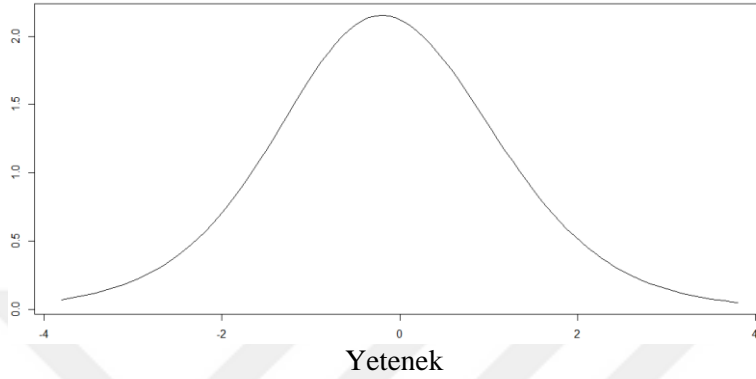
- Ramsay, J. O. and Abrahamowicz M. (1989). Binomial regression with monotone splines: A psychometric application. *Journal of the American Statistical Association*, 54 (408), 906-915
- Ramsay, J. O. (1991). Kernel smoothing approaches to nonparametric item characteristic curve estimation. *Psychometrika*, 56 (4), 611-630.
- Ramsay, J. O. (1997). A functional approach to modeling test data. In W. J. Man der Linden and R. K. Hambleton (Ed.). *Handbook of Modern Item Response Theory* (s. 381-394). New York: Springer.
- Ramsay, J. O. (2000). TestGraf A computer program for nonparametric analysis of testing data. Unpublished manuscript. McGill University.
- Ramsay, J. O. and Winsberg, S. (1991). Maximum marginal likelihood estimation for semiparametric item analysis. *Psychometrika*, 56 (3), 365-379.
- Rossi, N., Wang, X. and Ramsay, J. O. (2002). Nonparametric item response function estimates with the EM algorithm. *Journal of Educational and Behavioral Statistics*, 27, 291-317.
- Roussos, L. A., and Ozbek, O. Y. (2006). Formulation of the DETECT population parameter and evaluation of DETECT estimator bias. *Journal of Educational Measurement*, 43 (3), 215-243.
- Santor, D. and Ramsay, J. O. (1998). Progress in the technology of measurement: Applications of item response models. *Psychological Assessment*, 10, 345-359.
- Schermelleh-Engel, K. and Moosbrugger, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8 (2), 23-74.
- Sijtsma, K. (1998). Methodology review: Nonparametric IRT approaches to the analysis of dichotomous item scores. *Applied Psychological Measurement*, 22 (1), 3-31.
- Sijtsma, K., and Molenaar, I. W. (2002). *Introduction to nonparametric item response theory*. Newbury Park, CA: Sage.
- Sijtsma, K. and Junker, B.W. (1996). A survey of theory and methods of invariant item ordering. *British Journal of Mathematical and Statistical Psychology*, 49, 79-105.
- Sijtsma, K. and Junker, B.W. (2006). Item Response Theory: Past Performance. Present Developments and Future Expectations. *Behaviormetrika*, 1, 75-102.
- Sijtsma, K. and Meijer, R. R. (2007). Nonparametric item response theory. C. R. Rao. & S. Sinharay (Ed.) *Handbook of statistics 26: psychometrics*. (s.719-746). Amsterdam: Elsevier.
- Sijtsma, K., Meijer, R.,R. and Van der Ark, L.A. (2011). Mokken scales analysis as time goes by: an update for scaling practitioners. *Personality and Individual Differences*, 50, (31-37).

- Stocking, M. L. (1997). Item response theory. In J. P. Keeves (Ed.) *Educational Research. Methodology. and Measurement: an International Handbook*. Cambridge: Elsevier Services Ltd. 836-840.
- Stout, W. F. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52, 589–617.
- Stout, W. F. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*, 55, 293–325.
- Stout, W., Habing, B., Douglas, J., Kim, H. R., Roussos, L., & Zhang, J. (1996). Conditional covariance-based nonparametric multidimensionality assessment. *Applied Psychological Measurement*, 20 (4), 331-354.
- Stout, W. (2002). Psychometrics: from practice to theory and back. *Psychometrika*, 67 (4), 485-518.
- Straat, J. H., Van der Ark, L.A. and Sijtsma, K. (2013). Comparng optimization algorithmsfor item selection in Mokken scale analysis. *Journal of Classification*, 30, 75-99.
- Straat, J. H., Van der Ark, L.A. and Sijtsma, K. (2014). Minimum sample size requirements for Mokken scale analysis. *Educational and Psychological Measurement*, 74, 809-822.
- Suppes, P. and Zanotti, M. (1981). When are probabilistic explanations possible? *Synthese*, 48, 191-199.
- Tabachnick, B.G., and Fidell L.S. (2007). *Using multivariate statistics*. USA, Pearson Education.
- Thissen, D., Steinberg, L. & Mooney, J. (1989). Trace lines for testlets: A use of multiple-categorical response models. *Journal of Educational Measurement*, 26, 247-260.
- van der Ark, L. A. , Croon, M.A. and Sijtsma, K. (2008). Mokken scale analysis for dichotomus items using marginal models. *Psychometrika*, 73, 183-208.
- van der Ark, L. A., van der Palm, D. W., & Sijtsma, K. (2011). A latent class approach to estimating test-score reliability. *Applied Psychological Measurement*, 35 (5), 380-392.
- van der Ark, L. A. (2012). New developments in Mokken scale analysis in R. *Journal of Statistical Software*, 48 (5), 1-27.
- van der Linden, W. J. and Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York: Springer-Merlag.

- Van Schuur, W. H. (2003). Mokken scale analysis between the Guttman scale and parametric item response theory. *Political Analysis*, 11, 139-163.
- Van Schuur, W. H. (2011). *Ordinal item response theory: Mokken scale analysis*. Los Angeles: Sage Publications.
- Watson, R., Dreary, I. And Shipley, b. (2008). A hierarchy of distress: Mokken scaling of the GHQ-30. *Psychological Medicine*, 38 (4), 575-579.
- Wismeijer, A.A.J., Sijtsma, K, Van Assen, M.A.L.M. and Vingerhoets, J.J.M. (2007). A comparative study of the dimensionality of the self concealment scale using principal component analysis and Mokken scale analysis. *Journal of Personality Assessment*, 90 (4), 323-334.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three parameter logistic model. *Applied Psychological Measurement*. 8 (2), 125-145.
- Zhao, Y. (2008). *Approaches for addressing the fit of item response theory models to educational test data*. (Doctoral dissertation) Available from ProQuest Dissertations and Theses database.

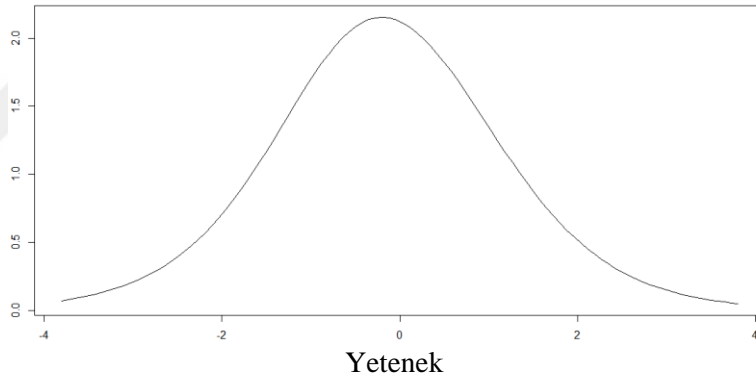
EK-1

Test Bilgi Fonksiyonu



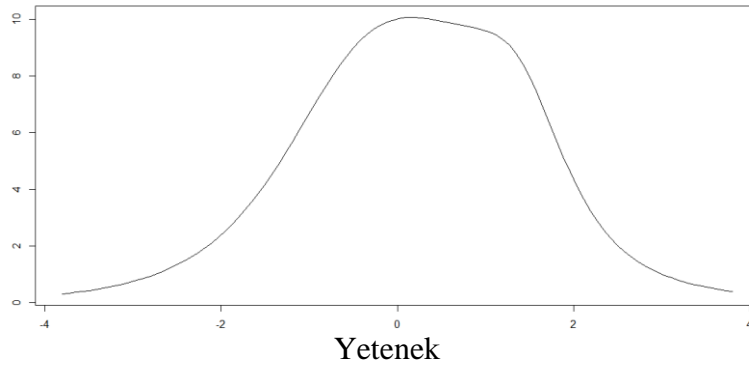
Şekil 1. Beş Maddeden ve 500 Kişiden oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu

Test Bilgi Fonksiyonu

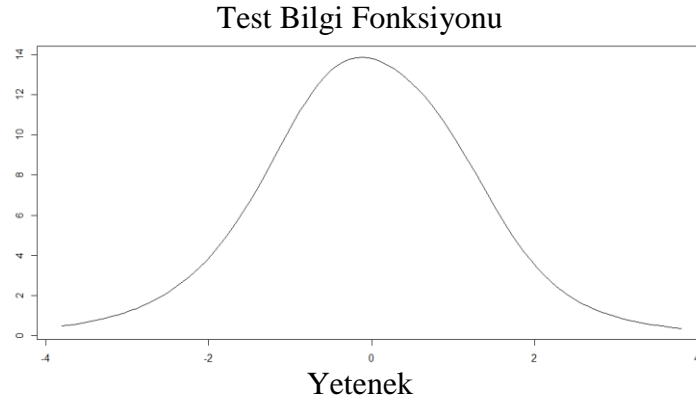


Şekil 2. On Beş Maddeden ve 500 Kişiden oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu

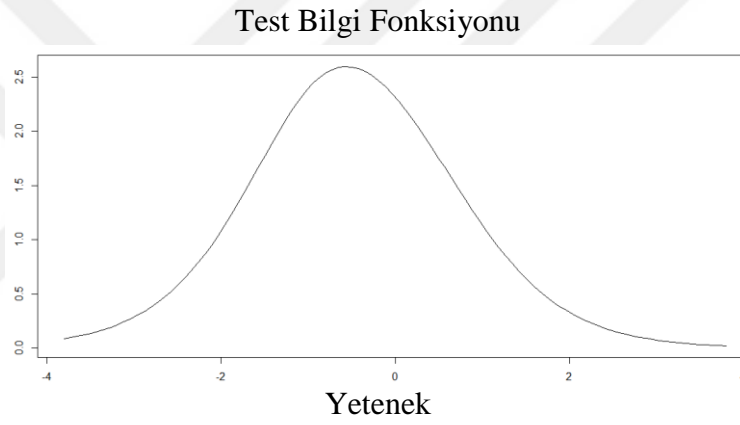
Test Bilgi Fonksiyonu



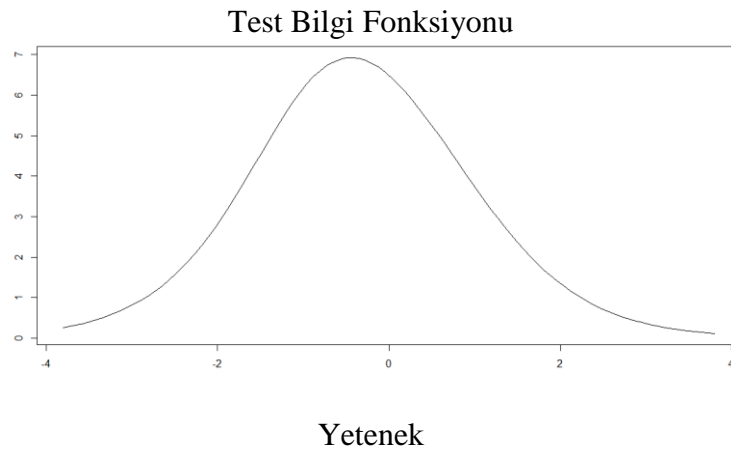
Şekil 3. Yirmi Beş Maddeden ve 500 Kişiden oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



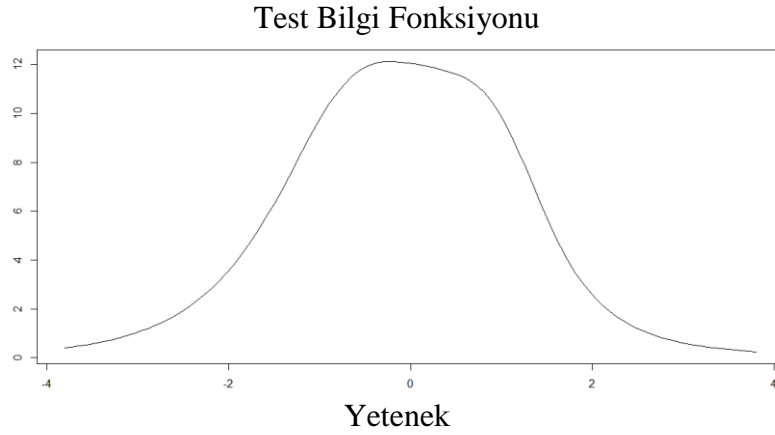
Şekil 4. Otuz Maddeden ve 500 Kişiden Oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



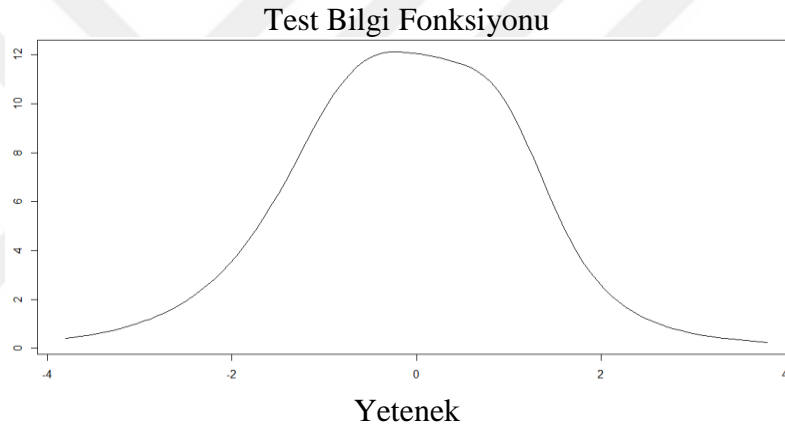
Şekil 5. Beş Maddeden ve 1000 Kişiden oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



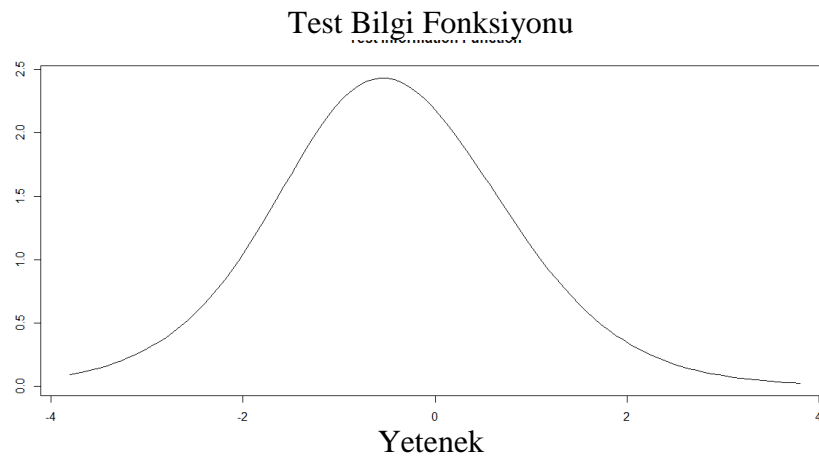
Şekil 6. On Beş Maddeden ve 1000 Kişiden Oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



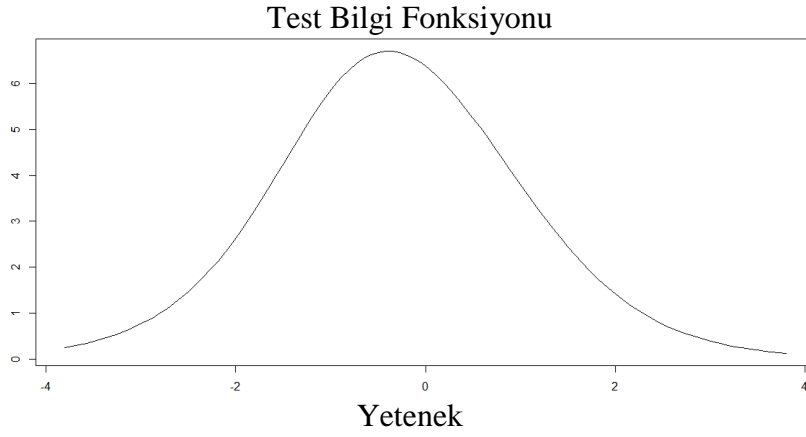
Şekil 7. Yirmi Beş Maddeden ve 1000 Kişiden oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



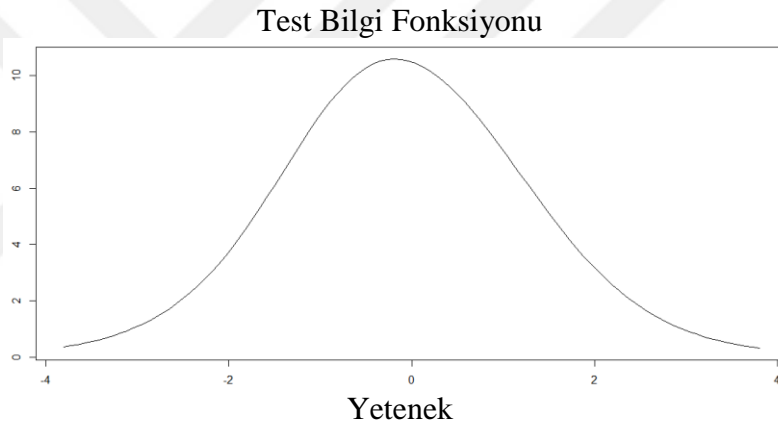
Şekil 8. Otuz Maddeden ve 1000 Kişiden oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



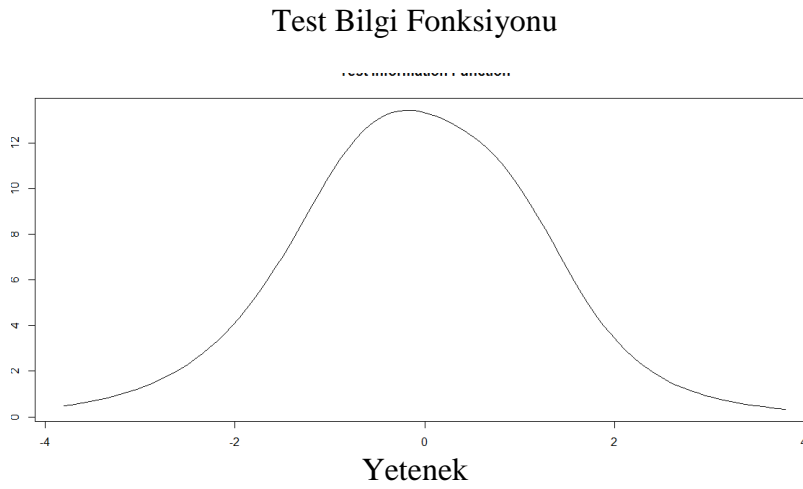
Şekil 9. Beş Maddeden ve 3000 Kişiden oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



Şekil 10. On Beş Maddeden ve 3000 Kişiden Oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu

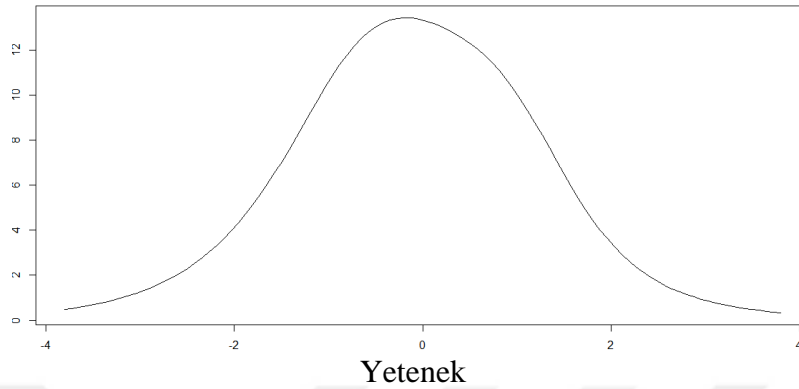


Şekil 11. Yirmi Beş Maddeden ve 3000 Kişiden Oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu



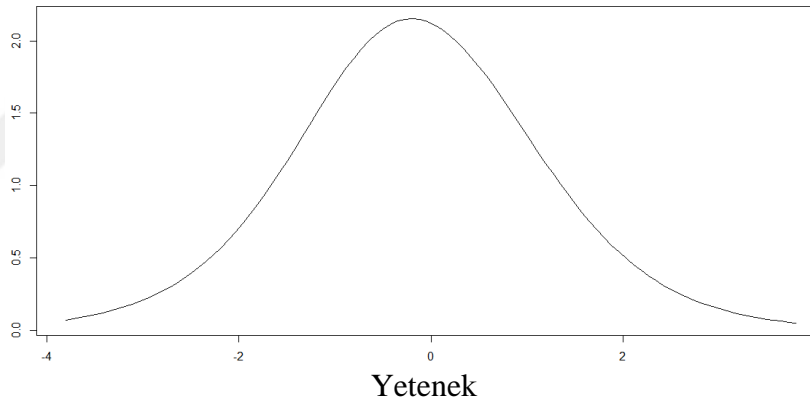
Şekil 12. Yirmi Beş Maddeden ve 3000 Kişiden Oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu

Test Bilgi Fonksiyonu



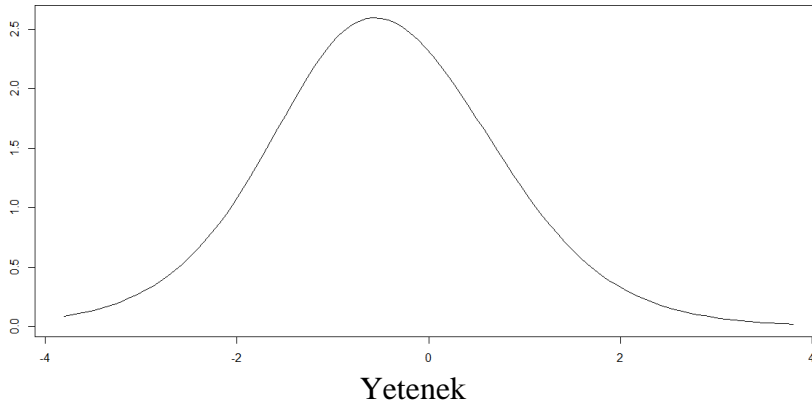
Şekil 13. Otuz Maddeden ve 3000 Kişiden Oluşan Veri Setinden Kestirilen Test Bilgi Fonksiyonu

Test Bilgi Fonksiyonu

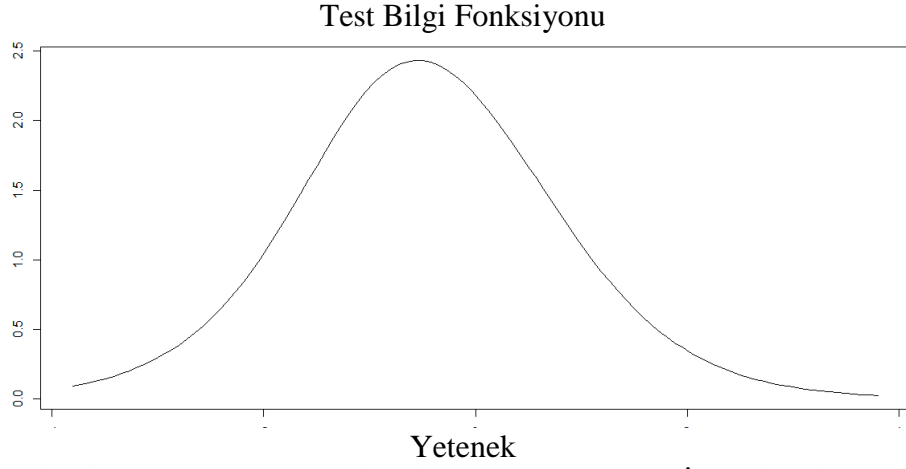


Şekil 14. Beş Maddeden ve 500 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu

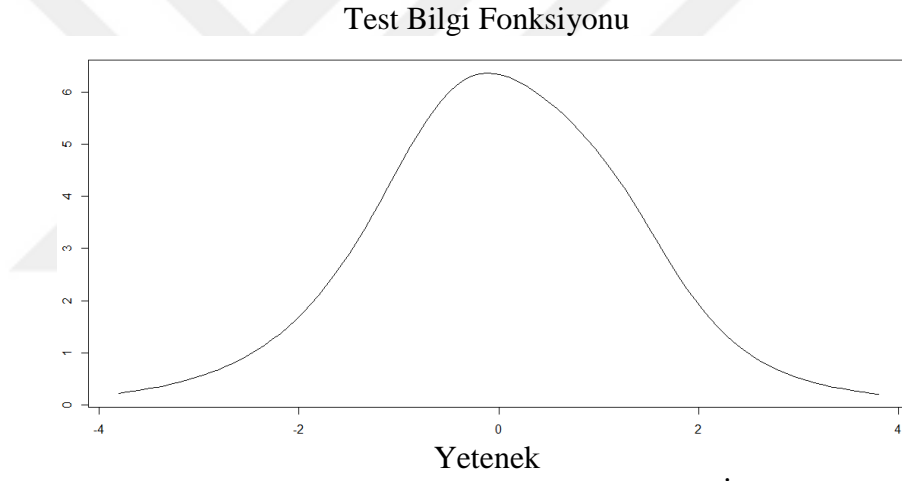
Test Bilgi Fonksiyonu



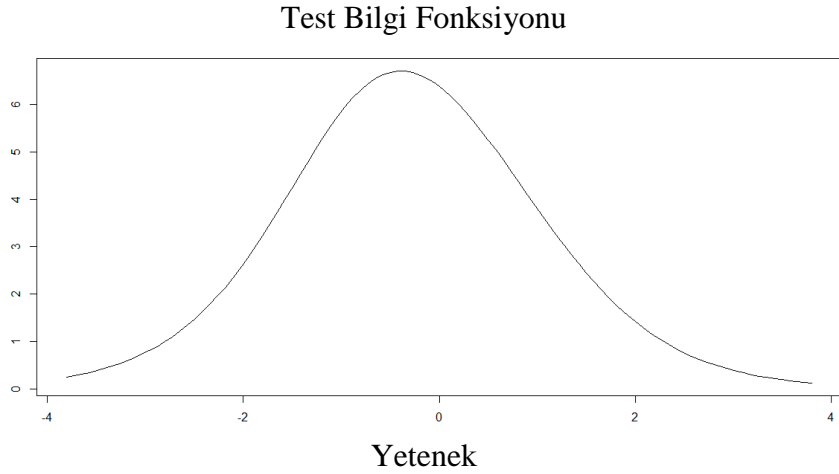
Şekil 14. Beş Maddeden ve 1000 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu



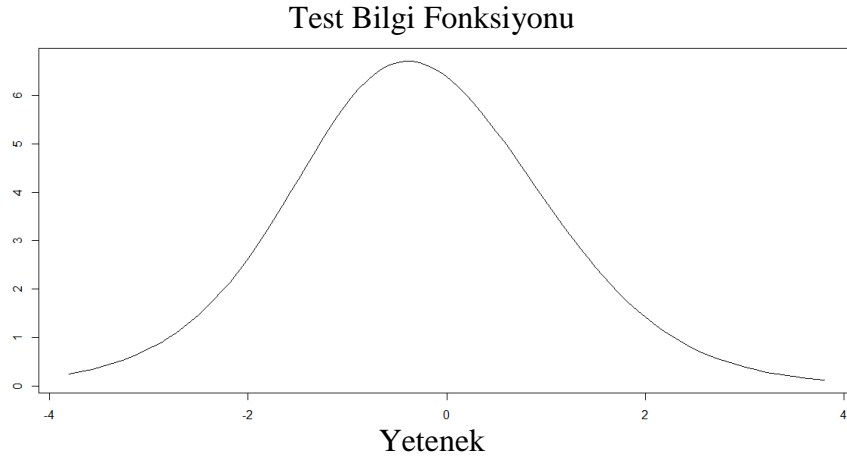
Şekil 14. Beş Maddeden ve 3000 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu



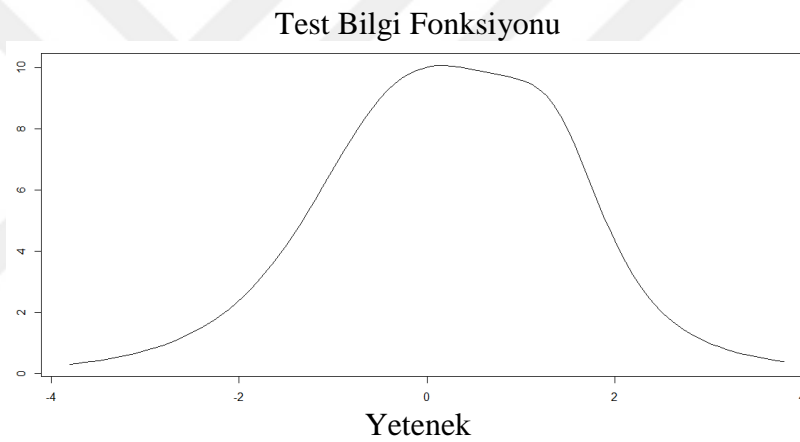
Şekil 15. On Beş Maddeden ve 500 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu



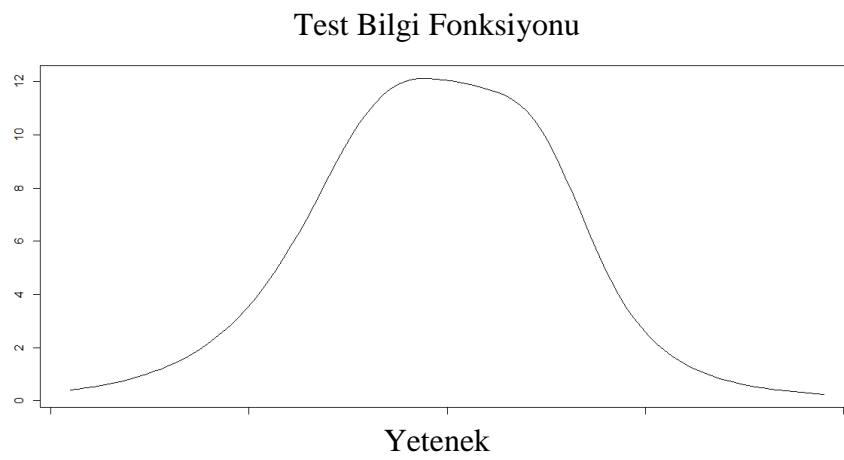
Şekil 16. On Beş Maddeden ve 1000 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu



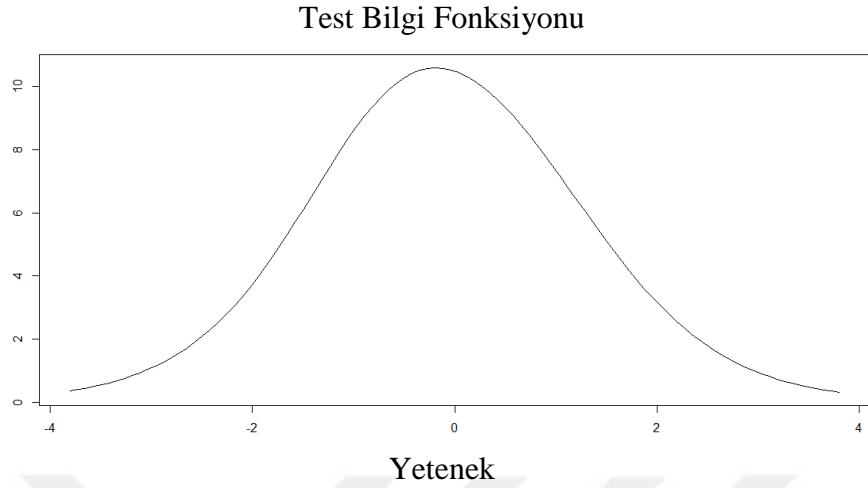
Şekil 17. On Beş Maddeden ve 3000 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu



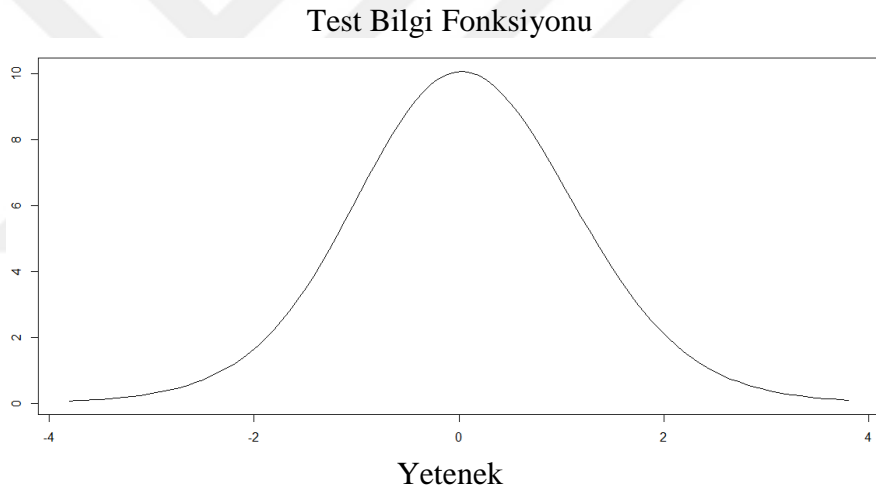
Şekil 18. Yirmi Beş Maddeden ve 500 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu



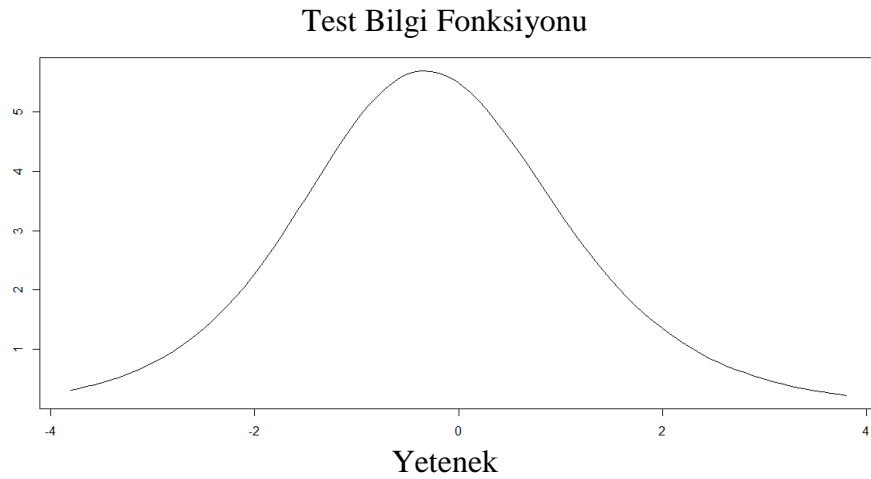
Şekil 19. Yirmi Beş Maddeden ve 1000 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu



Şekil 19. Yirmi Beş Maddeden ve 3000 Kişiden Oluşan Veri Setine İlişkin Kestirilen Test Bilgi Fonksiyonu

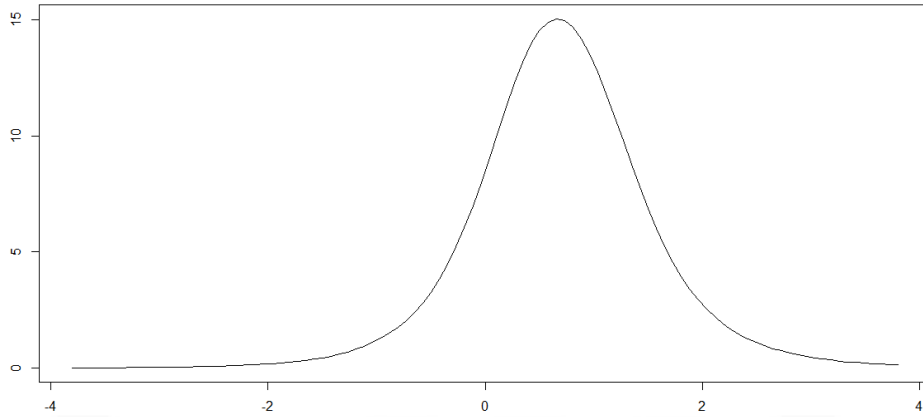


Şekil 20. A Parametresi Yüksek Maddelerden Oluşan Teste İlişkin Test Bilgi Fonksiyonu



Şekil 21. A Parametresi Düşük Maddelerden Oluşan Testin Test Bilgi Fonksiyonu

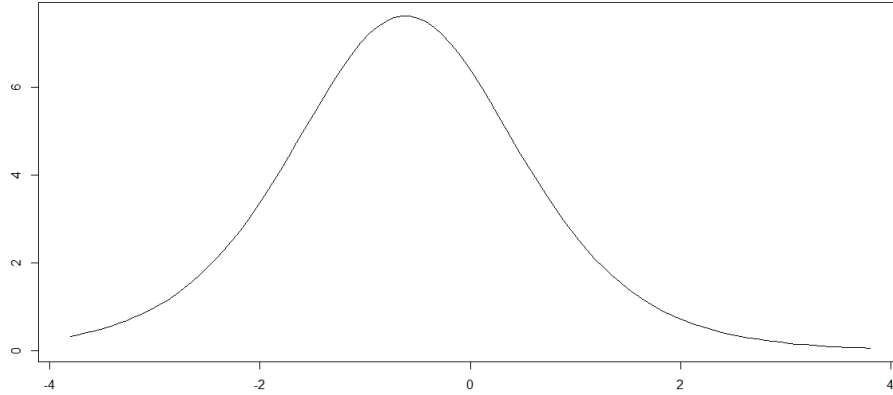
Test Bilgi Fonksiyonu



Yetenek

Şekil 22. B Parametresi Yüksek Maddelerden Oluşan Teste İlişkin Test Bilgi Fonksiyonu

Test Bilgi Fonksiyonu



Yetenek

Şekil 22. B Parametresi Düşük Maddelerden Oluşan Teste İlişkin Test Bilgi Fonksiyonu