



# HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ KURAMINDA  
FARKLI ÖRNEKLEM BÜYÜKLÜKLERİNE VE BOYUTLULUKLARINA GÖRE  
PARAMETRE DEĞİŞMEZLİĞİNİN İNCELENMESİ

Çiğdem REYHANLIOĞLU KEÇEOĞLU

Doktora Tezi

Ankara, 2018

Liderlik, arařtırma, inovasyon, kaliteli eđitim ve deđiřim ile

*Daha ileriye ... En İyiyeye ...*



**HACETTEPE ÜNİVERSİTESİ**  
**EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ KURAMINDA  
FARKLI ÖRNEKLEM BÜYÜKLÜKLERİNE VE BOYUTLULUKLARINA GÖRE  
PARAMETRE DEĞİŞMEZLİĞİNİN İNCELENMESİ

EXAMINING OF THE PARAMETER INVARIANCE ACCORDING TO DIFFERENT  
SAMPLE SIZES AND DIMENSIONALITIES IN PARAMETRIC AND  
NONPARAMETRIC ITEM RESPONSE THEORY

Çiğdem REYHANLIOĞLU KEÇEOĞLU

Doktora Tezi

Ankara, 2018

### Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,  
Çiđdem REYHANLIOđLU KEÇEOđLU'nun hazırladıđı "Parametrik ve Parametrik Olmayan Madde Tepki Kuramında Farklı Örnekleme Büy¼kl¼klerine ve Boyutluluklarına göre Parametre Deđiřmezliđinin İncelenmesi" bařlıklı bu çalıřma j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalında Doktora Tezi** olarak kabul edilmiřtir.

J¼ri Bařkanı

Prof. Dr. Selahattin GELBAL

İmza

J¼ri Üyesi (Danıřman)

Doç.Dr. Nuri DOđAN

İmza

J¼ri Üyesi

Prof. Dr. H¼lya KELECİOđLU

İmza

J¼ri Üyesi

Doç.Dr. İsmail KARAKAYA

İmza

J¼ri Üyesi

Dr. Öğr. Üyesi Hamide Deniz G¼LLEROđLU

İmza

Bu tez Hacettepe Üniversitesi Lisans¼st¼ Eđitim, Öğretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 27 / 04 / 2018 tarihinde uygun gör¼lm¼ř ve Enstit¼ Yönetim Kurulunca ..... / ..... / ..... tarihinde kabul edilmiřtir.

Prof. Dr. Ali Ekber řAHİN  
Eđitim Bilimleri Enstitüsü M¼d¼r¼

## Öz

Bu çalışmanın amacı farklı boyutluluk ve örneklem özelliklerinde madde tepki kuramı (MTK) uygulamalarına göre kestirilen parametrelerin değişmezliğini incelemektir. Bu amaç doğrultusunda 2015 yılındaki TEOG'un birinci uygulamasındaki A kitapçığını alan öğrenci cevapları araştırma verisi olarak kullanılmıştır. Çalışma evreninin büyüklüğü 63.871'dir. Evrenden rastgele seçilmiş 50, 100, 200, 500, 1000 ve 5000 kişilik gruplar çalışmanın örneklemini oluşturmaktadır. MTK'nin tek boyutlu uygulamalarında Matematik alt testi tek boyutlu model analizlerinde kullanılmıştır. Çok boyutlu MTK analizleri için ise araştırmacı tarafından çok boyutlu bir yapı oluşturulmuştur. Fen Bilgisi ile Din Kültürü ve Ahlak Bilgisi testlerinden belli maddelerin birleştirilmesiyle yapay olarak oluşturulan iki boyutlu testin sonuçlarından yararlanılmıştır. Tek boyutlu parametrik MTK (TBPMTK) için 2 Parametrelili Lojistik Model, tek boyutlu parametrik olmayan MTK (TBPOMTK) için Monoton Homojenlik Modeli ve çok boyutlu MTK (ÇBMTK) için Genişletilmiş 2 Parametrelili Lojistik Model'den yararlanılmıştır. Çalışmadan elde edilen bulgular sonucunda tek boyutlu testin TBPOMTK'ye göre analiz edilmesiyle madde parametresi ortalamalarının örneklem büyüklüğüne bağlı olarak büyük değişimler göstermediği sonucuna ulaşılmıştır. Bununla birlikte 200 örneklem büyüklüğü itibarıyla madde parametresi değişmezliği sağlanmıştır. TBPMTK'ye göre analiz edildiğinde ise evren değere yakın madde parametresi kestirimleri için en az 1000 örneklem büyüklüğü ile çalışılması gerektiği sonucuna ulaşılmıştır. İki boyutlu testin parametrik ve parametrik olmayan MTK'ye ve ÇBMTK'ye göre analiz edilmesi ile elde edilen madde parametreleri için değişmezliğin sağlanmadığı sonucuna ulaşılmıştır. Parametrik ve parametrik olmayan MTK'ye göre ve tek boyutlu testten elde edilen yetenek parametreleri arasında korelasyon türü farketmeksizin yüksek bir ilişki bulunmuştur. Tek boyutluluk varsayımının ihlal edilmesi durumunda da benzer sonuçlar elde edilmiştir. Benzer şekilde TBPMTK ve ÇBMTK'den elde edilen çok boyutlu teste ait yetenek parametreleri arasında örneklem büyüklüğü ve korelasyon türü farketmeksizin istatistiksel olarak anlamlı ve mükemmel yakın ilişki düzeylerinin olduğu tespit edilmiştir. Son olarak tek ve çok boyutlu testlere ait parametrik ve parametrik olmayan MTK ile ÇBMTK'den elde edilen güvenilirlik katsayılarının evren değerine yakın büyüklükte oldukları sonucuna ulaşılmıştır.

**Anahtar sözcükler:** MTK uygulamaları, örneklem büyüklüğü, parametre değişmezliği.

## Abstract

The aim of this study is to determine the invariance of the estimated parameters in regard to Item Response Theory (IRT) applications within different dimensionality and sample characteristics. To realize this aim, students' answers in 'A' test booklet of the first stage of TEOG applied in 2015 were used as the data of the study. The universe size of the study was determined as 63.871. The groups of 50, 100, 200, 500, 1000 and 5000 people who were randomly selected from the universe comprised the sample of the study.. mathematics subtest, which gives the assumptions that are necessary for one-dimensional applications of IRT, was used in one dimensional model analyses. Since all subtests are one-dimensional, a multidimensional construct was formed by the researcher to be used in multidimensional IRT analyses. Thus, for multidimensional IRT analyses, the results of a two-dimensional test, which was artificially formed by combining certain items from Science and Religion Culture and Moral Knowledge tests, were used. Two-parameter logistic model for one-dimensional parametric IRT (PIRT), Monotone Homogeneity Model for one-dimensional non-parametric IRT (NIRT) and multidimensional extension of the two-parameter Logistic Model for multidimensional IRT (MIRT) were used. As a result of the study in which one dimensional test was analyzed in regard to NIRT, it was found that item parameter means did not show significant changes based on the sample size. However, item parameter invariance was found when the sample size was 200 and larger. When it was analyzed in regard to PIRT, it was found that at least a sample size of 1000 have to be used for item parameter estimations nearly as high as the universe value (parameter). It was found that invariance for item parameters was not obtained by analyzing the two-parameter test in regard to parametric and nonparametric IRT and MIRT. Regardless of the correlation type, there was a high correlation between the parametric and nonparametric IRT and the ability parameters obtained from the one-dimensional test. Similar results were obtained when the assumption of unidimensionality was violated. Similarly, the ability parameters of the multidimensional test obtained from PIRT and MIRT were found to be statistically significant and have a correlation level close to perfect regardless of the sample size and correlation type. Finally, the parametric and nonparametric IRT of one-

dimensional and multidimensional tests and the reliability coefficients obtained from MIRT were found to be nearly as high as the universe value (parameter).

**Key words:** IRT applications, sample size, parameter invariance.



## Teşekkür

Lisansüstü eğitimim boyunca her zaman ve her koşulda dünyanın bir ucunda olsa bile benden yardımlarını esirgemeyen donanımıyla, birikimiyle, benzersiz kişiliği ile bu çalışmayı ortaya koymamı sağlayan Danışman Hocam **Prof. Dr. Nuri DOĞAN'a**,

Hacettepe Üniversitesi'ne adım attığım günden itibaren lisansüstü eğitim hayatımın önemli gelişmelerinde kilit rolleri üstlenen **Prof. Dr. Selahattin GELBAL'a**,

Bugüne kadar girmiş olduğum bütün kritik jürilerimde bulunan ve bana uğur getiren **Prof. Dr. Hülya KELECİOĞLU'na**,

Mütevazı kişiliğiyle, anneliğiyle, akademik bilgi birikimiyle kendime rol model aldığım **Doç. Dr. Burcu ATAR'a**,

Çalışmama önemli katkılar sağlayan çok değerli hocalarım **Doç. Dr. İsmail KARAKAYA** ve **Dr. Öğr. Üyesi H. Deniz Gülleroğlu'na**,

Edinimleriyle akademik ve saha çalışmalarına önemli katkılar getiren ve yoğun iş temposu içerisinde çalışmalarımı ilgili moral ve motivasyonumu her zaman yüksek tutmamı sağlayan **GKV Özel Okulları Genel Md. Yrd. İsmet TIRYAKI'ye**,

Günün hangi saati olursa olsun akademik ve teknik desteğini hiçbir zaman esirgemeyen değerli arkadaşım, meslektaşım **Dr. Eren Halil ÖZBERK'e**,

Bugünlere gelmemde kuşkusuz en büyük paya sahip olan anneciğim **Nermin REYHANLIOĞLU'na** ve babacığım **İsmail REYHANLIOĞLU'na**,

Önümüzde her zaman güzel ufuklar açan canımız ablamız **Uzm. Dr. Dilek Kafadar'a**,

Anne karnından doktora kadar hayatın ve eğitimin bütün kademelerinde beraber olduğum ikizim, diğer yarım **Dr. Didem ÖZDOĞAN'a**,

Lisansüstü eğitimim boyunca göstermiş olduğu sebatkârlığı ve özverisi için değerli eşim **Uzm. Dr. Serdar KEÇEOĞLU'na**,

Yaşama sebebim ve sevincim, en büyük güç kaynağım, canım oğlum **Serdar Emir KEÇEOĞLU'na**,

Doğduğumuz günden bu yana duaları hiç bir zaman üzerimizden eksik olmayan canımız anneannemiz **Necmiye PAKÖZ'e**,

Dünyaya geliřiyle hayatımıza bambařka renkler katan yeęenlerim **Mehmet Ali KAFADAR** ve **Zeynep Ela KAFADAR'a**,

Bu meřakkatli süreçte biręok zorluęu birlikte ařtıęımız dostlarım **Dr. Gizem SARAL ERTOPRAK'a** ve **Gülen ULUSOY'a**,

Son olarak buraya ismini sıędıramadıęım ama desteęini yüreęimde hissettięim herkese ve herřeeye sonsuz teřekkürlerimi sunarım.

Siz olmasaydınız asla bařaramazdım.

## İçindekiler

Öz.....	ii
Abstract.....	iv
Teşekkür.....	vi
Tablolar Dizini.....	x
Şekiller Dizini.....	xi
Simgeler ve Kısaltmalar Dizini.....	xii
Bölüm 1 Giriş.....	1
Problem Durumu.....	5
Araştırmanın Amacı ve Önemi.....	6
Araştırma Problemi.....	9
Sayıtlar.....	10
Sınırlılıklar.....	10
Bölüm 2 Araştırmanın Kuramsal Temeli Ve İlgili Araştırmalar.....	11
Madde Tepki Kuramı.....	11
Madde Tepki Kuramının Varsayımları.....	12
Madde Tepki Kuramı Modelleri.....	14
Madde Tepki Kuramında Güvenirlik.....	26
İlgili Araştırmalar.....	28
Bölüm 3 Yöntem.....	38
Araştırma Modeli.....	38
Araştırmanın Evreni ve Örneklemi.....	38
Veri Toplama Araçları.....	39
Değişkenler.....	40
Veri Toplama Süreci.....	40
Verilerin Analizi.....	40
Araştırmanın İç ve Dış Geçerliği.....	52
Bölüm 4 Bulgular Ve Yorumlar.....	54
1. Alt Problemin Çözümüne İlişkin Bulgular.....	54
2. Alt Problemin Çözümüne İlişkin Bulgular.....	57
3. Alt Problemin Çözümüne İlişkin Bulgular.....	61
4. Alt Problemin Çözümüne İlişkin Bulgular.....	64
5. Alt Problemin Çözümüne İlişkin Bulgular.....	66
6. Alt Problemin Çözümüne Yönelik İlişkin Bulgular.....	68

7. Alt Problemin Çözümüne İlişkin Bulgular .....	70
Bölüm 5 Sonuç, Tartışma Ve Öneriler .....	74
Kaynakça .....	83
EK-A: Türkçe Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu.....	90
EK-B: Matematik Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu ....	91
EK-C: T.C. İnkılap Tarihi ve Atatürkçülük Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu.....	92
EK-Ç: İngilizce Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu .....	93
EK-D: Fen Bilgisi Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu ...	94
EK-E: Din Kültürü ve Ahlak Bilgisi Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu.....	95
EK-F: Birleşik Test'e Ait Faktör Analizi Sonucu .....	96
EK-G: Birleşik Test'ten Elde Edilen %20'lik Dilime Ait Maddeler Arası Korelasyon Matrisi .....	97
EK-I: TBPMTK İçin Yazılan Kodlar .....	99
EK-İ: ÇBMTK İçin Yazılan Kodlar .....	100
EK-J: TBPOMTK İçin Yazılan Kodlar .....	102
EK-K: Etik Kurul İzin Muafiyeti Formu .....	106
EK-L: Etik Beyanı .....	108
EK-M: Doktora Tez Çalışması Orijinallik Raporu .....	109
EK-N: Dissertation Originality Report.....	110
EK-O: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı.....	111

## Tablolar Dizini

Tablo 1 KMO ve Bartlett Küresellik Testi.....	42
Tablo 2 Paralel Analiz Sonucunda Elde Edilen Uyum İndeksleri .....	43
Tablo 3 Matematik Testini Gizil Monotonluk Varsayımına Ait Analiz Sonuçları.....	46
Tablo 4 Matematik Testine ve Birleşik Teste Ait Betimsel İstatistik Sonuçları.....	47
Tablo 5 TBPMTK ve TBPOMTK'ye Göre Kestirilen Matematik Testi'ne Ait Madde Parametreleri ve Standart Hata Ortalamaları.....	54
Tablo 6 Tek ve Çok Boyutlu MTK'ye Göre Kestirilen Birleşik Teste Ait Madde Parametreleri ve Standart Hata Ortalamaları.....	57
Tablo 7 TBPMTK ve TBPOMTK'ya Göre Kestirilen Birleşik Teste Ait Madde Parametreleri ve Standart Hata Ortalamaları.....	61
Tablo 8 TBPOMTK ve TBPMTK'ya Göre Kestirilen Matematik Testine Ait Yetenek Parametreleri Arasındaki İlişki .....	64
Tablo 9 TBPMTK ve ÇBMTK'ye Göre Kestirilen Birleşik Teste Ait Yetenek Parametreleri Arasındaki İlişki .....	66
Tablo 10 TBPMTK ve TBPOMTK'ye Göre Kestirilen Birleşik Teste Ait Yetenek Parametreleri Arasındaki İlişki .....	68
Tablo 11 TBPMTK ve TBPOMTK 'ye Göre Matematik Testi için Kestirilen Güvenirlilik Katsayıları .....	71
Tablo 12 Birleşik Test için TBPMTK, TBPOMTK ve ÇBMTK'ye Göre Kestirilen Güvenirlilik Katsayıları .....	72

## Şekiller Dizini

Şekil 1. Monoton Azalmayan MTF Örnekleri .....	17
Şekil 2. 2 PLM'ye ait MKE .....	21
Şekil 3. M2PLM'ye ait Madde Karakteristik Yüzeyi .....	26

## Simgeler ve Kısaltmalar Dizini

**MTK:** Madde Tepki Kuramı

**MTF:** Madde Tepki Fonksiyonu

**MKE:** Madde Karakteristik Eğrisi

**POMTK:** Parametrik Olmayan Madde Tepki Kuramı

**PMTK:** Parametrik Madde Tepki Kuramı

**TBPOMTK:** Tek Boyutlu Parametrik Olmayan Madde Tepki Kuramı

**TBPMTK:** Tek Boyutlu Parametrik Madde Tepki Kuramı

**ÇBMTK:** Çok Boyutlu Madde Tepki Kuramı

**1PLM:** 1 Parametrelili Lojistik Model

**2PLM:** 2 Parametrelili Lojistik Model

**3PLM:** 3 Parametrelili Lojistik Model

**MHM:** Monoton Homojenlik Modeli

**IMM:** İkili Monotonluk Modeli

**M1PLM:** Genişletilmiş 1 Parametrelili Lojistik Model

**M2PLM:** Genişletilmiş 2 Parametrelili Lojistik Model

**M3PLM:** Genişletilmiş 3 Parametrelili Lojistik Model

**TEOG:** Temel Öğretimden Ortaöğretime Geçiş

**-2LL:** -2 LogLikelihood

**PMÇKK:** Pearson Momentler Çarpımı Korelasyon Katsayısı

**SHO:** Standart Hata Ortalaması

## Bölüm 1

### Giriş

Alanyazın incelendiğinde ölçmenin farklı tanımlarına rastlamak mümkündür. Lord ve Novick'e göre (1968) ölçme, davranış alanındaki ilişkileri koruyacak ve onların özelliklerini niteleyecek biçimde, deneysel birimlerin tanımlanmış özelliklerine göre sayılarla ayrılma işlemi; Crocker ve Algina'ya (1986) göre, bir gözlemci tarafından fiziksel olaylar üzerinde yapılan bir işlem; Turgut'a (1987) göre, bir niteliğin gözlenip gözlem sonucunun sayı veya başka sembollerle gösterilmesi; Tekin'e (1993) göre, belli bir nesnenin ya da nesnelerin belli bir özelliğe sahip olup olmadığını, sahipse oluş derecesinin gözlenip gözlem sonuçlarının sembollerle ve özellikle sayı sembolleriyle ifade edilmesi ve son olarak Murphy ve Davidshofer'a (2005) göre ise belli özelliklerin sayılarla temsil edilmesidir.

Bireylerin ölçmeye konu olan örtük özellikleri (psikolojik yapıları, akademik başarı, tutum vb. gibi) değişken yapıdadır. Bir başka ifadeyle bu özellikler zamanla ve çeşitli uygulamalar sonucunda değişim gösterebilir. Oysa ölçmeye konu olan fiziksel özellikler için bu değişkenlik daha azdır. Bu nedenle genelde fiziksel özelliklerin ölçme konusu olduğu fen bilimlerinde tanımlanan yapıların sosyal bilimlerde ölçülen örtük özelliklere göre daha kararlı bir yapıya sahip olduğu söylenebilir.

Psikolojik yapıların ölçülmesinde psikolojik ölçme araçlarından yararlanılır ve bu ölçme araçlarından belli özellikleri karşılamaları istenir. Ölçme araçlarında aranan özelliklerden bir tanesi yakın bir geçmişe kadar tek boyutlu olmalarıydı. Bunun nedeni genellikle ölçme modellerinin psikolojik yapıların tek boyutlu olduğu kabulü üzerine kurulmuş olmasıydı ancak bu durum bazı yapıları açıklamada yetersiz kalmıştı. Böylece psikolojik yapıların bazılarının çok boyutlu olması sebebi ile bu yapılara uygun ölçme modellerinin geliştirilmesi fikri ortaya çıktı ve günümüze kadar çok boyutlu yapıların açıklanmasında kullanılan birçok model geliştirildi. Bütün bu gelişmelere rağmen yapı çok boyutlu olsa bile tek boyutlu ölçme modellerinden yararlanılması sık rastlanan bir uygulamadır. Böyle durumlarda tek ve çok boyutlu modellerden elde edilen test ve madde parametreleri değişmekte ve bu değişimin ölçülen özellikten mi yoksa yanlış ölçme modelinin kullanılmasından mı



kaynaklandığının ortaya konması gerekir. Aksi halde ölçme sonuçlarının geçerliği tehlikeye girmiş olur.

Tek boyutluluk genel olarak bireyin madde üzerindeki performansını etkileyen tek bir yetenek türünün olması durumu şeklinde tanımlanır (Hambleton, Swaminathan ve Rogers 1991). Crocker ve Algina'ya göre (1986) tek boyutlu yapılarda maddeler uygulanan çalışma grubu için istatistiksel olarak bağımlıdır ve paylaştıkları ortak bir varyansın olduğu söylenebilir. Paylaşılan ortak bir varyansın olması ve bu varyansın olabildiğince büyük olması tek boyutlu yapıları ölçen psikolojik ölçme araçları için istenen bir durumdur. Buradan hareketle tek boyutlu psikolojik ölçme araçlarından elde edilen verilerin analizi için birçok tek boyutlu ölçme modeli ortaya konmuştur.

Tek boyutlu yapılara ilişkin ölçme sonuçlarının analizinde kullanılan tek boyutlu parametrik modeller, uzun yıllardır alanyazında ve saha çalışmalarında kendilerine geniş bir yer edinmişlerdir. Ancak tek boyutlu parametrik modellerin en az eşit aralık ölçek düzeyinde olan ve normal dağılım gösteren verilere uygulanma zorunluluğu, parametrik modeller için önemli birer sınırlılık olarak görülmektedir. Çünkü psikolojik yapıların ölçülmesi ile elde edilen sonuçlar yaygın olarak sınıflama ve sıralama ölçeği düzeyindedir. Bu sınırlılığın üstesinden gelebilmek için sıralama düzeyindeki ölçümlere de uygulanabilen yine tek boyutlu ancak parametrik olmayan modellerin geliştirilmesi önemli bir ihtiyaç haline gelmiştir. Böylece tek boyutlu parametrik modellere göre daha esnek modeller olan tek boyutlu parametrik olmayan modeller geliştirilmiştir. Yine de ister parametrik ister parametrik olmayan tek boyutlu modeller olsun, kullanılan verilerin tek boyutluluk varsayımını karşılaması gerektiği unutulmamalıdır. Ancak her zaman ölçme sonuçlarından elde edilen verilerin tek boyutlu olması mümkün olmamaktadır. Çok boyutlu psikolojik değişkenlere uygulanamaması veya uygulansa bile yorumlamada güçlük çıkarması tek boyutlu modellerin sınırlılığı olarak belirtilebilir.

Eğitim uygulamalarında bir yeteneği ölçmek için geliştirilen bir test nadiren sadece o yeteneği ölçüyordur ve sıklıkla başka yetenekler ya da psikolojik özellikler farklı düzeylerde ölçme sonuçlarını etkilemektedir. Örneğin, bir matematik başarı testini ele alındığında, soruları cevaplamak için öğrencilerin matematik dışında Türkçe bilgisinin (dil yeteneğinin) gerekliliğini de göz önünde bulundurmaları gerekir.

Bu durumda, öğrencilerin matematik başarısını ölçen maddeler aynı zamanda onların Türkçe dilindeki yeteneklerini de yansıtabilir. Eğer öğrencilerin dil yetenekleri birbirinden farklı ise, öğrencilerin matematik dersindeki başarısı onların matematik alanındaki yeteneklerinin yanında Türkçe dilindeki yeterliklerine bağlı olacaktır (Özbek Baştuğ, 2012). Örnekten de anlaşılacağı üzere, özellikle başarı testlerinde bulunan problemlerin çözümünde gerekli olan beceriler için, ölçeklerde ideal olan tek boyutluluk özelliğinin tam olarak karşılanması pek mümkün değildir (Traub, 1983). Ackerman'ın (1994) de belirttiği gibi, başarı testinde yer alan maddeler gerekli beceri düzeylerinde farklılaşma gösteriyorsa, kullanılacak test çok boyutlu bir yapı altında modellenmelidir.

Eğitimde ve psikolojide kullanılan testler farklı alt bölümlerden oluşabilir. Örneğin, bir fen bilgisi testi fizik, kimya ve biyoloji olmak üzere başlı başına üç farklı disipline yönelik sorulardan oluşur. Ya da bir matematik testi ayrı ayrı cebir ve geometri bilgisini ölçen sorulardan meydana gelir. Bu soruların her biri kendi içinde bir araya gelerek alt testleri oluşturur. Bu alt testlerden alınan puanların her biri alt test puanı olarak tanımlanır (Sinharay, Haberman ve Puhan, 2007).

Öğrencilerin başarı durumlarının bir göstergesi olan test sonuçları toplam test puanı veya alt test puanları olmak üzere iki şekilde hesaplanmaktadır. Türkiye'de öğrenciler hakkında önemli kararların alındığı üniversiteye giriş sınavları gibi ulusal düzeyde uygulanan geniş ölçekli sınavlardan alınan puanlar, testin bütünü üzerinden hesaplanmaktadır. Toplam test puanını hesaplama yollarından biri, bireylerin ölçmeye konu olan özelliklerini, kendilerinde bulunma derecesine göre verdikleri yanıtlar doğrultusunda doğru ya da yanlış kabul edilme ölçütüne göre puanlanmasıdır (Özer Özkan, 2012). Bu hesaplama işlemi, üniversiteye giriş sınavları gibi geniş ölçekli sınavlarda kullanılan testlerin eşit aralık ölçek düzeyinde kabul edilmesiyle yapılabilir. Bireylerin hesaplanan bu puanları standart puanlara çevrilir ve en yüksek puan alandan en düşük puan alan öğrenciye doğru bir başarı sıralaması yapılır. Öğrenciler başarı puanlarına göre değil, başarı sıralamalarına bağlı olarak bir yükseköğretim kurumuna yerleştirilirler. TEOG uygulamaları için de benzer bir durum söz konusudur. TEOG'da yer alan alt testlerin eşit aralıklı ölçek düzeyinde oldukları kabulünden hareketle öğrencilerin sınav puanları hesaplanır. Bu puanlar standart puanlara çevrilerek öğrenciler, en yüksek puan alandan en

düşük puan alan öğrenciyeye doğru sıralanırlar ve başarı sıralarına göre ortaöğretim kurumlarına yerleştirilirler. Yükseköğretime giriş sınavlarında olduđu gibi ölçme sonuçları eşit aralık ölçek düzeyindeki bir kabulden yola çıkarak hesaplanır, ancak yerleştirme işlemi ölçme sonuçlarının sıralama ölçeğine dönüştürülmesiyle gerçekleştirilir. Böyle durumlarda metrik ölçüm gerektirmeyen sıralama düzeyindeki ölçme sonuçlarına da uygulanabilecek kuram ve modellerin geliştirilmesi ve kullanılması gerekir.

Hemen hemen bütün bilim dallarının bir deneysel bir de kuramsal yapısı vardır. Ölçme hem kuramsal olarak hem de uygulama olarak birçok bilim insanının ilgisini çekmiştir. Ölçmeye kuramsal yaklaşım formal incelemeye yol açmış, sonuçta bazı ölçme kuramları geliştirilmiştir (Baykul ve Turgut, 1992). Geliştirilen bu kuramlar sayesinde farklı ölçek türlerinden elde edilen sonuçlardan geçerliđi ve güvenirliliđi yüksek bulgular elde etmek mümkündür.

Ölçme kuramları incelendiđi zaman bir çok varsayım gerektiren parametrik yöntemlerin kullanılması oldukça yaygındır. Ancak parametrik koşulların oluşmadıđı ve tek boyutluluđun sağlanmadıđı durumlara eğitim uygulamalarında rastlamak muhtemeldir. Böyle durumlarla karşılaşma sıklıđı parametrik olmayan ve çok boyutluluk için kullanılan uygulamaların geliştirilmesini zorunlu kılmıştır. Ancak geliştirilen bu kuramların işlerliđinin ortaya konması gerekir. Bu amaçla geliştirilen kuramlar çerçevesinde elde edilen sonuçlar, mevcut kuramlardan elde edilen sonuçlarla farklı deđişkenler açısından karşılaştırılabilir. Bu deđişkenlerden bir tanesi örneklem büyüklüđüdür.

Araştırmaların veri kaynađı olan örneklem büyüklüđü istatistik seçimini etkiler. Büyük gruplar üzerinden toplanan verilerin normallik gibi parametrik koşulları sağladığı kabul edilir ve buna uygun istatistikler kullanılabilir (Çokluk ve diđerleri, 2012). Kullanılan örneklemin büyüklüđü özellikle parametrik yöntemlerde, elde edilen sonuçların doğruluđunu etkilemektedir. Parametrik yöntemler analiz sonuçlarını örneklem büyüklüđüne bađlı olarak araştırmacıya sunduđu için, kullanılan yöntemeye uygun olan örneklem büyüklüđünü belirlemek araştırmacının özellikle geçerliđi açısından önemlidir. Bu da demektir ki kullanılan istatistiksel yöntemeye uygun olan örneklem büyüklüđü ile çalışmak gerekir. Özellikle MTK'nin madde ve yetenek parametrelerinin birbirinden bađımsız kestirilmesi iddası için

örneklem büyüklüğü önemli bir faktördür (Hambleton, 1989). Parametre değişmezliğinin sağlanması için kullanılan modele uygun olan örneklem büyüklüğünün seçilmesi gerekir. Ancak Literatürde örneklem büyüklüğünün ne olması gerektiğine dair yapılmış birçok çalışma olmasına rağmen mutabık kalınmış ortak ve kesin bir görüş yoktur. Linacre (1994), örneklem büyüklüğü ile parametre sayısı arasında doğru bir orantı kurmuştur. Lord (1968) parametrik yöntemler için 1000'den az olan örneklemelerin doğru sonuçlar vermediğini belirtmiş hatta örneklem büyüklüğünün 10.000 olması gerektiğini vurgulamıştır. Buna karşılık Hulin, Lissak ve Drasgow (1982) ise parametrik MTK analizlerinde 2000'den daha büyük bir örneklemelerden elde edilen sonuçlar arasında anlamlı farklılıkların olmadığını belirtmiştir. Molenaar (2001) ise küçük örneklemelerle çalışılacağı zaman parametrik yöntemlerin tercih edilmemesi gerektiğini savunmuştur. Molenaar'a (2001) göre parametrik olmayan MTK modellerinde doğru sonuçlar elde etmek için örneklem büyüklüğünün 300 civarında olması gerekir. Sonuçta literatürde örneklem büyüklüğüne dair kesin bir görüş yoktur. Bu nedenle en iyi sonucu veren örneklem büyüklüğünün belirlenmesi için farklı örneklem büyüklüklerinden elde edilen sonuçların karşılaştırılmasına ilişkin yeni çalışmalara ihtiyaç duyulmaktadır. Bu çalışmada sonuçların örneklem büyüklükleri açısından karşılaştırılması hedeflenen uygulamalar madde tepki kuramına ait tek boyutlu parametrik ve parametrik olmayan uygulamalar ile çok boyutlu parametrik uygulamalardır.

## **Problem Durumu**

Bulguların elde edildiği örneklemin büyüklüğüne karar vermek bilimsel çalışmaların en çok tartışılan konularından biridir. Bu nedenle bu çalışmanın odağındaki problemlerden bir tanesi örneklem büyüklüğüdür. Herhangi bir örneklem büyüklüğünün uygunluğuna karar vermek için, kullanılan örneklemin sonuçları ile evrenden elde edilen bulgular arasında istatistiksel olarak bir farklılığın olup olmadığının incelenmesi gerekir. Bir başka ifadeyle, bir örneklem büyüklüğü evreni yansıttığı düzeyde, ilgili çalışma için uygun bir büyüklüktür. Bu durum için ilk akla gelen çözümlerden bir tanesi evrene yakın örneklem büyüklükleri ile çalışmak olabilir. Ancak her zaman evrene yakın örneklem büyüklüklerine ulaşamayabilir. Bu nedenle küçük örneklem büyüklükleri ile de evren değerine yakın sonuçlar üretebilen parametrik olmayan kuram ve modeller geliştirilmiştir. Ancak bu

modellerden elde edilen sonuçların birçok bilimsel çalışmada test edilmiş mevcut kuram ve modellerden elde edilen sonuçlar ile karşılaştırılması gerekir. Bu çalışmada örneklem büyüklüğü problemine açıklık kazandırmak amacıyla küçük örneklem büyüklükleri ile elde edilen bulguların değerlendirilmesi için geliştirilen parametrik olmayan madde tepki kuramı modelleri ile bu kuramın parametrik karşılığı olan modellerden elde edilen sonuçlar karşılaştırılmıştır.

Eğitim ve psikolojide karşılaşılan bir diğer problem ise ölçülen yapının boyutluluğuna uygun model seçimidir. Ölçülen yapı çok boyutlu olsa bile tek boyutlu ölçme modellerinden yararlanılması sık karşılaşılan bir uygulamadır. Böyle durumlarda tek ve çok boyutlu modellerden elde edilen test ve madde parametreleri değişmektedir. Bu değişimin sebebi ölçülen özellik ya da yanlış ölçme modelinin kullanılması olabilir. Bu probleme bir çözüm sunmak amacıyla bu çalışmada, ölçülen yapının tek boyutlu olmadığı durumda tek boyutlu yapılar için geliştirilen ve kullanılmakta olan tek boyutlu madde tepki kuramına ait modellerden elde edilen sonuçlar ile çok boyutlu modellerden elde edilen sonuçlar karşılaştırılmıştır.

### **Araştırmanın Amacı ve Önemi**

Eğitim ve psikolojide çeşitli amaçlarla uygulanan ölçeklerden elde edilen sonuçlara göre bireyler hakkında kararlar verilirken çeşitli ölçek düzeylerindeki verilerden yararlanır. Ölçme araçlarında yer alan sorular farklı güçlük, ayırıcılık ve bilgi düzeyine sahip olmasına rağmen soruların tümüne 1-0 puan verilmesi sık karşılaşılan bir uygulamadır. Bu şekilde maddelere eşit puan vermek onların eşdeğer olduğu varsayımına dayanmaktadır. Bu noktadan hareketle kullanılan ölçme aracının eşit aralık ölçek düzeyinde olduğu varsayılır. Bu varsayım doğrultusunda, bireylerin testlerden aldıkları puanlar işaretledikleri doğru ve yanlış cevap sayısına göre hesaplanabilmektedir.

Türkiye’de seçme ve yerleştirme amaçlı uygulanan geniş ölçekli sınavlardan alınan puanlar standart puanlara dönüştürülerek, en yüksek puan alan bireyden en düşük puan alan bireye doğru bir başarı sıralaması yapılır. Bir ölçüde ölçme sonuçları eşit aralık ölçek düzeyindeki bir kabulden yola çıkarak hesaplanır ancak yerleştirme işlemi ölçme sonuçlarının sıralama ölçeğine dönüştürülmesiyle gerçekleştirilir. Buradan anlaşılacağı üzere eğitim uygulamalarının araştırma

kapsamında incelenen alıřmalar, maddelerin analizi ve kiři ve maddeleri sıralamaya dayalı lekleme alıřmaları olabilmektedir. Bu tr alıřmalarda llmek istenen zellikler birbirlerinden bağımsız yapılar olarak llebilir. Bunun yanı sıra yapılan lmenin amacı, yalnızca kiřileri ya da maddeleri sıralamak olabilmektedir. Byle durumlarda metrik lm gerektirmeyen sıralama dzeyindeki lme sonularına da uygulanabilecek modellerin geliřtirilmesi ve kullanılması gerekir. Bu alıřma sıralama dzeyindeki lme sonularına da uygulanabilen parametrik olmayan modellerden elde edilen sonuları ortaya koyması aısından nemlidir.

lkemizde uygulanan birok geniř lekli sınav vardır. Bu sınavlardan elde edilen sonular doėrultusunda bireyler hakkında kritik kararlar verilmektedir. Uygulanan bu sınavlarda birden fazla alt test bulunmaktadır. Bir bařka deyiřle sınavlarda kullanılan testler ok boyutlu bir yapıya sahiptir. Bu nedenle bu alt testlere ait puanların ayrı ayrı hesaplandıktan sonra alınan kararlara dayanak oluřturmaları gerekir. nk bir testin ok boyutlu olduėu bir durumda bireylerin performansını belirlemede tm teste ait puanın kullanılıp kullanılmayacaėı veya nasıl kullanılacaėı tartiřılmaktadır (Meara, Robin ve Sireci, 2000). Dolayısıyla bir testin tek boyutlu olmadıėı bir durumda, testten elde edilen bireylere ait yetenek ve madde parametrelerinin kestirilmesi srecinde kullanılan modeller de nemlidir. Bu modeller tek boyutluluk varsayımı gerektirmeyen ok boyutlu verileri analiz edebilecek nitelikteki modeller olmalıdır. Bu alıřma, ok boyutlu veri analizinde kullanılan modellere yer verilmesinden dolayı nem arz etmektedir.

lme kuramları incelendiėi zaman bir ok varsayım gerektiren parametrik yntemlerin kullanılması olduka yaygındır. Ancak parametrik kořulların oluřamadıėı ve tek boyutluluėun saėlanmadıėı durumlara da eėitim uygulamalarında rastlamak muhtemeldir. Byle durumlara karřılařma sıklıėı parametrik olmayan ve ok boyutluluk iin kullanılan uygulamaların geliřtirilmesini zorunlu kılmıřtır. Ancak geliřtirilen bu kuramların iřlerliėinin ortaya konması gerekir. Bu amala geliřtirilen kuramlar erevesinde elde edilen sonular, mevcut kuramlardan elde edilen sonularla karřılařtırılabilir.

Parametrik MTK modellerinin geliřtirilmesi ve yaygınlařması modern test teorisinin geliřtirilmesinde ve kullanılmasında kuřkusuz nemli bir ařamadır. Bu

modellerin kullanılması için uygun örneklem büyüklüğüne dair bir görüş birliği olmasa da geçerli ve güvenilir düzeyde parametre kestirimleri için büyük örneklemelere ihtiyaç duyulduğuna dair önemli bulgular elde edilmiştir. Bu durum parametrik modellerin okullarda ve diğer saha çalışmalarında uygulanabilirliği açısından önemli bir sınırlıktır. Saha çalışmalarında özellikle okullarda küçük çalışma gruplarına uygulanabilmesi açısından parametrik olmayan MTK modellerinin parametrik modellere göre daha avantajlı olduğu düşünülmektedir. Ancak bunun için parametrik modeller kadar iyi çalıştığına ilişkin kanıt toplamaya ihtiyaç vardır ve bu çalışmadan elde edilen sonuçlar bu açıdan önemli kanıtlar ortaya koyacaktır. Bu çalışmada Türkiye’de kullanımı yeni yaygınlaşan bir kuram tek boyutlu parametrik olmayan MTK’den elde edilen sonuçlar ile bu kuramın parametrik karşılığı tek boyutlu MTK ve çok boyutlu MTK kapsamında elde edilen sonuçlar karşılaştırılmıştır.

Türkiye’deki alanyazın incelendiğinde, parametrik olmayan tek boyutlu MTK ile ilgili sınırlı sayıda çalışmaya rastlanmıştır. Koğar (2014) “Madde Tepki Kuramının Farklı Uygulamalarından Elde Edilen Parametrelerin ve Model Uyumlarının Örneklem Büyüklüğü ve Test Uzunluğu Açısından Karşılaştırılması” başlıklı çalışmasında parametre değişmezliğini farklı değişkenler açısından yapay veri üzerinden incelemiştir. Mor-Dirlik (2017) de benzer bir çalışmayı gerçek veri üzerinden gerçekleştirmiştir. Bu çalışma, gerçek veri üzerinden gerçekleştirilen Türkiye literatüründeki ilk çalışmalardan biri olması bakımından önemlidir.

Bilgi tekrarlanabilir bir yapıya sahiptir. Bilimde kesinlikçi bir anlayış hâkimdir; yani aynı nedenler, aynı koşullarda aynı sonuçları verir. Yinelenmeyen, neden-sonuç ilişkisine dayanmayan hiçbir bilgi bilimsel sayılmaz. Bilimsel bilgi, incelenen alanların değişik ortamlarda tekrarlanmasıyla elde edilen sonuçlarından oluşur. Aynı zamanda bilimsel bilgi birikimlidir. Her bilim insanının yaptığı bir çalışma, kendinden önceki çalışmaları destekler nitelikte olmalıdır. Bu bağlamda Türkiye ve dünya literatürü incelendiği zaman MTK uygulamalarından elde edilen sonuçların, farklı değişkenler açısından karşılaştırıldığı çok sayıda çalışmanın olduğu görülmüştür. Bu değişkenlerden bir tanesi örneklem büyüklüğüdür. Örneklem büyüklüğü, aynı amaca hizmet eden parametrik ve parametrik olmayan modellerden hangisinin kullanılması gerektiğine karar verirken göz önünde bulundurulması gereken önemli

bir faktördür. Bu çalışmada elde edilen sonuçların farklılaşıp farklılaşmadığı kuramsal uygulamaların yanı sıra örneklem büyüklüğüne göre de incelenmiştir. Bu bağlamda bu çalışma, MTK uygulamalarından elde edilen sonuçların örneklem büyüklüğüne göre karşılaştırıldığı çalışmalardan biri olması açısından da önemlidir. Bu çalışmada TEOG sınavının farklı boyutlardaki alt testlerinden elde edilen başarı puanlarının, örneklem büyüklükleri de göz önünde bulundurularak parametrik ve parametrik olmayan MTK ile çok boyutlu MTK kapsamında farklılaşıp farklılaşmadığı incelenmiştir.

### **Araştırma Problemi**

“Farklı boyutluluk ve örneklem büyüklüğü değişkenlerine göre parametrik ve parametrik olmayan MTK için parametrelerin değişmezliği ne düzeyde sağlanmaktadır?”

### **Alt Problemler**

1. Tek boyutlu testlerde, parametrik ve parametrik olmayan MTK'ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrenden kestirildiği durumda nasıldır?

2. Çok boyutlu testlerde, tek ve çok boyutlu MTK'ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrenden kestirildiği durumda nasıldır?

3. Çok boyutlu testlerde, parametrik ve parametrik olmayan MTK'ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrenden kestirildiği durumda nasıldır?

4. Tek boyutlu testlerde, parametrik ve parametrik olmayan MTK'ye göre hesaplanan yetenek kestirimleri arasında, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıl bir ilişki vardır?

5. Çok boyutlu testlerde, tek ve çok boyutlu MTK'ye göre hesaplanan yetenek kestirimleri arasında, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıl bir ilişki vardır?



6. Çok boyutlu testlerde, parametrik ve parametrik olmayan MTK'ye göre hesaplanan yetenek kestirimleri arasında, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıl bir ilişki vardır?

7. Tek ve çok boyutlu testlerde, TBPMTK, TBPOMTK ve ÇBMTK'ye göre hesaplanan güvenilirlik düzeyleri örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıldır?

### **Sayıtlılar**

Farklı örneklem büyüklüklerindeki verilerde aynı bireylerin bulunma oranının düşük olduğu ve sonuçları etkilemediği varsayılmıştır.

### **Sınırlılıklar**

Bu araştırmada elde edilen bulguların genellenebilirliği, sonuçlarının kullanıldığı sınavın ilgili olduğu alan ve psikometrik özellikleri, uygulanan kitlenin özellikleri, kullanılan madde tepki kuramı modelleri ve kullanılan örneklem büyüklükleri ile sınırlıdır.

## Bölüm 2

### Araştırmanın Kuramsal Temeli Ve İlgili Araştırmalar

#### Madde Tepki Kuramı

Eğitim ve psikolojide ölçme alanındaki en önemli gelişmelerden birisi Madde Tepki Kuramı'na (MTK) ait uygulamalardır. MTK, bir ölçeğin uygulandığı kitleden elde edilen verileri özel bir bağıntı kurarak modelleyen bir örtük özellikler kuramıdır (Smith, 2009). MTK, bireyin bir maddeye verdiği tepki ile o bireyin yetenek veya örtük özellik olarak da adlandırılan özelliği arasındaki ilişkiyi matematiksel olarak açıklayan bir modeldir (Reeve, 2006). Bir başka ifadeyle MTK, bir bireyin maddeye verdiği tepki ile maddenin yokladığı yeteneğin ya da örtük özelliğin arasındaki ilişkiyi matematiksel bir model ile tanımlayan bir örtük özellik kuramıdır (Embretson ve Reise, 2000). Bu tanımlama ise Madde Tepki Fonksiyonu (MTF) adı verilen olasılıksal modeller sayesinde gerçekleştirilir. Bu fonksiyon MTK'nin özünü ifade eder (Ackerman,1992).

MTF, MTK'nin özünü ifade eden, testi alan bireyin örtük özelliği ile madde parametreleri arasında matematiksel bir ilişki kuran olasılıksal bir modeldir (Ackerman, 1992). MTF, belli bir yetenek düzeyindeki cevaplayıcının maddeyi doğru cevaplama olasılığını veren matematiksel bir fonksiyondur. MTF'nin en önemli özelliği, madde parametrelerinin değişmezliğidir (Hambleton, Swaminathan ve Rogers 1991).

MTK'de bireylerin maddeye verdiği tepkiyi yetenek düzeylerinin bir fonksiyonu olarak gösteren grafiklerden yararlanır. Bu grafiklere Madde Karakteristik Eğrisi (MKE) adı verilir. MKE'nin şekli sadece testi alan bir bireyin, bir maddeye doğru cevap verme olasılığına bağlıdır. Bu olasılık, testi alan diğer bireylerin yetenek dağılımlarından tamamen bağımsızdır. Başka bir ifade ile maddeyi karakterize eden parametreler cevaplayıcıların yetenek dağılımına bağlı değildir. Bunun tersi de doğrudur. Şöyle ki: Bir cevaplayıcıyı tanımlayan yetenek parametresi bir grup test maddesine bağlı değildir. MTF'nin bu değişmezlik özelliği bütün madde tepki modellerinin bir özelliğidir (Köse, 2010).

Bütün MTK modelleri için geçerli olan bir başka özellik ise MTK'nin belli varsayımlarının karşılanması gerekliliğidir. Bu varsayımların karşılanma gerekliliği

ise MTK'nin modellerine göre farklılık göstermektedir. Bu varsayımlar tek boyutluluk, yerel bağımsızlık, madde tepki fonksiyonlarının monotonluğu ve madde tepki fonksiyonlarının kesişmezliği.

### **Madde Tepki Kuramının Varsayımları**

Bireyin test performansını açıklamak için gerekli olan tek bir yetenek ya da özelliğin olduğu varsayılır. MTK bu tek bir gizil yeteneği *tek boyutluluk* olarak kabul eder. Ancak bu varsayımın her zaman kesin bir şekilde karşılanması mümkün değildir. Çünkü test performansına diğer bilişsel, kişisel ve daha birçok faktör etki edebilir. Bu nedenle bu varsayımın karşılanması için yeterli ve gerekli olan, test maddeleri ile ölçülen ve test performansını etkileyen baskın bir bileşenin ya da faktörün olmasıdır. Bu baskın bileşen ya da faktör test tarafından ölçülen yetenek olarak adlandırılır (Hambleton ve Swaminathan, 1985; Crocker ve Algina, 1986).

Genel olarak, herhangi bir MTK modeli kullanılacağı zaman, kullanılan testin tek boyutlu bir yeteneği ya da özelliği ölçüp ölçmediğinin mutlaka kontrol edilmesi gerekir ve sonuçta ölçülen baskın olan tek bir yetenek türünü ya da özelliği ölçtüğünün ya da ölçmediğinin belirlenmesi gerekir. Ancak bireyin test performansını ortaya koymak için tek bir yetenekten daha fazla yeteneğin ölçülmesini gerektiren model varsayımları da vardır ki bu modeller çok boyutlu modeller olarak bilinir (Hambleton ve Swaminathan, 1985).

Tek boyutluluğun belirlenmesi için birçok metod önerilmektedir. Bunlardan bir tanesi test puanlarının koşullu dağılımını incelemektir (Hambleton ve Swaminathan, 1985). Test puanlarının her bir yetenek düzeyindeki dağılımı, test puanının o yetenek düzeyindeki koşullu dağılımı olarak bilinir. Tek boyutlu bir test, bütün bir popülasyonun alt popülasyonlarına uygulandığı zaman test sonuçlarından elde edilen koşullu dağılımların birbirine benzer olması beklenir. Eğer alt popülasyonlara ait dağılımlar birbirinden farklıysa o zaman ölçülen tek boyutlu olduğundan söz edilemez (Hambleton ve Swaminathan, 1985).

Tek boyutluluk varsayımının karşılanması için, testi oluşturan maddelerin doğru cevap olasılıklarının belirlenmesinde iki veya daha fazla yetenek boyutu etkili olmamalıdır. Maddenin cevaplanma olasılığında eğer birden fazla özellik etkinse, tek boyutluluk varsayımı karşılanamayacaktır. Tek boyutluluk varsayımı MTK

modellerini büyük ölçüde sınırlar, çünkü saf tek boyutluluğun pratik durumlarda sağlanması çok zordur. Bireylerin maddelere tepki verirken tek bir boyut doğrultusunda hareket etmeleri beklenemez; dolayısıyla ölçülmek istenen değişken kompleks bir yapıya sahip olabilir (Sünbül,2011).

MTK'nin ikinci varsayımı olan yerel bağımsızlık, bireylerin bir testin farklı maddelerine verdikleri cevapların istatistiksel olarak birbirinden bağımsız veya ilişkisiz olmasıdır (Hambleton ve Swaminathan, 1985). Ancak maddelere verilen tepkilerin istatistiksel olarak birbirinden bağımsız olması için test maddeleri tarafından ölçülen yeteneğin sabit tutulması gerekir (Lord ve Novick, 1968). Bu varsayımın gerçekleşmesi için bireyin bir test maddesine verdiği cevabın, diğer maddeye vereceği cevabı etkilememesi gerekir. Bir başka ifadeyle bir sorunun cevabının diğer sorunun cevabı için herhangi bir ipucu niteliği taşıması gerekir (Hambleton ve Swaminathan, 1985).

Yerel bağımsızlık varsayımının sağlanabilmesi için ölçülen yapının tek boyutlu olması yeterlidir. Bu noktadan hareketle, MTK'nin iki varsayımı arasında bir ilişki kurulabilir. Şöyle ki: Test maddeleri ortak bir yapıyı ölçerken, bir başka ifadeyle test maddeleri arasında yüksek düzeyde pozitif yönlü bir ilişki varken, test maddelerine verilen cevaplar istatistiksel olarak birbirinden bağımsız ve ilişkisizdir (Crocker ve Algina, 1986; Hambleton ve Swaminathan, 1985). Sünbül'ün de (2011) ifade ettiği gibi, bu durum ilk bakışta bir çelişki yaratıyor gibi görünmesine rağmen gerçekte öyle değildir. Tek boyutluluk varsayımında maddelerin gösterdiği ilişkiler incelenirken herhangi bir yetenek kısıtlaması veya koşulu yoktur. Yerel bağımsızlık varsayımında ise, maddeler arasındaki ilişki veya bağımsızlık belirli bir yetenek koşulu altında incelenmektedir. Eğer maddeye verilecek olan tepki birden fazla yetenekten kaynaklanıyorsa, maddeleri bağımsız yapmak için süreçte etkili olan yeteneklerden birini sabitlemek yeterli olmayacaktır ve sabitlenmemiş olan diğer yetenek, maddeler arasında bağımlılık veya ilişki oluşturacaktır.

Yerel bağımsızlığın sağlanamaması, özellikle madde ve yetenek parametrelerinin kestirilmesinde kullanılan maksimum olabilirlik (maximum likelihood) fonksiyonunun temelini oluşturan olabilirlik (likelihood) fonksiyonlarından hesaplanacak olan olasılıkları tehlikeye sokacaktır. Çünkü olabilirlik fonksiyonları,

belirli bir yetenek koşulu altında, maddelere bağımsızmış gibi davranarak olasılık sonuçları üretmektedir (Sünbül,2011).

MTK modellerinin hepsi için gerekli olmayan ancak bazı modellerin kullanılması için karşılanması gereken iki varsayım vardır: Bunlar MTF'lerin monotonluğu ve MTF'lerin kesişmezliği varsayımlarıdır. Bu iki varsayımı gerektiren modeller, MTK modelleri başlığı altında detaylı olarak açıklanmıştır.

MTF'lerin monotonluğu varsayımı  $P_j(\theta)$  koşullu olasılığının  $\theta$  üzerinde monoton olarak azalmadığını ifade eder (Sijtsma ve Molenaar, 2002). Monotonluk varsayımı (2.1.) denklemi ile gösterilmiştir.

$$P(X_i=0 | \theta) = 1 - P(X_i=1 | \theta) \quad (2.1)$$

Bütün  $\theta$  çiftleri için yazılan ve  $\theta_a < \theta_b$  ve  $j=1,2,3,\dots j$  olmak üzere,  $P_j(\theta_a) \leq P_j(\theta_b)$  ile ifade edilen ilişkinin olması,  $P_j(\theta)$  ve  $\theta$  arasındaki ilişkinin belirli sınırlamalar tarafından yönetildiği varsayımının bir sonucudur. Bu eşitsizlik  $\theta$ 'nın azalmayan bir fonksiyonunun olduğunu ifade etmektedir.  $P_j(\theta)$  ve  $\theta$  arasındaki ilişki monoton ve azalmayan bir ilişki olduğu sürece, parametrik olsun ya da olmasın herhangi bir fonksiyonel forma izin verilir (Sijtsma ve Molenaar, 2002).

Son olarak MTF'lerin kesişmezliği varsayımı ise, k tane madde karakteristik eğrisinin ilgili  $\theta$  değerleri üzerinde kesişmemesi gerekliliğini ifade eder. Bu varsayım maddelerden ziyade bireylere ait ölçümlerin yapıldığı durumda test edilmesi gereken bir varsayımdır (Sijtsma ve Molenaar, 2002). MTF'lerin kesişmezliği varsayımı bütün  $\theta$  değerleri için denklem (2.2) ile ifade edilebilir:

$$P_1(\theta) \leq P_2(\theta) \leq \dots \leq P_k(\theta) \quad (2.2)$$

### **Madde Tepki Kuramı Modelleri**

MTK modelleri ile ilgili günümüze kadar birçok sınıflandırma yapılmıştır. Bu sınıflandırmalar modellerin boyutluluğu, puanlama, doğrusallık yapısı... vb. gibi özelliklere göre yapılmıştır.

Sijtsma ve Molenaar (2002), MTK modellerini parametrik ve parametrik olmayan modeller olarak sınıflandırmıştır. Bu sınıflandırmanın detaylandırılması

şöyledir:

1. Parametrik MTK Modelleri (PMTK)
  - a) Tek boyutlu modeller
    - i. Bir parametrelili lojistik model
    - ii. İki parametrelili lojistik model
    - iii. Üç parametrelili lojistik model
    - iv. Dört parametrelili lojistik model
  - b) Çok boyutlu modeller
    - i. Telafisel model
    - ii. Telafisel olmayan model
2. Parametrik Olmayan MTK Modelleri (POMTK)
  - a) Monoton homojenlik modeli
  - b) İkili monotonluk modeli

Sijtsma ve Meijer (2007)'ye göre MTK modellerinin parametrik ve parametrik olmayan şeklinde sınıflandırılmasının nedeni, öncelikli olarak tepki fonksiyonları üzerindeki sınırlılıklar ve bir modelin belirli bir parametrik fonksiyonu varsayması ya da varsaymamasıdır (lojistik ya da normal ogive gibi). Genel olarak MTF,  $\theta$ 'nın örtük özellik değeri ne kadar fazla olursa,  $\theta$  değerini ölçen bir maddeye doğru cevap verilme olasılığı da yüksektir fikrini desteklemektedir. Dolayısıyla MTF,  $\theta$ 'nın artan bir fonksiyondur.  $X_j$  rastgele seçilen bir  $j$  maddesine ait madde puanını göstermek ve bu puan yanlış cevap için 0 ve doğru cevap için 1 olmak üzere tanımlanan MTF denklem (2.3) ile ifade edilmiştir (Sijtsma,2005):

$$P_j(\theta)=P(X_j=1 | \theta) \quad (2. 3)$$

MTF için en uygun matematiksel fonksiyon lojistik fonksiyondur. Bu fonksiyon  $\theta$  değerinde ilk olarak yavaşça ardından hızlıca ve hemen ardından yeniden yavaşça monoton bir şekilde yükselir. Bu fonksiyon sağa doğru itilmiş bir "S" şeklinde görülür ve dolayısıyla doğru cevaplama olasılıkları yorumlanabilir. Bu fonksiyona ilişkin sorunlardan biri doğru cevabı bulma olasılığı yalnızca  $\theta$ 'ya bağlıdır ve bu olasılık başka bir duruma bağlı değildir. Ancak şurası bir gerçektir ki bazı maddeler diğerlerinden daha zordur ve bu durum doğru cevaplama olasılığını etkiler. Bu sorunu gidermek için öncelikle  $\theta$  ile bir maddenin yer parametresi ( $\delta_j$ )

arasındaki farka bağılı bir lojistik fonksiyon tanımlanır ve böylece bir MTF elde edilir (Sijtsma ve Molenaar, 2002). Elde edilen MTF (2.4) denklemi ile ifade edilmiştir:

$$P_j(\theta) = \frac{\exp[(\theta - \delta_j)]}{1 + \exp[(\theta - \delta_j)]} \quad (2.4)$$

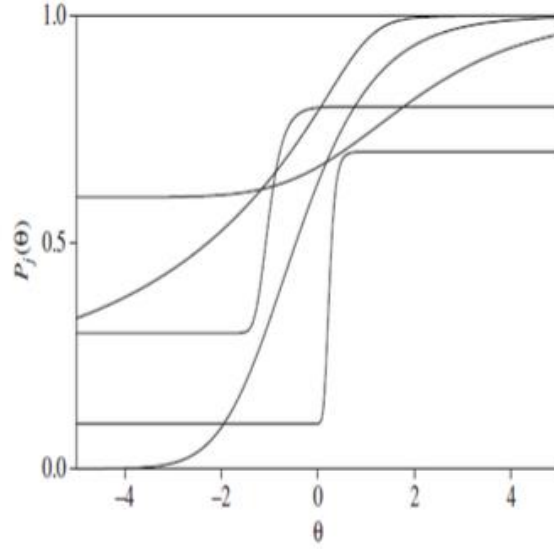
POMTK ve PMTK modelleri arasındaki önemli farklılıklardan biri sahip oldukları MTF'lerle ilgilidir. Örneğin POMTK modeli tipik olarak ikili madde puanı ile gizil değişken arasında pozitif bir monoton ilişki olduğunu varsayar. Bir başka ifadeyle bütün  $\theta$  çiftleri için  $\theta_a < \theta_b$  ve  $j=1,2,3,\dots j$  olmak üzere,

$$P_j(\theta_a) \leq P_j(\theta_b) \quad (2.5)$$

olduğu varsayılır. Ancak 2PLM gibi PMTK modelleri bu varsayıma ek olarak  $\delta_j$  gibi bir yer parametresi ve  $\alpha_j$  gibi bir eğim parametresiyle birlikte bu ilişkinin lojistik olmasını gerektirir. POMTK'de ise böyle bir zorunluluk yoktur. Buna göre, her bir madde için  $P_j(\theta)$  (2.6) denklemi ile ifade edilmiştir:

$$P_j(\theta) = \frac{\exp[\alpha_j(\theta - \delta_j)]}{1 + \exp[\alpha_j(\theta - \delta_j)]} \quad (2.6)$$

(2.4) ve (2.6) denklemleri parametrik fonksiyonları göstermektedir, çünkü bu fonksiyonlar parametrik bir yöntemle  $P_j(\theta)$  ve  $\theta$  arasındaki ilişkiyi belirlerler. İlgili parametreler biliniyorsa fonksiyonel ilişki tamamen belirlenebilir. Denklem (2.5)'e ait MTF'ler ise tam olarak belirlenemez, yapılan tanım parametrik değildir. POMTK'ye göre bütün  $\theta$  çiftleri için yazılan ve (2.5) denklemi ile ifade edilen ilişkinin olması,  $P_j(\theta)$  ve  $\theta$  arasındaki ilişkinin belirli kısıtlamalar tarafından yönetildiği varsayımının bir sonucudur. Bu eşitsizlik  $\theta$ 'nın azalmayan bir fonksiyonunun olduğunu ifade etmektedir. Bu ifade bireylerin yetenek düzeyi arttıkça, ilgili maddeye doğru cevap verme olasılığının değişmemesi ya da artması şeklinde yorumlanır (Sijtsma,2005). Bu tür düzen kısıtlanmalı MTF'lerine dayanan modeller POMTK modelleri olarak incelenir. Denklemi (2.5)'i sağlayan MTF örnekleri şekil 1'de görülmektedir (Sijtsma ve Molenaar, 2002):



Şekil 1: Monoton azalmayan MTF örnekleri

Şekil 1'de (2.5) denkleminde verilen eşitsizliği sağlayan MTF örnekleri görülmektedir. Buna göre, MTF'lerin 2PLM'deki gibi lojistik fonksiyonlar olabileceği, ancak kısmen doğrusal veya üstel olabileceği, daha düşük (0'dan büyük) ve daha büyük (1'den küçük) asimptotlara sahip olabileceği ve daha düşük (0 dan büyük) ve daha büyük (1 den küçük) simetrik olması gereken asimptotlara sahip olabilecekleri yorumu yapılabilir. MTF'ler son derece düzensiz, ayrık veya basamaklı fonksiyonlar olabilirler.  $P_j(\theta)$  ve  $\theta$  arasındaki ilişki monoton ve azalmayan bir ilişki olduğu sürece, herhangi bir fonksiyonel forma izin verilir (Sijtsma ve Molenaar, 2002).

POMTK ve PMTK arasındaki bir diğer farklılık, POMTK öncelikli olarak,  $\theta$ 'nın boyutluluğunu açıklamayı amaçlar. Bir başka ifadeyle, POMTK bütün öngörülerden uzak bir şekilde  $\theta$ 'nın boyutluluğunu ortaya koymayı amaçlar. Bununla birlikte PMTK modelleri sıklıkla boyutluluğun belirli bir biçimi olduğunu varsayar ve daha sonra verinin boyutluluğu üzerine hipotezlenen modelin uyumunu test eder. Buna göre POMTK  $\theta$ 'nın boyutluluğunu açıklayıcı bir yaklaşımla ortaya koyarken, PMTK mevcut hipotezlerden hareket ederek doğrulayıcı bir yaklaşım sergiler (Sijtsma ve Meijer, 2007).

POMTK modelleri, PMTK'yi daha iyi anlaşılması açısından daha geniş bir teorik alanla içine alır. Bunun yanı sıra PMTK belirli bir modelin veriye uyum sağlaması yönünden doğrulayıcı araçlarla uğraşırken, POMTK ölçek verileri ve test



analizleri için öncelikli olarak açılımlayıcı araçlarla ilgilenir. Bu farklılıklara göre POMTK'nin PMTK'ye göre daha genel olduğu ve POMTK'nin veri analizinde açılımlayıcı, PMTK'nin ise doğrulayıcı olduğu söylenebilir (Sijtsma ve Meijer, 2007).

MTK modellerinin parametrik ve parametrik olmayan modeller olarak iki kategoride incelenmesinin ardından, parametrik modeller de kendi içinde iki grupta sınıflandırılmıştır. McDonald (1982) tarafından yapılan bu sınıflama için boyutluluk faktörü göz önünde bulundurulmuştur. MTK için boyutluluk kavramı, kullanılan modelden elde edilen sonuçların geçerliği için önemli bir faktördür. Daha önce belirtildiği gibi tek boyutlu MTK modelleri madde parametreleri ve teste giren kişinin yeteneğinin belirlenmesinde testlerin yalnızca tek bir boyutu ölçtüğünü varsayar. Ancak tek boyutlu MTK modellerinin kullanımından önce kullanılan verinin tek boyutlu olduğu kanıtlanmalıdır. Nitekim Gao (1997) güvenilir sonuçlar elde edebilmek için verilerin tek boyutluluğunu gerektiren psikometrik yöntemlerin uygulanmasından önce, tek boyutluluk varsayımının doğrulanması gerektiğini savunur. Aksi halde bu modellerin kullanılmasıyla elde edilen sonuçların geçerliliği tartışmalı olacaktır. Bu doğrulamanın yapılması durumunda tek boyutlu MTK modellerinin kullanılması doğru olur. Tek boyutluluk varsayımının karşılanmadığı durumlarda tek boyutlu MTK modelleri ile madde parametreleri ve teste giren kişiye ait beceri değerlendirmelerine ait ciddi sapmalar elde edilebilir (Ackerman, 1989).

Tek boyutlu MTK modelleri bireyin madde üzerindeki performansını etkileyen tek bir yetenek türü olduğunda kullanılan MTK modelleridir. Tek boyutlu MTK modelleri parametrik ve parametrik olmayan tek boyutlu MTK modelleri şeklinde sınıflandırılabilir.

Tek boyutlu parametrik MTK (TBPMTK) modelleri ikili puanlanan verilerde (1-0) kullanılan; bir parametrelili, iki parametrelili, üç parametrelili ve dört parametrelili lojistik model başta olmak üzere farklı türde modelleri kapsamaktadır. Tek boyutlu parametrik olmayan MTK (TBPOMTK) modelleri ise ilk olarak Mokken (1971) tarafından ortaya konmuştur. Bu modellerde madde karakteristik fonksiyonu ile ilgili matematiksel bir form tanımlanmamaktadır. Bu durumda her bir madde gelişigüzel bir madde karakteristik fonksiyonuna sahip olabilmektedir ve parametrelerdeki değişimlerin matematiksel formüllerle verilmesi gibi bir gereklilik yoktur. Bu modelin ortaya konmasının literatüre yaptığı en önemli katkı gizil değişkenlere ilişkin

parametre dağılımını sınırlayan bir varsayımın olmayışıdır (Sijstma ve Molenaar, 2002).

TBPOMTK modellerinin TBPMTK modellerine göre sahip olduğu en büyük avantaj ilgilendiği çok az varsayımın olmasıdır. Bu durum TBPOMTK modellerini oldukça kullanışlı bir istatistiksel model konumuna getirmektedir. TBPOMTK'nin sahip olduğu algoritmalar, TBPMTK'ye göre daha basittir (Sijstma ve Molenaar, 2002).

TBPOMTK modellerinin varsayımlar açısından daha esnek olması, gizil değişken ile cevap kategorisine verilen cevaba ait madde karakteristik fonksiyonunun parametrik olarak tanımlanmamasının bir sonucudur. TBPOMTK modelleri sıralama düzeyindeki ölçmelerdir ve klasik testler ile anketlerdeki her bir katılımcıyı değerlendirme için oldukça uygundur (Stochl, 2007).

Bu çalışmada, amacı bakımından iki kategorili veriler kullanıldığı için detaylı olarak incelenen modeller de iki kategorili veri analizinde kullanılan modellerdir. Bu modeller TBPMTK, TBPOMTK ve ÇBMTK çerçevesinde incelenen iki kategorili modellerdir.

**TBPMTK modelleri.** TBPMTK, bir kişinin test maddeleriyle olan etkileşimlerini, kişinin örtük özelliklerini tanımlayan tek bir parametreyi içeren bir matematiksel ifade ile yeterli ölçüde temsil edilebileceğinin öngörüsüne dayalı olarak bir dizi model içermektedir. TBPMTK modelinin temel gösterimi (2.7) 'de verilmiştir.

$$P(U = u | \theta) = f(\theta, \eta, u) \quad (2.7)$$

Bu denklemde,  $\theta$  kişinin örtük özelliğini ya da yetenek düzeyini,  $\eta$  test maddesinin özelliklerini tanımlayan parametrelerin bir vektörünü,  $U$  test maddesindeki puanı ve  $u$  bir test maddesinden alınabilecek bir puan için olası bir değeri ve son olarak  $F$ , parametreler ile yanıt olasılıkları arasındaki ilişki fonksiyonunu göstermektedir (Reckase, 2009).

TBPMTK modellerinin matematiksel olarak en temel gösterimi (2.7)'deki gibidir. Özellikle iki kategorili bir başka ifadeyle ikili puanlanan maddeler için kullanılan TBPMTK modelleri bu ifadenin genişletilmesi ile elde edilmiştir.

TBPMTK modelleri doğru/yanlış şeklinde puanlanan maddeler için sıklıkla uygulanır. Bu maddelerden doğru olanları genel olarak “1” ile yanlış olanları ise “0” ile kodlanır. İki kategorili olan bu maddeler arasında doğru olanların yanlış olanlara göre daha yüksek bir yetenek düzeyini gösterdiği varsayılır. Bu yüzden monoton azalmayan matematiksel fonksiyonlar, kişiler ve maddeler arasındaki etkileşimi modellemek için kullanılabilirler (Reckase, 2009).

Bu çalışmada kullanılan iki kategorili modeller sahip oldukları parametre sayısına göre isimlendirilen modellerdir. Bu modeller 1 parametrelili lojistik model (1PLM), 2 parametrelili lojistik model (2PLM), 3 parametrelili lojistik model (3PLM) ve 4 parametrelili lojistik modellerdir (4PLM). Bu çalışmada TBPMTK modellerinden 2PLM kullanıldığı için sadece 2PLM hakkında kuramsal bilgiye yer verilmiştir.

**2 parametrelili lojistik model.** Birnbaum (1968) tarafından önerilen bu model 2 parametreye sahiptir. Bu parametreler kişinin örtük özelliği ile test maddelerinin özelliğini tanımlamak için kullanılır ve (2.8) ile ifade edilir (Reckase, 2009).

$$P(U_{ij} = u_{ij} | \theta_j) = f(\theta_j, b_i, u_{ij}) \quad (2.8)$$

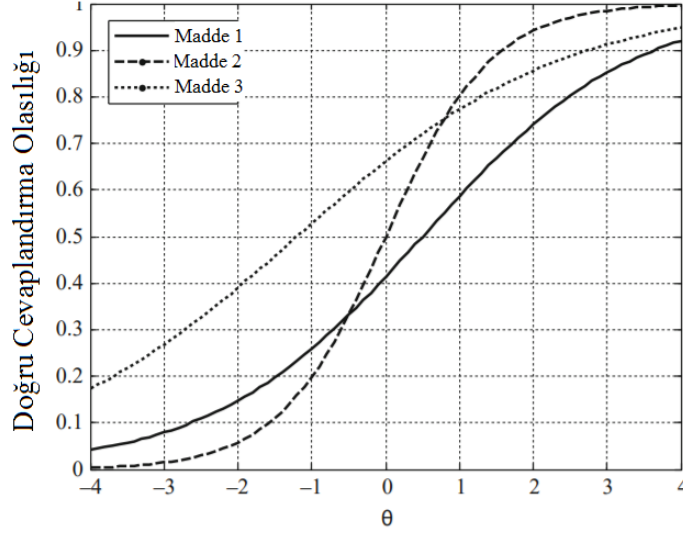
Verilen (2.8) denklemi ile herhangi bir j kişinin i maddesi üzerindeki puanının  $u_{ij}$  olduğu yerde,  $\theta_j$  parametresi j. kişinin ilgili olan özelliğini gösterir ki bu özellik genelde yetenek ya da bireyin i maddesi üzerindeki performansını etkileyen örtük özellik olarak düşünülür. Bununla birlikte  $b_i$  genelde madde güçlüğünü gösterir (Reckase, 2009). Çoğu iki kategorili puanlanan bilişsel test maddeleri için, parametrelerle doğru cevaplama olasılığı arasında monotonluk varsayımının karşılanmasını gerektiren bir yolla ilişki kuran bir fonksiyona ihtiyaç vardır. Bir başka ifade ile,  $\theta_j$  arttıkça modelin fonksiyonel formu doğru cevaplama olasılığının artmasını belirler. Bu model (2.9) ile gösterilmiştir (Reckase, 2009).

$$P(u_{ij} = 1 | A_j, B_i) = \frac{A_j B_i}{1 + A_j B_i} \quad (2.9)$$

Eşitlikte verilen  $A_j$ , kişi parametresidir ve genel olarak  $\theta_j$  olarak adlandırılır.  $B_i$  ise madde parametresidir ve  $b_i$  olarak adlandırılır. (2.9) eşitliği için gerekli logaritmik dönüşümler yapıldıktan sonra (2.10) denklemi elde edilir (Reckase, 2009).

$$P(u_{ij} = 1 \mid \theta_j, a_i, b_i) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}} \quad (2.10)$$

2PLM'nin birtakım özellikleri sayesinde basit çıkarımların yapılması mümkündür. Bunu iyi anlayabilmek için şekil 2'de verilen madde karakteristik eğrisinin (MKE) incelenmesi gerekmektedir (Reckase, 2009).



Şekil 2. 2 PLM'ye ait MKE

Şekil 2'de  $\theta$ 'nın bazı değerleri için eğrinin daha dik, bazıları için daha eğik olduğu görülmektedir. Grafikteki eğim görece olarak daha dik olduğu zaman maddenin doğru cevaplanma olasılığının bireyler arasında çeşitlendiği anlamına gelir. Bu durum ilgili maddenin ilgili yetenek düzeyindeki bireyleri maddeye doğru cevap verme olasılıkları bakımından birbirinden ayırmada başarılı olduğunu gösterir. Bir başka ifadeyle maddenin ayırt edicilik gücünün fazla olduğunu gösterir. Bir maddenin ayırt edicilik gücü arttıkça o maddeye ait MKE dikleşir, ayırt etme gücü azaldıkça eğri düzleşir.

Herhangi bir  $\theta$  düzeyinde maddeye doğru cevap verme olasılığı 0,5 olduğu zaman MKE en dik şeklini alır. Bir başka ifadeyle,  $\theta$  doğru cevaplama olasılığının 0,5 olduğu durumdaki değerine eşit olduğu zaman MKE en dik şeklini alır. Doğru cevaplama olasılığının 0,5 olması ise  $\theta = b_i$  durumunda mümkün olur. Böylece  $b$  parametresi  $\theta$  üzerinde ayırt ediciliği en yüksek olan noktayı işaret eder. Yüksek  $b$  parametresine sahip olan maddeler yüksek yetenek düzeylerindeki bireyler için en

yüksek ayırt ediciliğe sahip olan maddelerdir. Düşük  $b$  parametresine sahip olan maddeler ise, düşük yetenek düzeyindeki bireyler için en fazla ayırt ediciliği olan maddelerdir.  $b$  parametresi aynı zamanda bireylerin bir maddeyi doğru ya da yanlış cevaplandırmaları için potansiyel yetenek düzeylerine ait yetenek ranjı hakkında bilgi verir.  $b_i$ 'den daha yüksek bir  $\theta_j$ 'ya sahip olan bireylerin  $i$  maddesini doğru cevaplandırma olasılığı 0,5'ten büyüktür.  $b_i$ 'den daha düşük bir  $\theta_j$ 'ya sahip olan bireylerin  $i$  maddesini doğru cevaplandırma olasılığı ise 0,5'ten küçüktür.  $b$  parametresi herhangi bir  $i$  maddesi tarafından en iyi şekilde ölçülen yetenek düzeyi ve maddenin muhtemel doğru cevaplandırma olasılığı hakkında bilgi verdiği için  $b$  parametresi *güçlük parametresi* olarak adlandırılır (Reckase, 2009).

Şekil 2'de MKE'lerin 0,5 olasılık çizgisinde paralel olmadıkları ve eğrilerin kesiştiği görülmektedir. Kesişen eğriler bir yetenek düzeyindeki bireyler için bir test maddesinin diğer test maddesine göre daha zor olduğu anlamına gelir. Ancak aynı madde, diğer bir yetenek düzeyindeki bireyler için daha kolay olabilir.

2 PLM'deki  $a$  parametresi, her bir birey için  $\theta$  parametresinin tahmin edilmesine izin verir. 2 PLM için aynı ağırlıklandırılmış puana sahip olan bireyler  $\theta$ 'nin aynı maksimum olabilirlik ihtimaline sahiptir (Reckase, 2009).

**TBPOMTK modelleri.** POMTK her ne kadar öncelikli olarak açıklayıcı araçlarla ilgilenirse de, veri matrisinin yapısı ile ilgili öngörüler POMTK modellerinden de türetilir. Öngörülere dayalı olarak tahmin edilen bu veri yapısı gerçek verilerle karşılaştırılabilir. Eğer gözlenen ve beklenen yapılar eşleşiyorsa, örneklem özelliklerinden kaynaklanan farklılıklar dışında, model veriyi açıklar ya da veriye uyum sağlar ve kullanılan testin ya da ölçeğin öngörülen ölçek özelliklerini (sıralama ölçeği, aralık ölçeği gibi) karşıladığı varsayılır. Örneklem özelliklerinden kaynaklanan veri yapıları arasındaki farklar çok fazla ise, model istatistiksel standartlara göre verileri açıklayamaz (Sijtsma ve Meijer, 2007).

Modelin gözlenen veriyi açıklama derecesi maddelerin cevaplandırma süreci ile ilgili varsayımların karşılanmasına bağlıdır. Bu varsayımlar detaylı bir şekilde daha önce açıklanmıştı. Karşılanması gereken varsayımlara ve kullanım amacına göre farklılaşan iki POMTK modeli ikili monotonluk modeli ve monoton homojenlik modelidir.

Monoton homojenlik modeli (MHM), tek boyutlu olma, yerel bağımsızlık ve monotonluk varsayımlarına dayanmaktadır. MHM, örtük özellik ile monoton olarak ilişkili olan MTF'lere sahip, homojen (tek boyutlu) bir madde seti tarafından üretilen madde-tepki örüntülerini açıklar. MHM'nin pratikteki önemi, bireylerin  $\theta$  ölçeğinde  $X_+$  vasıtasıyla sıralanmalarını gerektirmesidir. Dolayısıyla MHM, bireyleri sıralı bir ölçekte ölçmek için kullanılan bir MTK modelidir. Teknik olarak, kişileri sıralamak şu anlama gelir:  $c$ ,  $\theta$  ölçeğinde rastgele bir sabit ve  $X_+$ 'in iki sabit değeri  $s$  ve  $t$  ile gösterilmiş olsun. Burada  $k$  toplamda elde edilen puan sayısını göstermek üzere;  $0 \leq s < t \leq k$  koşulu için, MHM (2.11) eşitliğini ifade eder:

$$P(\theta > c|X_+ = s) \leq P(\theta > c|X_+ = t) \quad (2.11)$$

Burada (2.11) eşitliği bireylerin  $\theta$ 'ya göre  $X_+$ 'ya bağlı olarak rastgele sıralanabileceğini gösterir. (2.11) eşitliğinden yola çıkarak (2.12) eşitliğine ulaşılır:

$$E(\theta|X_+ = s) \leq E(\theta|X_+ = t) \quad (2.12)$$

(2.12) eşitliği, toplam puan ortalaması  $t$  olan bir grupta,  $\theta$ 'nın en azından  $s$  gibi düşük bir toplam puana sahip bir grupta olduğu kadar yüksek bir yetenek düzeyi olduğunu ifade eder. POMTK modelinden beklenen sıralama tam olarak (2.12) eşitliği ile ifade edilir.

PMTK modellerinde olduğu gibi, MHM'nin uygulamadaki önemi  $\theta$ 'ların sayısal tahminlerinin elde edilemediği durumlarda,  $X_+$  üzerindeki sıralamaları kullanılarak  $\theta$ 'lara ait sıralama sonuçlarının elde edilmesidir. Bu ifade, MHM'nin bireylere yönelik bir ölçüm modeli olarak kullanımını haklı çıkaran en iyi kilit sonuçtur.

Diğer POMTK modeli ikili monotonluk modelidir (IMM). IMM, tek boyutlu olma, yerel bağımsızlık, monotonluk ve kesişmezlik üzerine kuruludur. IMM, ilk üç varsayımı MHM ile paylaştığı ve dördüncü bir varsayımı eklediği için MHM'nin özel bir hali olarak düşünülmektedir. Bu, IMM tarafından açıklanabilen her veri setinin daha zayıf bir MHM tarafından açıklanabileceği anlamına gelir, ancak tersi doğru değildir (Sijtsma ve Meijer, 2007).

**ÇBMTK Modelleri.** ÇBMTK modellerinin iki önemli türü vardır. Bu iki çeşit, maddenin doğru cevaplama olasılığını belirlemek için,  $\theta$  koordinatlarına ait vektörden gelen bilgi ile madde karakteristik eğrisinin birleştirilmesi yoluyla tanımlanır. Bu türlerden bir tanesi  $\theta$  koordinatlarının lineer kombinasyonuna dayanır. Bu lineer birleşim, maddenin doğru cevaplama olasılığını belirlemek amacıyla normal ogive ya da lojistik form ile birlikte kullanılır.  $\theta$  koordinatlarının lineer birleşimi  $\theta$  değerlerinin çeşitli bileşimleri ile birlikte aynı toplam puanın elde edilmesini sağlar. Eğer bir  $\theta$  koordinatı düşükse, diğer bir  $\theta$  koordinatının yeterince yüksek olmasına bağlı olarak toplam  $\theta$  değeri aynı kalır. Bundan dolayı bu tür modeller telafi etme özelliğine sahiptir ve bu özelliğe sahip olan modeller tipik olarak *telafisel modeller* olarak adlandırılır (Reckase, 2009). Bir yeteneğin yokluğunu, başka bir yeteneğin telafi edebildiği durumlarda kullanılması önerilen modeller telafisel modellerdir.

İkinci model türü ise bir test maddesinde yer alan bilişsel görevleri kısımlara ayırır ve her bir kısım için tek boyutlu model kullanılır. Bir maddenin doğru cevaplanma olasılığı her bir kısım için elde edilen olasılıkların bileşkesidir. Olasılıkların bileşkesinin kullanılması, bu modeller için doğrusal olmayan özelliklerin olmasına neden olur. Aynı zamanda doğru cevaplama olasılığı bileşke içinde yer alan en yüksek olasılığı aşamayacağı için, düşük  $\theta$  koordinatları için yüksek  $\theta$  koordinatının telafiselliği azalmaktadır. Bu nedenle bu modeller sıklıkla *telafisel olmayan modeller* olarak adlandırılırlar. Ancak kısmi telafisel modeller olarak da isimlendirilirler. Çünkü bir boyutta yer alan yüksek bir  $\theta$  koordinatı, aynı boyutta yer alan düşük  $\theta$  koordinatına göre daha yüksek bir doğru cevaplama olasılığı sağlayabilir. Bir başka ifadeyle tam olarak telafisel modellerde olduğu gibi olmasa da bazı telafisellikler olabilir (Reckase, 2009).

Tıpkı TBPMTK'de olduğu gibi ÇBMTK'de telafisel ve kısmi telafisel modeller de kendi içlerinde birçok modeli kapsamaktadır. Bu modeller kişi ve madde arasındaki etkileşimi mantıklı ve doğru bir şekilde belli matematiksel ifadelerle modelleyebilen temel öncüllere sahiptirler. ÇBMTK modelleri için çeşitli matematiksel ifadeler geliştirilmiştir.

ÇBMTK modelleri varsayımlar açısından değerlendirilecek olursa, MTK için önemli olan varsayımlardan sadece monotonluk varsayımının sağlanması gerektiği

söylenbilir. MTK için önemli olan varsayımlardan tek boyutluluk varsayımının sağlanamamasından dolayı zaten çok boyutlu modeller ortaya çıkmıştır. MTK için ikinci önemli varsayım olan yerel bağımsızlık varsayımı ise Reckase (2009)'a göre test edilmesi ve sağlanması gereken bir varsayımken, Crocker ve Algina (1986)'ya göre tek boyutluluğun bozulduğu durumda yerel bağımsızlık da bozulacaktır. Bu nedenle tek boyutluluğun sağlanmadığı durumda yerel bağımsızlığın sağlanmasını beklememek gerekir.

Bu çalışmada kullanılan veriler TEOG uygulamalarından elde edildiği için TEOG'un doğasına uygun olan ÇBMTK modelinin seçilmesi gerekir. TEOG uygulamasında yer alan alt maddeler incelendiğinde, maddeler arasındaki telafiselliği görmek mümkündür. Bu nedenle kullanılan veri setinin doğasına uygun olarak bu çalışmada, iki kategorili puanlanan telafisel ÇBMTK modellerine yer verilmiştir. Bu modeller, psikometri literatüründe görece olarak uzun bir geçmişe sahip oldukları için, uygulama noktasında önemli tecrübeler edinilmiştir. ÇBMTK modellerinin gelişim sürecindeki bazı uyarıcılar, iki kategorili verilere uygulanan faktör analizinde ortaya çıkan bazı problemleri belirlemek amacıyla yapılan çalışmalar sonucunda elde edilmiştir. Bu nedenle iki kategorili ÇBMTK modelleri ile 1980'li yıllardan beri literatürde çalışılmaktadır. Bu ilk çalışmaların öneminden dolayı iki kategorili ÇBMTK modelleri TBPMTK modeli olan 2PLM'nin çok boyutlu modellere genişletilmesiyle başlar. Bu model ÇBMTK modelinin en temel mantığını gösterir. Sonrasında ise aynı temel forma sahip modeller geliştirilmiş ve sunulmuştur. Bu modeller biraz daha az ya da biraz daha fazla parametre sayısına sahip olmakla birlikte, normal ogive formu değil lojistik formu içeren modellerdir.

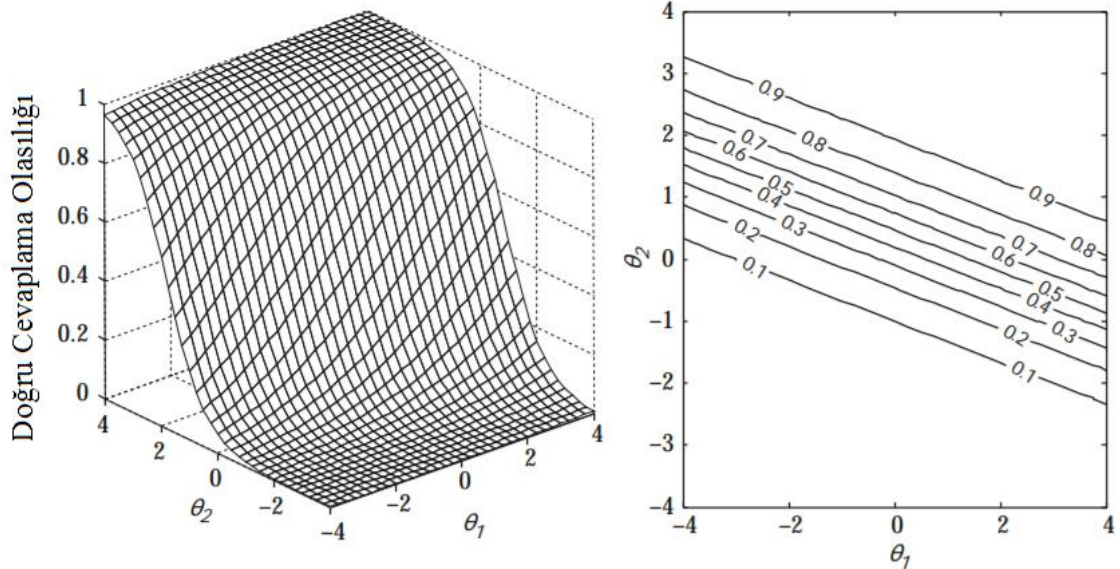
**Genişletilmiş 2 paramereli lojistik model (M2PLM).** 2 PLM'nin denkleminde hatırlanacağı üzere  $a(\theta-b)$  ifadesi yer almaktadır. 2PLM'nin M2PLM'ye genişletilmesi için  $a$ 'nın paranteze dağıtılması gerekir. Bu işlemden sonra elde edilen  $a\theta-ab$  ifadesinde  $-ab$  yerine  $d$  parametresinin gelmesiyle oluşan  $a\theta+d$  ifadesi eğim/kesişim parametresi olarak adlandırılır (Reckase 2009). Burada  $a$  parametresi eğim parametresi olup  $d$  parametresi kesişim parametresi olarak adlandırılmıştır.  $U_{ij}$  j bireyinin i maddesine verdiği cevabı,  $\theta_j$  j kişinin yetenek düzeyini,  $a_i$  i maddesinin eğim parametresini ve  $d_i$  i maddesinin kesişim parametresini göstermek üzere 2 PLM'nin en genel şekliyle ifadesi (2.13)'de görülmektedir.



$$P(u_{ij} = 1 \mid \theta_j, a_i, d_i) = \frac{e^{(a_i\theta_j+di)}}{1+e^{(a_i\theta_j+di)}} \quad (2.13)$$

M2PLM ile 2 PLM arasındaki en temel farklardan biri d parametresidir. d parametresi maddenin güçlüğünü gösteren önemli bir betimleyici istatistiktir. Genişletilmiş 2 PLM'deki d parametresi ile 2PLM'de bulunan b parametresi benzer şekilde yorumlanırlar (Reckase 2009). İki model arasındaki bir diğer önemli fark ise tanımlanan M2PLM'de her boyut için ayrı bir eğim parametresi hesaplanmaktadır. Örnek olarak, iki boyutlu bir maddenin  $a_1$  ve  $a_2$  olmak üzere iki eğim parametresi bulunmaktadır.

TBPMTK modellerinde birey ile madde etkileşimi arasındaki matematiksel ilişkiyi gösteren madde karakteristik eğrilerinden yararlanıldığı daha önce belirtilmişti. ÇBMTK kapsamında ise bu ilişkiyi göstermek için *madde karakteristik yüzeylerinden* yararlanılır. 2 boyutlu bir maddenin, maddeye ait iki boyutta yeteneği tanımlanmış bir bireyin maddeyi doğru cevaplama olasılığını gösteren madde karakteristik yüzeyi şekil 3'te görülmektedir (Reckase 2009). Madde karakteristik yüzeyi verilen iki boyutlu bir maddenin birinci boyuttaki eğim parametresinin  $a_1=0,5$ , ikinci boyuttaki eğim parametresinin  $a_2=1,5$  ve kesişim parametresinin ise  $d=0,7$  olduğunu belirtmekte yarar vardır.



Şekil 3. M2PLM'ye ait madde karakteristik yüzeyi

### Madde Tepki Kuramında Güvenirlik

MTK'daki puan güvenilirliği, bir katsayıdan ziyade  $\theta$ 'nın bir fonksiyonu olarak hesaplanabilir. Bulut (2013)'ün aktarımlarına göre, klasik test kuramındaki test güvenilirliği yaklaşımını takiben Wang ve arkadaşları (2004) MTK temelli güvenilirlik için bir metod belirlemişlerdir. Bu MTK temelli test güvenilirliği marginal güvenilirlik olarak adlandırılır. Tek bir yeteneği ölçtüğü düşünülen her bir alt testin MTK'ye göre ölçme kesinliğinin belirlenmesinde marginal güvenilirlikten yararlanılabilir. TBPMTK modellerinin kullanıldığı durumlarda marginal güvenilirlik sıklıkla kullanılan bir güvenilirlik kestirme metodudur.

Her ne kadar MTK'de güvenilirlik kestirimi bir katsayıdan ziyade  $\theta$ 'nın bir fonksiyonu olarak hesaplanıyorsa da genel olarak güvenilirlik, maddenin ve o maddelerden oluşan testin bütünüdürün ayırtedicilik düzeyi ile doğrudan ilişkili olan bir kavramdır. Herhangi bir test maddesi ölçülen özelliğe sahip olan öğrenci ile sahip olmayan öğrenciyi birbirinden ayırabildiği ölçüde ayırtedicidir. Böyle maddelerden oluşan bir testten alınan puanlardaki çeşitlilik fazladır dolayısıyla puanlara ait varyans oranları yüksek olur. Bu durum testin güvenilirliğini artırıcı yönde bir etki yapar. Parametrik olmayan modellerde maddelere ait ayırtedicilik düzeylerinin ağırlıklarına göre hesaplanan bir güvenilirlik kestirme yöntemi kullanılır. TBPMOTK için madde ayırt edicilik katsayısı olarak  $H_i$  katsayılarından yararlanılır. Bu katsayı aynı zamanda madde geçerlik katsayısıdır. Bir de tüm test için hesaplanan  $H$  ile gösterilen ölçeklenebilirlik katsayılarından yararlanılır ki bu katsayı aynı zamanda tüm teste ait güvenilirlik katsayısını verir. Buradaki  $H$  katsayısında hesaba katılan hata türü Guttman hatasıdır. Bu nedenle  $H$  katsayısının hesaplandığı güvenilirlik belirleme yöntemine Guttman'ın homojenlik modeli adı verilir.  $H$  katsayısı testin iç tutarlılığın bir ölçüsü olarak kullanılır (Loevinger, 1948, Mokken 1971; Van Nearssen, 1972; Akt: Molenaar ve Sijtsma, 1984).

Tek boyutlu modellerde güvenilirliğin belirlenmesi için  $\theta$ 'nın fonksiyonları da, madde ayırtedicilik katsayıları da güvenilirliğin belirlenmesinde göz önünde bulundurulur. Ancak çok boyutlu modellerde göz önünde bulundurulması gereken bir özellik daha vardır ki o da alt testlere ait puanlar arasındaki değişimdir.

Akademik başarı gibi ölçülen özelliklerle ilgili belirli bir gruptaki bireylerin güçlü ve zayıf yönlerini belirlemek amaçlandığında ve bu amaçla çok boyutlu bir test uygulandığında, bireylerin alt testlerden aldıkları alt puanlar arasındaki değişimler,

bütün bireylerin alt puanlarındaki değişimlerinden daha önemli olur. Bir başka ifadeyle farklı örneklem gruplarındaki bireylerin alt testlerden aldıkları alt puanlar arasındaki değişim, evrende yer alan bütün bireylerin alt testlerden aldıkları alt puanlar arasındaki değişimden daha önemlidir. Çok boyutlu bir testin uygulanması sonucunda alınan çoklu alt puanlar test puanı profili olarak düşünülen bir değerlendirmeden türetilmektedir. Test puanı profili belirli bir grup öğrencinin test puanlarının toplamı olarak tanımlanabilir (Bulut, 2013).

Bulut'un (2013) aktarımlarına göre, test puanı profillerinin ölçme kesinliğinin belirlenmesinde Davison, Chang ve Davenport (2012) toplam değişim, bireyler arasındaki değişim ve alt puanlar arasındaki değişime bağlı olarak test puanı profillerinde bireysel farklılıkların güvenilirliğini tahmin etmeye yarayan bir yaklaşım önermişlerdir. Bu yaklaşım, esas olarak, Conger ve Lipshits (1973) tarafından önerilen canonical test güvenilirliğinin genişletilmiş bir halidir (bir uzantısıdır). Test puanı profilleri için canonical güvenilirlik, geleneksel tek değişkenli güvenilirliğin çok değişkenli versiyonu olarak gerçek puan varyansının gözlenen puan varyansına oranıdır (Conger ve Lipshitz (1973); Akt: Bulut (2013)). Sonuç olarak, çok boyutlu modellemelere göre gerçekleştirilen analizlerin güvenilirlik düzeyini belirlemek için genişletilmiş klasik güvenilirlikten yararlanılabilir.

### **İlgili Araştırmalar**

Bu bölümde, ilgili alanyazında yapılan çalışmalara ve bu çalışmaların sonuçlarına yer verilmiştir.

**Türkiye'de yapılan araştırmalar.** Kan (2004), yetenek düzeyinin ve örneklem büyüklüğünün madde tepki kuramından kestirilen madde parametreleri üzerindeki etkisini araştırdığı çalışmasında, madde değişmezliği prensibinin sağlanıp sağlanmadığını ortaya koymayı amaçlamıştır. Bu amaçla 553108 öğrencinin katılımıyla gerçekleşen Öğretmen Kurumları Seçme ve Yerleştirme Sınavına (ÖKSYS) ait 25 maddelik Türkçe alt testinin sonuçlarından yararlanmıştır. Madde parametrelerinin değişmezliği prensibinin korunma derecesini belirlemek için yüksek, düşük ve orta yetenek düzeyine sahip gruplardan kestirilen madde parametreleri hem birbirleriyle hemde evrenden kestirilen madde parametreleriyle karşılaştırılmıştır. Pearson moment çarpımlar korelasyon ve regresyon analizi

sonucunda deęişik yetenek gruplarından kestirilen madde parametreleri arasında doęrusal ve anlamlı bir iliřki bulunmuřtur. Aynı zamanda a-) dūřuk ve yūksek yetenek grubundan kestirilen, b-) orta ve yūksek yetenek grubundan kestirilen, c-) evren ve yūksek yetenek grubundan kestirilen a parametreleri arasında istatistiksel olarak anlamlı korelasyonlar elde edilmemiřtir. Őrneklem geniřlięi ve yetenek dūzeylerinin madde parametreleri űzerindeki etkisini belirlemek iin deęişik yetenek dūzeyleri ve űrneklem geniřliklerinden kestirilen madde parametreleri űzerinde korelasyon ve regresyon analizi uygulanmıřtır. Bu analizler sonucunda deęişik űrneklem geniřliklerinde ve yetenek dūzeylerinde kestirilen madde parametreleri arasında doęrusal ve anlamlı iliřkiler bulunmuřtur.

Kōse (2010), kendisi tarafından geliřtirilmiř olan Tūrke testi verilerini tek ve ok boyutlu Madde Tepki Kuramı modelleri altında madde, yetenek parametreleri ve model-veri uyum deęerlerini farklı deęiřkenler aısından karřılařtırmıřtır. Bu alıřmada Kōse (2010), ok boyutlu MTK modelleri kapsamında, elde edilen madde ve yetenek parametrelerinin daha az hata ierdięini, daha duyarlı űlümlere ulařıldıęını, model veri uyumunun ok boyutlu kuram lehine saęlandıęını ortaya koymuřtur. Bunun yanında űrneklem bűyüklüęü deęiřkeninin madde parametreleri kestirimlerini etkileyen en űnemli deęiřken olduęunu, yetenek kestiriminde ise en űnemli etkenin test uzunluęu olduęunu belirtmiřtir. Őrneklem bűyüklüęü ve test uzunluęu deęiřkenlerinin birlikte ele alınmasının, ok boyutlu kuram kapsamında sadece madde parametreleri űzerinde pozitif bir etkisi olduęuna ulařmıř, yetenek parametreleri ve model veri uyumunda tek ve ok boyutlu MTK kuramları kapsamında belirgin bir etkisi olduęu sonucuna ulařmamıřtır.

Sūnbűl (2011), eřitli boyutluluk űzelliklerine sahip yapılardaki parametre deęiřmezlięini, Klasik Test Kuramı (KTK), Tek Boyutlu Madde Tepki Kuramı (MTK) ve ok boyutlu Madde Tepki Kuramı (BMTK) erevesinde inceledięi siműlatif alıřmasında, parametre deęiřmezlięinin gōstergesi olarak standart hata deęerlerini kullanmıřtır. KTK iin elde ettięi sonulara gōre, literatűrűn aksine parametre deęiřmezlięinin bűtűn parametreler ve kořullar iin saęlandıęı bulgusuna ulařmıřtır. MTK iin elde ettięi sonulara gōre hem a hem de b parametresi iin parametre deęiřmezlięi bűyűk oranda saęlanmıřtır ancak bazı yapılarda parametre deęiřmezlięinin saęlanamadıęına ifade etmiřtir. Bu yapılar; ikinci boyutun birinci

boyuta kuvvetçe eşdeğer (madde sayısı ve a parametre ortalaması) ve boyutların birbirine dik konumlandığı deneysel hücreler olduğunu belirtmiştir. ÇBMTK çerçevesinde yapılan değerlendirmeler sonucunda parametre değişmezliğinin büyük oranda sağlandığı ancak, bazı deneysel hücreler için sağlanamadığı sonucuna ulaşmıştır. ÇBMTK kapsamında parametre değişmezliğini tehdit eden iki önemli etken olduğunu ifade etmiştir: bunlardan birincisi; ÇBMTK kestirimleri, örneklem büyüklüğüne karşı duyarlıdır ve ancak geniş örneklemelerde değişmez madde parametreleri üretebilmektedir. İkincisi ise; boyut oluşturma gücünden yoksun madde öbeklerinin boyut olarak tanımlanması, madde parametrelerinin değişmezliği için problem yaratmaktadır.

Şahin (2012), örneklem büyüklüğü ve test uzunluğunun model veri uyumu üzerine etkisini incelediği çalışmasında,  $x^2/sd$  oranı, standartlaştırılmış artık analizi,  $G^2$  olabilirlik oranı, test karakteristik eğrileri incelenmesi şeklinde dört farklı yöntemle farklı örneklem ve test uzunluklarında model veri uyumunu incelemiştir. Standartlaştırılmış artık analizi yönteminin örneklem büyüklüğünden ve kısmen de olsa test uzunluğundan etkilendiği; ancak bu yöntemin 10 maddelik bir testte 500, 20 ve 30 maddelik testlerde ise 750 kişilik örneklemekten fazla örneklem büyüklüğünden etkilenmediği belirlenmiştir.  $x^2/sd$  yönteminin ise test uzunluğu ve örneklem büyüklüğünden kısmen etkilendiği; 20 maddelik testlerde örneklem büyüklüğünden etkilenmeyen bu yöntemin, madde sayısı 10 olduğu durumlarda, 3000 kişilik örneklem büyüklüğünden sonra; madde sayısı 30 olduğu durumlarda ise 5000 kişilik örneklem büyüklüğünden sonra örneklem büyüklüğünden etkilenmeye başladığı görülmüştür.  $G^2$  Olabilirlik Oranı, test karakteristik eğrilerinin incelenmesi yöntemlerinin test uzunluğu ve örneklem büyüklüğünden etkilenmediği ve dolayısıyla çalışmada kullanılan farklı test uzunlukları ve örneklem büyüklüklerinde kullanılmalarının uygun olabileceği sonucuna varılmıştır. Test uzunluğu ve örneklem büyüklüğünün madde parametreleri üzerindeki etkisinin araştırılması ile 150 kişilik örneklemde, 10, 20 veya 30 maddelik test uzunluklarında 1PL modelde parametre kestirimi yapabilmek için 150 kişilik örneklemin yeterli olabileceği sonucuna varılmıştır. Bunun yanında 2PL modelde madde parametresi kestirimi yapılabilmesi için, 10 maddelik bir test için en az 750; 20 maddelik bir test için en az 500 ve 30 maddelik bir test için en az 250 kişilik örneklemelerin yeterli olabileceği sonucuna varılmıştır. 3PLM'de ise madde parametreleri kestiriminde 10

maddelik bir test için en az 750, 20 maddelik bir test için en az 750 ve 30 maddelik bir test için en az 350 kişilik örneklemelerin yeterli olabileceği sonucuna varılmıştır.

Özer Özkan (2012), bir test bataryasındaki başarı ölçüleri kestiriminin doğruluğunun belirlenmesi ve ampirik olarak Klasik Test Kuramı (KTK), tek ve çok boyutlu Madde Tepki Kuramı (MTK) modellerinin Öğrenci Başarılarının Belirlenmesi Sınavı'nın (ÖBBS) Türkçe ve matematik alt testi verilerine uygulanarak elde edilen başarı ölçülerinin karşılaştırılmasını amaçladığı çalışmada başarı ölçülerini daha az hata ile kestiren en iyi modeli ortaya koymaya çalışmıştır. Araştırmada kullanılan veriler, 2008 yılında sekizinci sınıflara uygulanan ÖBBS'nin 25 maddelik çoktan seçmeli Türkçe ve matematik testlerinin 9876 öğrenciye uygulanmasıyla elde edilmiştir. Gerekli varsayımlar test edildikten sonra, çok boyutlu olan testlerin alt boyutları belirlenmiştir. Türkçe ve matematik testi verilerinin analizi sonucunda tüm testten çok boyutlu MTK ile kestirilen yetenek parametrelerinin alt boyutlar bazında tek boyutlu MTK'ye göre kestirilen yetenek parametreleri ve KTK'ye göre elde edilen test puanlarına kıyasla kısmen daha düşük standart hataya sahip olduğu belirlenmiştir. KTK'ye göre elde edilen puanlar ile tek ve çok boyutlu MTK'ye göre belirlenen yetenek kestirimlerinin güvenilirlikleri incelendiğinde ise çok boyutlu MTK'ye göre elde edilen güvenirliliğin kısmen daha yüksek olduğu görülmektedir. Puanlar ve yetenek kestirimleri arasındaki ilişkiler incelendiğinde ise KTK'ye göre elde edilen puanlar ile tek ve çok boyutlu MTK'ye göre belirlenen yetenek parametreleri arasında yüksek ve manidar bir ilişki bulunmuştur. Benzer şekilde tek ve çok boyutlu MTK'ye göre elde edilen yetenek kestirimleri arasında da yüksek ve manidar bir ilişki belirlenmiştir.

Koçar (2014), iki boyutlu bir özelliğe ve ikili puanlanan bir yapıya sahip veri setleri üreterek gerçekleştirdiği monte carlo çalışmada, TBPMTK, TBPOMK ve ÇBMTK kapsamında model uyumlarının örneklem büyüklüğü ve test uzunluğunun parametre değişmezliği üzerinde etkisini incelemiştir. Parametre değişmezliği üzerinde etkisi incelenen örneklem büyüklükleri 100, 500, 1000 ve 5000 olarak belirlenmiştir. Ayrıca test uzunluğunun etkisini incelemek amacıyla üretilen ilk faktör 25 madde ile sınırlandırılırken; ikinci faktör 5, 15 ve 25 maddeden oluşturulmuştur. Boyutlar arası korelasyon ise üretilen veri setleri ile tahmin edilen yetenek parametreleri arasındaki ilişkilerin 0.00, 0.25 ve 0.50 olarak değişimlenmesi ile elde

edilmiştir. Her bir deneysel desenden, Monte Carlo çalışması kapsamında, 20 tekrar (replications) ile bulgular elde edilmiştir. Bu tekrarlar, tek ve çok değişkenli normal dağılım altında üretilmiştir. Çalışmanın bulguları incelendiğinde her üç model için de örneklem büyüklüğü ve test uzunluğundaki artış ile birlikte, daha iyi model veri uyumunu gösterdiği görülmüştür. TBPOMTK'da  $H_i$  ve  $p$  parametreleri ile TBMTK'da  $a$  parametresine ait standart hata değerleri, örneklem büyüklüğündeki artış ile azalma eğilimi göstermektedir. Hiçbir simülasyon koşulundaki değişim,  $b$  parametresi üzerinde etkili olamamıştır. ÇBMTK'da  $a_1$  ve  $a_2$  parametreleri, örneklem büyüklüğündeki artış ile birlikte azalma eğilimi göstermektedir. Test uzunluğu ve boyutlar arası korelasyondaki değişim ile  $a_1$  ve  $a_2$  parametrelerindeki değişim arasında bir örüntü bulunamamıştır.

Çakıcı Eser (2015); çok boyutlu madde tepki kuramının farklı modellerinden çeşitli koşullar altında kestirilen parametreleri incelediği çalışmada, farklı madde sayısı ve örneklem büyüklüğü koşullarına sahip değişen boyutluluk özelliklerindeki veri setlerinin madde ve birey parametreleri tek boyutlu ve çok boyutlu madde tepki kuramı ile kestirmiş; kestirime ilişkin olarak elde edilen standart hata ortalamalarının karekökü (RMSE), yanlılık ve kestirilen parametreler ile gerçek parametreler arasındaki korelasyon değerlerini incelemiştir. Elde edilen sonuçlara göre, madde parametrelerine ait RMSE değerlerinin ve korelasyonların hem madde sayısındaki hem de örneklem büyüklüğündeki değişimden etkilendiği görülmüştür. Buna göre, test uzunluğunun ve birey sayısının artması madde parametrelerinin daha düşük hatalı kestirilmesini, gerçek parametrelerle daha yüksek korelasyon vermesini sağlamıştır. Birey parametresi kestirimleri örneklem büyüklüğünün değişiminden etkilenmezken, madde sayısının artışı etkilenmiştir. Madde sayısının artması birey parametreleri kestirimlerine ait RMSE değerlerinin düşmesini, korelasyon değerlerinin artmasını sağlamıştır. Ancak madde sayısı sabit tutularak örneklem büyüklüğünün arttığı koşullarda birey parametrelerine ilişkin RMSE ve korelasyon değerlerinde önemli bir değişiklik gözlenmemiştir. Madde parametreleri çok boyutlu basit yapılarda çok boyutlu karmaşık yapılardan daha düşük hata ile kestirilmiştir. İki boyutlu karmaşık yapılarda birey parametreleri basit yapılardan daha düşük RMSE değerleri almıştır. Yanlılıklar iki boyutlu basit ve karmaşık yapılarda benzer değerlerde olmakla beraber 0,1'in altındadır. İki boyutlu yapılardan kestirilen birey parametreleri ile gerçek parametrelerin korelasyonları

kısa testlerde orta, uzun testlerde iyi derecede ilişkiye işaret etmektedir. Üç boyutlu yapılarda hem madde hem de birey parametresi kestirimleri iki boyutlu yapılardan daha düşük kararlılıkla yapılmıştır. Üç boyutlu karmaşık yapının kısa test koşulunda madde parametreleri için oldukça yüksek ( $>1,0$ ) RMSE değerleri hesaplanmıştır. Üç boyutlu karmaşık yapının kısa test koşulunda madde parametreleri için elde edilen korelasyonlar düşük değerlerdedir. Üç boyutlu karmaşık yapıda üç boyutlu basit yapıdan daha yanlış kestirimler yapılmıştır.

Mor-Dirlik (2017), Madde Tepki Kuramı kapsamında yer alan iki farklı yaklaşım olan parametrik ve parametrik olmayan madde tepki kuramı modellerinden kestirilen madde ve yetenek parametrelerini test uzunluğu, örneklem büyüklüğü ve maddelerin psikometrik nitelikleri faktörlerine göre karşılaştırılmayı amaçlamıştır. Temel araştırma modelinde yer alan araştırmanın verileri Uluslararası Matematik ve Fen Eğilimleri Araştırması'nın (Trends of International Mathematic and Science-TIMSS) 2011 yılındaki uygulamasından elde 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 5, 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.

Bıkmaz-Bilgen (2017), çok kategorili maddeler için Parametrik Madde Tepki Kuramı (PMTK) kapsamındaki Aşamalı Tepki Modeli (ATM) ve Parametrik olmayan Madde Tepki Kuramı (PoMTK) kapsamındaki Monoton Homojenlik Modeli (MHM) ile elde ettiği kestirimlerde örneklem büyüklüğünün, örneklem dağılımının, testte yer alan madde sayısının ve testte yer alan maddelerin yanıt kategorisi sayılarının etkilerini incelemiştir. Bu amaçla örneklem büyüklüğü, örneklem dağılımı, madde sayısı, maddenin kategori sayısı değişkenleri ile belirlenen 192 simülasyon koşulunu desenlemiştir. Örneklem büyüklüğü (N=100, 250, 500, 1000), örneklem dağılımı (normal dağılım, çarpıklık katsayısı -1,0 olan dağılım), madde sayısı (10, 20, 40, 80), maddenin yanıt kategorisi sayısı (3, 5, 7) koşulları için ATM ve MHM ile yaptığı kestirimleri sırasıyla model veri uyumlarını, güvenilirlik değerlerini ve madde parametrelerinin standart hatalarını hesaplayarak incelemiştir. Araştırma sonucunda ATM'de model veri uyumu hesaplarken değerlerin değişken artışından etkilenmesi, tek başına yorumlanamaması bu değerlerin karşılaştırılmasını ve genellenmesini zorlaştırdığı sonucuna ulaşmıştır. MHM'de model veri uyumunun pratik olarak hesaplanmasının, başka bir değere ihtiyaç duyulmadan tek başına yorumlanmasının ATM'ye göre üstünlük sağladığı sonucuna ulaşmıştır. Aynı zamanda güvenilirlik değerlerinin iki model için benzer sonuç verdiğini belirtmiştir.

MHM için hesaplanan parametrelerin, küçük örneklem ve kısa test koşullarında standart hataların ATM kestirimlerine göre oldukça düşük olduğu ve MHM için hesaplanan parametrelere ait standart hataların tüm koşullarda birbirine yakın değerler aldığı sonucuna ulaşmıştır.

**Yurtdışında yapılmış çalışmalar.** Ansley ve Forsyth (1985) yapay olarak farklı koşullarda üretilen iki boyutlu bir veriyi tek boyutlu ve çok boyutlu MTK'ye göre analiz etmişler ve bu analiz sonucunda veri setinde yer alan maddelere ait madde parametrelerini kestirmişlerdir. Çalışmanın sonucunda elde edilen ayırtedicilik parametreleri arasında anlamlı bir farklılık çıkmamıştır ancak tek boyutlu parametrik MTK'ye göre kestirilen güçlük parametreleri, çok boyutlu MTK'ye göre kestirilen güçlük parametrelerine göre anlamlı bir şekilde daha büyüktür.

Darasgow ve Parsons (1983) çok boyutlu verilerin tek boyutlu kurama dayalı olarak analiz edilebileceği ve böylece çok boyutlu verilere ait madde ve yetenek parametresi kestirimi yapabileceği öngörüsüyle gerçekleştirdikleri çalışmada simülasyon yolu ile oluşturdukları 1500 kişilik bir çalışma grubundan elde ettikleri verileri kullanmışlardır. Elde edilen bu yapay verinin analizinde 1PLM, 2PLM ve 3PLM kullanılmıştır. Kullanılan üç modelden en iyi madde ve yetenek parametresi kestiriminin 2PLM tarafından yapıldığını ortaya koymuştur. Ancak her ne kadar en iyi parametre kestirimi 2PLM'ye ait olsa da tek boyutluluk varsayımının ihlal edildiği durumda parametre kestirimi sonuçlarının etkilendiğini de ortaya konmuştur. Tek boyutluluğun ihlalinin yanı sıra boyutlar arası korelasyon büyüklüklerinin parametre kestirimi üzerinde etkileri olduğu sonucuna ulaşılmıştır.

Meijer, Sijstma ve Smid (1990) tek boyutlu parametrik ve parametrik olmayan modellerin karşılaştırılmasını amaçladıkları çalışmada 990 kişiden elde edilen sözel zeka testine ait sonuçları kullanmışlardır. Sözel zeka testinden elde edilen sonuçlar İMM, MHM ve 1PLM'ye göre analiz edilerek, analiz sonuçları kullanılan modellerin kısıtlılıklarına göre karşılaştırılmıştır. Çalışmanın sonunda 1PLM'nin İMM ve MHM'ye göre daha sınırlayıcı bir model olduğu ortaya konmuştur. Özellikle de bireylerin yetenek düzeylerine göre sıralanması amaçlanıyorsa, İMM'nin kullanılması önerilmiştir.

Chang (1992) gerçek ve yapay verilerden elde edilen sonuçları bir arada kullandığı çalışmasında, tek ve çok boyutlu MTK'ya göre analiz edilen gerçek ve yapay verilere ait yetenek kestirimi sonuçlarını karşılaştırmıştır. Belirli koşullar altında üretilen yapay veri 500 ve 1000 kişilik örneklemelerden oluşmaktadır. Çalışmada kullanılan gerçek veri ise iki farklı alt testi bulunan Educational Battery-Revised testinin 1000 kişiye uygulanmasıyla elde edilen sonuçlarından oluşmaktadır. İki boyutlu olan gerçek ve yapay verinin tek ve çok boyutlu MTK'ya göre analiz edilmesi ile çok boyutlu MTK'dan elde edilen yetenek kestirimlerinin tek boyutlu MTK'dan elde edilen sonuçlara göre daha küçük bir standart hata ile kestirildiği sonucuna ulaşılmıştır. İki kuramdan elde edilen yetenek kestirimlerine ait standart hata büyüklüklerinden dolayı, çok boyutlu MTK'dan elde edilen yetenek kestirimlerinin tek boyutlu MTK'dan elde edilen yetenek kestirimlerine göre daha büyük bir ölçme kesinliğine sahip olduğunu belirtmişlerdir.

Douglas (2001), parametrik olmayan madde karakteristik eğrisi ile belirlenen madde tepki kuramı modellerinin tanımlanması üzerine yaptığı derleme çalışmasında, sabit bir gizil yetenek dağılımı ile sadece tek bir dizi madde karakteristik eğrisi tarafından madde cevapları dağılımı manifestosunun üretilmesi mümkün olduğu zaman, modellere ait net bir tanımlama yapmayı başarmıştır. Madde karakteristik eğrisi çok genel bir sınıfa ait olduğu zaman bu tanımlama yapılamamaktadır. Fakat çok sayıda madde ile yapılan değerlendirmeler için kullanılan bütün modellerde, madde sayısı arttıkça birbirine yaklaşan ya da aralarındaki çok küçük farklılıkları gizil yeteneğin bütün değerlerinde 0'a birleştiren madde karakteristik eğrisine sahip olan bir dağılım manifestosu olduğu gösterilmiştir. Bu çalışmanın asıl sonucu ise, asimptotik tanımlama özelliğine sahip uzun değerlendirme süreçlerinde kullanılan parametrik olmayan MTK modellerinin uygulamasına yönelik teorik bir destek sağlamasıdır.

Walker ve Baretvas (2003), çok boyutlu olduğunu düşünerek puanlanan maddelerin tek boyutlu bir modele göre analiz ederek, tek boyutlu modellerin kullanımının çok boyutlu maddelerin puanlanması üzerindeki etkisini incelemişlerdir. Böylece çok boyutlu madde tepki kuramının tanılayıcı bir araç olarak nasıl kullanılacağını göstermişlerdir. Geniş ölçekli bir matematik testine ait gerçek verileri kullanarak, bir ve iki boyutlu doğrulayıcı model uyum gösterdiği zaman yetenek

düzeyleleriyle ilgili sınıflama yeterliklerindeki farklılığı araştırmışlardır. Tek boyutlu modelin kullanıldığı durumda yetenek tahmininin, genel matematik yetenek düzeyini temsil ettiğini düşünmüşlerdir. İki boyutlu model altında ise iki boyuttan bir tanesinin genel matematik yeteneğini, ikinci boyutun ise matematikteki iletişim yeteneğini temsil ettiğini düşünmüşlerdir. Sonuçta bireylerin yetenek düzeylerinin farklı modeller altında farklı sınıflama düzeylerinde olma olasılıklarının daha yüksek olduğunu ortaya koymuşlardır. Araştırmanın bir başka sonucu ise çok boyutlu bir verinin tek boyutlu bir model altında incelenmesinin doğru olmayan sonuçlara sebep olmasıdır.

**İlgili araştırmaların özeti.** Yapılan incelemeler sonucunda alanyazında ağırlıklı olarak KTK ve MTK uygulamalarından elde edilen parametre kestirimlerinin karşılaştırıldığı, MTK'nin özellikle parametrik uygulamalarından elde edilen parametre kestirimlerinin ve model uyumlarının karşılaştırıldığı çalışmaların olduğu belirlenmiştir. Gerçek ve yapay verinin kullanılmasıyla gerçekleştirilen kuramsal ve uygulama düzeyindeki bu karşılaştırma sonuçları genel olarak örneklem büyüklüğü ve test uzunluklarına göre elde edilmiştir. Alanyazın incelemesi sonucunda iki kategoriden oluşan gerçek (simülasyon koşulları ile üretilmeyen) verinin kullanılarak TBPOMTK modellerinden elde edilen sonuçların, kuramın parametrik karşılığı olan TBPMTK modellerinden elde edilen sonuçlar ile karşılaştırıldığı çok az sayıda çalışmanın olduğu görülmektedir. Aynı zamanda yine iki kategoriden oluşan gerçek verinin kullanıldığı ve tek boyutluluk varsayımının ihlal edilmesiyle tek boyutlu parametrik ve parametrik olmayan MTK'den elde edilen sonuçların çok boyutlu modellerden elde edilen sonuçlarla karşılaştırıldığı bir çalışmaya rastlanmamıştır. Bu nedenle tek ve çok boyutlu yapıların bir arada kullanıldığı ve parametrik ve parametrik olmayan MTK ile tek ve çok boyutlu MTK modellerinden elde edilen sonuçların karşılaştırıldığı bu çalışmanın alanyazına önemli katkılar sağlayacağı düşünülmektedir.

## Bölüm 3

### Yöntem

Bu bölümde, araştırmanın modeli, araştırmanın evreni ve örnekleme, veri toplama araçları, değişkenler, veri toplama süreci, verilerin analizi ile araştırmanın iç ve dış geçerliği ile ilgili açıklamalara yer verilmiştir.

#### Araştırma Modeli

Bu çalışma tek boyutlu parametrik ve parametrik olmayan MTK ile çok boyutlu MTK modellerine ilişkin betimleyici istatistikler elde etme, iki ya da daha fazla değişken arasındaki ilişkinin varlığını ve derecesini ortaya koyma açısından betimsel bir çalışmadır. Betimsel araştırmalar, olayların, objelerin, varlıkların, kurumların ve çeşitli alanların "ne" olduğunu açıklamaya çalışır (Kaptan, 1977).

#### Araştırmanın Evreni ve Örnekleme

Çalışmada kullanılan veriler, 2015 yılındaki TEOG'un birinci uygulamasındaki yer alan testlerin her biri için A kitapçığını alan öğrenci cevaplarından elde edilmiştir. A kitapçığında Türkçe, Matematik, Fen Bilgisi, Din Kültürü ve Ahlak Bilgisi, İnkılap Tarihi, İngilizce, Almanca ve Fransızca alt testleri yer almaktadır. Bu çalışmada bütün alt testlerde A kitapçığını cevaplayan öğrenciler çalışmanın evrenini oluşturmaktadır. Çalışmanın bundan sonraki kısmında çalışma evreni "evren" şeklinde ifade edilmiştir. Evren büyüklüğü 63.871'dir. Bu çalışmada elde edilen sonuçlar örneklem büyüklüklerine göre karşılaştırıldığı için, evrenden rastgele seçilmiş 50, 100, 200, 500, 1000 ve 5000 kişilik gruplar çalışmanın örneklemini oluşturmaktadır. Rastgele seçilen bu örneklem evrenden sadece bir kez çekilmiştir. Buradan anlaşılacağı üzere evren büyüklüğü 63.871 ve kullanılan örneklem büyüklükleri ise 50, 100, 200, 500, 1000 ve 5000'dir. Örneklem büyüklükleri belirlenirken özellikle küçük örneklem için saha uygulamaları göz önünde bulundurulmuştur. Örneklem büyüklüğünün alt sınırı belirlenirken 2017-2018 eğitim ve öğretim yılı içerisinde Gaziantep'teki özel kurumlarda okumakta olan ortalama 8. sınıf öğrenci sayısı göz önünde bulundurulmuştur. Gaziantep İl Milli Eğitim Müdürlüğü'ne bağlı olan Ar-Ge biriminden alınan bilgiye göre Gaziantep'te bulunan 24 özel ortaokulda okumakta olan toplam 1188 tane sekizinci sınıf öğrencisi bulunmaktadır. Dolayısıyla her bir okul başına düşen öğrenci sayısının ortalama

değeri 49,5'tir. Bu nedenle örneklem büyüklüğünün alt sınırı 50 olarak belirlenmiştir. Diğer küçük örneklem büyüklükleri (100, 200 ve 500) de 50'nin çift katları olacak şekilde belirlenmiştir. Büyük örneklem büyüklüklerinin belirlenmesinde ise Hullin, Lissak ve Drasgow (1982), Goldman ve Raju (1986) ve Thissen ve Wainer (1982) tarafından yürütülen çalışma sonuçları göz önünde bulundurulmuştur.

Evrenden rastgele seçilen 50, 100, 200, 500, 1000 ve 5000 kişilik örneklem replikasyon yapılmaksızın sadece bir kez seçilmiştir. Örneklem seçiminde replikasyon yapılmaması bir başka ifadeyle her bir örneklemin sadece bir kez çekilmesi bu çalışmanın bulgularını sınırlayan bir durum olmasına rağmen replikasyon yapılmasının parametre değişmezliği ile ilgili olumsuz sonuçları da vardır. Örnek olarak 5000 kişilik bir örneklem seçerken 50-100 replikasyon yapıldığında her bir örnekleme çok sayıda aynı birey bulunabilir ve bu bireyler parametre değişmezliğini şişirebilir bir başka ifadeyle değişmezliğin sağlanmasına otomatik neden olurlar veya kestirimlerin yanlı olmasına neden olabilirler. Bu nedenle bu çalışmada replikasyon yapılmamıştır.

### **Veri Toplama Araçları**

Bu çalışmada 2015 yılı TEOG birinci sınavının A kitapçığında yer alan Türkçe, Matematik, Fen ve Teknoloji, T.C. İnkılap Tarihi, Yabancı Dil, Din Kültürü ve Ahlak Bilgisi alt testlerinden elde edilen verilerden yararlanıldığı daha önce belirtilmişti. Her bir alt test 20 maddeden oluşmaktadır. Çalışmada TEOG verilerinden yararlanılmasının sebebi alt testlerden en az ikisinin tek boyutlu ve iki boyutlu çıkması yönündeki beklentidir. TEOG'da yer alan alt testlerden çalışmanın amacına uygun olarak tek boyutlu ve iki boyutlu olan alt testlere ait sonuçların analiz sürecinde kullanılması amaçlanmıştır. Ancak boyutluluk analizi sonucunda TEOG'un bütün alt testlerinin tek boyutlu olduğu belirlenmiştir. TBPMTK ve TBPOMTK analizlerinde tek boyutluluk kanıtlarının daha güçlü olduğu belirlenen Matematik alt testinin sonuçlarından yararlanılmıştır. Ancak ÇBMTK analizlerinde Fen Bilgisi ile Din Kültürü ve Ahlak Bilgisi alt testleri birer boyutu temsil etmek üzere, iki alt testin birleştirilmesi ile oluşan 40 maddelik testten 20 madde seçilmesine karar verilmiştir. Bu 20 madde faktör analizi sonuçlarına göre seçilmiştir. Faktör analizi sonuçlarına göre oluşturulan iki boyutlu testin bir boyutu fen bilgisi ve ikinci boyutu din kültürü ve ahlak bilgisi sorularından oluşmaktadır. Faktör analizi sonucunda

birinci ve ikinci boyuta yük veren maddeler arasından, 10 madde Fen Bilgisi testinden ve 10 madde Din Kültürü ve Ahlak Bilgisi testinden seçilmiştir. Sonuçta 2 boyutlu ve 20 maddeden oluşan birleşik testin nihai formu elde edilmiştir. Sonuçta ÇBMTK analizleri için iki boyutlu olan birleşik testin sonuçlarından yararlanılmıştır.

## **Değişkenler**

Araştırmanın bağımlı ve bağımsız değişkenlerinin belirlenmesinde alanyazında yer alan önceki çalışmalar esas alınmıştır. Bununla birlikte araştırmanın bağımsız değişkenlerinin bağımlı değişkenler üzerinde önemli etkileri olduğu öngörülmektedir. Araştırmanın bağımlı değişkenleri, farklı koşullardan nasıl etkilendiği incelenen madde ve yetenek parametrelerinin kestirimi ve kestirime ait güvenilirlik düzeyleridir. Bağımlı değişkenleri nasıl etkilediği incelenen bağımsız değişkenler ise örneklem büyüklüğü ve madde tepki kuramının farklı uygulamaları olan TBPMTK, TBPOMTK ve ÇBMTK modelleridir.

## **Veri Toplama Süreci**

MEB Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü'nden alınan A kitapçığına ilişkin öğrenci cevaplarının her biri A kitapçığı cevap anahtarı dikkate alınarak, 1 doğru cevapları ve 0 ise yanlış cevaplar ile boş bırakılan soruları gösterecek şekilde yeniden düzenlenmiştir. TEOG uygulamalarında düzeltme formülünün kullanılmadığını belirtmekte yarar vardır. Sonuçta verilerin analizinde kullanmak üzere iki kategorili (1-0) veri seti elde edilmiştir.

## **Verilerin Analizi**

**MTK varsayımlarının test edilmesi.** Madde tepki kuramı belli varsayımların karşılanmasını gerektiren güçlü bir kuramdır. Bu varsayımlardan ilki ve en önemlisi tek boyutluluktur. Daha önce belirtildiği gibi herhangi bir MTK modeli kullanılmadan önce testin tek boyutlu olup olmadığının belirlenmesi gerekir. Testin tek boyutlu olduğu durumda tek boyutlu modellerden, aksi durumda ise çok boyutlu modellerden yararlanır. Bu nedenle verilerin analizi için yapılan ilk işlem boyutluluk analizidir.

Bu çalışmada sonuçlarından yararlanan TEOG'un alt testlerinin boyutluluğunu belirlemek için faktör analizinden yararlanılmıştır. Her bir alt test için

uygulanan faktör analizi sonucunda testlerin tek boyutlu olup olmadığı belirlenmiştir.

TEOG'da yer alan alt testlerin eşit aralıklı ölçek düzeyinde oldukları kabulünden hareketle öğrencilerin sınav puanının hesaplandığı ve sonrasında bu puanların standart puanlara çevrilmesiyle öğrencilerin, en yüksek puan alandan en düşük puan alan öğrenciye doğru sıralandıkları ve başarı sıralamalarına göre ortaöğretim kurumlarına yerleştirildikleri daha önce belirtilmişti. Buna göre TEOG'da yer alan alt testlerin eşit aralık ölçek düzeyinde olduğu kabul edilen sıralama ölçekleri, alt testlerde yer alan maddelerin ise sıralama maddeleri olduğu söylenebilir. Normal dağılımlı, sürekli sıralı değişkenler arasındaki ilişki belirlenirken polichorik korelasyon matrislerinden yararlanılır (Choi, Peters ve Mueller, 2010; Coenders ve Saris, 1995). Pearson tarafından geliştirilen bu yöntem, incelenen değişkenler sürekli iken ölçme araçlarının sıralı olabilecek veri sağladığı durumlarda Pearson korelasyon katsayısına alternatif olarak kullanılır (Öztuna, Elhan ve Kurşun, 2008). Yapılan analiz sonucunda maddelere ait çarpıklık katsayıları göz önünde bulundurularak, bütün test maddelerinin tek değişkenli normal dağılıma sahip olduğu görülmüştür. Bu nedenle bu çalışmada, polichoric korelasyon matrisinin kullanıldığı paralel analize dayalı boyutluluk sonuçları veren FACTOR 10.5.01 programından yararlanılmıştır.

Genel olarak faktör analizi sürecinde testte yer alan her bir madde birer faktör gibi düşünülür. Kullanılan döndürme yöntemine bağlı olarak (dik açılı döndürme yöntemleri) her bir değişken seti için varyans oranı hesaplanır. Varyans oranı her bir faktördeki maddelerin faktör yük değerleri karelerinin toplamının, toplam madde sayısına bölünmesiyle elde edilir (Çokluk, Şekercioğlu, Büyüköztürk, 2010). Açıklanan bu varyans oranının büyüklüğü ölçeğe ait faktör sayısının belirlenmesinde kullanılan önemli göstergelerden biridir. Bir ölçeğin faktör yapısı, açıklanan varyans oranının büyüklüğünce güçlüdür (Gorsuch, 1974).

Faktör analizi ile ilgili alan yazın incelemesi yapıldığı zaman, örneklem büyüklüğünün faktör analizi sonuçlarını önemli ölçüde etkilediğine dair bir görüş birliği olduğu görülmüştür. Genel olarak örneklem büyüklüğü 1000'in üzerinde olduğu zaman faktör analizi mükemmel sonuçlar vermektedir (Tabachnick ve Fidell, 2001). Bu çalışmada, evrenden rastgele çekilmiş 10.000 kişiye ait veriye faktör analizi uygulanmıştır. Bu örneklem büyüklüğü faktör analizinden mükemmel



sonular elde etmek iin yeterlidir. Ancak yine de faktör analizi iin örneklem büyüklüğünün yeterli olup olmadığını ve veri yapısının faktör analizi iin uygunluğunu ortaya koyan KMO ve Bartlett küresellik testi sonuçları incelenmiştir. TEOG’da yer alan her bir alt test iin elde edilen Bartlett testinin istatistiksel olarak anlamlı olması, verilerin ok deęişkenli normal dağılımdan geldiğini ve dolayısıyla verilerin, faktör analizinin uygulanması iin uygun bir yapıya sahip olduğunu gösterir (okluk, Şekerciođlu, Büyüköztürk, 2010). Bunun yanında örneklem büyüklüğünün faktör analizine uygunluğu aısından KMO deęerinin 0,60’dan büyük olması istenir (Tabachnick ve Fidell, 2001).

Faktör analizine geçmeden önce TEOG’da yer alan her bir alt testin örneklem büyüklüğü ve veri yapısı aısından faktör analizi iin uygunluğunu gösteren KMO deęeri ile Bartlett Küresellik Testi sonucu elde edilmiştir. Bu sonuçlar Tablo 1’de görölmektedir.

Tablo 1

*KMO ve Bartlett Küresellik Testi*

	KMO	Bartlett Küresellik Testi (p)
Matematik	0,985	0,00
Türke	0,946	0,00
Fen Bilgisi	0,942	0,00
İngilizce	0,943	0,00
İnkılap Tarihi	0,947	0,00
Din Kültürü	0,952	0,00

Tablo 1’de görüldüğü gibi TEOG’da yer alan bütün alt testlere ait KMO deęerleri 0,60’ın üzerindedir. Daha önce belirtildiği gibi bu durum, her bir alt testten elde edilen verinin örneklem büyüklüğü aısından faktör analizi iin uygun olduğunu gösterir. Benzer şekilde TEOG’da yer alan bütün alt testlere ait Bartlett Küresellik testi sonucu 0,01 manidarlık düzeyinde anlamlı çıkmıştır ( $p < 0,01$ ). Buna göre TEOG’da yer alan bütün alt testlerin ok deęişkenli normal dağılımdan gelmesi aısından faktör analizi iin uygun olduđu söylenebilir.

TEOG’da yer alan her bir alt testin faktör analizi iin uygunluğunun ortaya konmasının ardından, alt testlerin boyutluluğunu belirlemek amacıyla faktör analizi

uygulanmıştır. Polichoric korelasyon matrisine dayalı paralel analiz sonucunda her bir alt testin tek boyutlu olduğu ortaya konmuştur. Elde edilen sonuçlar eklerde yer almaktadır. Mevcut verinin, elde edilen tek faktörlü modelle olan uyumunu ortaya koymak için kullanılan uyum indeksleri Tablo 2'de görülmektedir.

Tablo 2

*Paralel Analiz Sonucunda Elde Edilen Uyum İndeksleri*

	GFI	RMSR
Matematik	1	0,009
Türkçe	0,999	0,01
Fen Bilgisi	0,999	0,013
İngilizce	0,998	0,001
İnkılap Tarihi	1	0,022
Din Kültürü	0,998	0,023

Mevcut verinin, elde edilen tek faktörlü modelle olan uyumunu ortaya koymak için iki göstergeden yararlanılmıştır. Bunlar GFI ve RMSR'dir. GFI'nın 1'e yakınlığı ölçüsünde model ile veri uyumludur. RMSR ise Kelly'nin ölçüt değeri olan (0,0316)'dan küçük olursa model ile verinin iyi uyum sağladığı söylenebilir (Harman, 1962). Buna göre Tablo 2'de görüldüğü gibi her bir alt teste ait GFI uyum indeksi 1'e çok yakındır. Öyle ki matematik ve inkılap tarihi alt testlerinden elde edilen veriler ile kurulan model mükemmel yakın bir uyum göstermektedir. Benzer şekilde her bir alt testten elde edilen verilerin RMSR açısından kurulan tek faktörlü model ile iyi uyum gösterdiği Tablo 2'de görülmektedir.

Sonuç olarak TEOG'da yer alan bütün alt testlerin tek boyutlu bir yapıya sahip olduğu ortaya konmuştur. Bu sebepten dolayı ÇBMTK analizlerine uygun olan bir alt test bulunmamaktadır. Bu nedenle ÇBMTK analizleri için kullanılmak üzere, iki tane alt testten madde seçerek 2 boyutlu ve tek boyutlu MTK analizlerinde kullanılan testle aynı uzunluğa sahip (20 madde) yeni bir test oluşturulmuştur. Bu yeni testin oluşturulma süreci veri toplama araçları başlığı altında kısaca bahsedilse de bu sürecin detaylı bir şekilde açıklanması gerekmektedir.

ÇBMTK analizleri için gerekli olan testin oluşturulmasında yararlanılan alt

testlerin seçiminde alt testlerin ikili kombinasyonlarından elde edilen faktör analizi sonuçları göz önünde bulundurulmuştur. Elde edilen ikili kombinasyonlara uygulanan faktör analizi sonucunda aynı alt testte bulunan maddelerin birlikte açıkladığı varyans miktarının yüksek, farklı alt testlerde yer alan maddelerin birlikte açıkladığı varyans miktarının düşük olması koşuluna bağlı olarak birleşik testin oluşturulmasında kullanılan alt testlere karar verilmiştir. Bu şartı en iyi şekilde sağlayan alt testler Fen Bilgisi ile Din Kültürü ve Ahlak Bilgisi alt testleridir.

Fen Bilgisi ile Din Kültürü ve Ahlak Bilgisi testleri birleştirildiğinde 40 maddelik birleşik bir test elde edilmiştir. Bu teste faktör analizi uygulandığında test iki boyutlu çıkmıştır. Alanyazında yer alan çalışmalar incelendiğinde parametrik ve parametrik olmayan MTK ile tek ve çok boyutlu MTK için, kullanılan testlerin uzunluğunun parametre kestirimlerini etkilediği görülmüştür. Karşılaştırmaların anlamlı olabilmesi için çok boyutlu ve tek boyutlu MTK analizlerinde kullanılan testlerin uzunluklarının aynı olması gerekmektedir. Buradan hareketle Fen Bilgisi ile Din Kültürü ve Ahlak Bilgisi testlerinin birleştirilmesiyle oluşan 40 maddelik testten 20 madde seçilmesine karar verilmiştir. Bu 20 madde faktör analizi sonuçlarına göre seçilmiştir. Faktör analizi sonucunda birinci ve ikinci boyuta yük veren maddeler arasından, 10 madde Fen Bilgisi testinden ve 10 madde Din Kültürü ve Ahlak Bilgisi testinden seçilmiştir. Sonuçta 2 boyutlu ve 20 maddeden oluşan birleşik testin nihai formu elde edilmiştir. Birleşik testin nihai formuna uygulanan polichoric korelasyon matrisine dayalı paralel analiz sonucunda, testte yer alan her bir maddenin dolayısıyla boyutun açıkladığı varyans miktarına göre testin iki boyutlu olduğu ortaya konmuştur. Bununla birlikte iki boyut arasındaki korelasyon katsayısını 0,70 olduğu tespit edilmiştir. Yapılan boyutluluk analizi sonucunda her ne kadar testin 2 boyutlu bir yapıya sahip olduğu ortaya konmuş olsa da boyutlar arasındaki korelasyon katsayısı yüksektir. Mevcut verinin, elde edilen iki faktörlü modelle olan uyumunu ortaya koymak için kullanılan iki gösterge arasından GFI değeri 0,999 ve RMSR değeri ise 0,01'dir. Sonuçta birleşik teste ait verilerin 2 faktörlü yapıya uyum sağladığı sonucuna ulaşılmıştır.

MTK'nin ilk varsayımı olan tek boyutluluk varsayımının test edilmesinden sonra sıra yerel bağımsızlıktadır. Tek boyutluluk varsayımının karşılandığı durumda yerel bağımsızlığın da sağlandığı kabul edilir (Embretson ve Reise, 2000). Bu

nedenle bu çalışmada tek boyutlu yapılar için yerel bağımsızlık ayrıca test edilmemiştir. Böylece TBPMTK için önemli olan iki varsayım da bütün testler için sağlanmıştır. TBPMTK analizlerinde TEOG'da yer alan alt testlerden herhangi biri kullanılabilir. Bununla birlikte yerel bağımsızlık varsayımı iki boyutlu test için test edilmiştir. Yerel bağımsızlığı test etmek için kullanılan yöntemlerden bir tanesi koşullu maddeler arası korelasyonların incelenmesidir (Ferrara, Huyny&Baghi, 1997; Akt: Bulut, 2015). Bu çalışmada yerel bağımsızlığı test etmek için belli bir yetenek ranji arasındaki (yüksek yetenek ve düşük yetenek grupları) maddelerarası korelasyonlardan yararlanılmıştır. Alt grup ile üst grupları belirlemek için ham puanların %20 ve %80'lik dilimleri kullanılır. Sınırlı yetenek düzeyindeki bireyler için elde edilen varyans ve kovaryans veya korelasyon matrislerinin köşegeninde yer alan elementlerin 0 veya 0'a çok yakın olması yerel bağımsızlık varsayımının karşılandığını göstermektedir (Hambleton, 1991; McDonald, 1981; Akt: Bulut, 2015). Bu noktadan hareketle bu çalışmada düşük ve yüksek yetenek gruplarındaki bireylerin maddelere verdikleri cevaplar üzerinden elde edilen maddeler arası korelasyonlar elde edilmiştir. Elde edilen sonuçlara göre her iki yetenek grubunda da maddeler arası korelasyonlar çok düşük çıkmıştır. Böylece birleşik test için yerel bağımsızlık varsayımının karşılandığı sonucuna ulaşılabilir. Yerelbağımsızlık için kullanılan maddeler arası korelasyon matrisi ekte sunulmuştur.

Test edilmesi gereken diğer bir varsayım gizil monotonluk varsayımıdır. Gizil monotonluğun ortaya konması test verilerinin gözlenebilen bir özelliğidir. MTF'lerin monotonluğu  $P_j(\theta)$  koşullu olasılığının  $\theta$  üzerinde monoton olarak azalmadığını ifade eder (Sijtsma ve Molenaar, 2002).  $P_j(\theta)$  ve  $\theta$  arasındaki ilişki monoton ve azalmayan bir ilişki olduğu sürece, herhangi bir fonksiyonel forma izin verilir (Sijtsma ve Molenaar, 2002).

Bu çalışmada gizil monotonluk varsayımının test edilmesi için R 3.0.2. yazılımında kullanılan Mokken paketinden yararlanılmıştır. Monotonluk varsayımının sonuçlarını yorumlamaya geçmeden önce önemli sembol ve kısaltmaları açıklamakta yarar vardır. (#AÇ) her bir madde için aktif çift sayısını, (#GMİ) gizil monotonluğun ihlalinin miktarı, (#GMİ/#AÇ) her madde çifti için ortalama olarak monotonluk ihlalinin miktarını, (maxGMİ) monotonluk ihlalinin miktarının en büyük değerini, (TOP) toplam monotonluk ihlalinin miktarını, (TOP/#AÇ) her madde

çifti için toplam monotonluk ihlalinin miktarını göstermektedir. Bütün bu değerler 0'dan anlamlı bir şekilde büyük olduğu takdirde gizil monotonluk varsayımı ihlal edilmiş olur (Van der Ark, 2007). Yorumlama için önemli olan bir diğer gösterge ise madde ölçeklenebilirlik katsayısı olan  $H_j$ 'dir. Her bir maddeye ait olan  $H_j$ 'lerin 0,30'dan küçük olması maddenin ayırteçilik bakımından zayıf olduğunu gösterir.  $H_j$  madde ayırt edicilik katsayısı olarak yorumlanır (Sijtsma ve Molenaar, 2002).

TBPOMTK modelleri için önemli olan gizil monotonluk varsayımı test edilmiş ve Din Kültürü ve Ahlak Bilgisi, İngilizce ve Fen bilgisi alt testleri için ise bazı maddeler tarafından varsayımın karşılanmadığı görülmüştür. Türkçe ve Matematik alt testlerinde yer alan tüm maddelerin ise varsayımı karşıladığı tespit edilmiştir. Her bir alt test için elde edilen gizil monotonluk varsayımı sonuçları eklerde sunulmuştur. Matematik testinin gizil monotonluk varsayımına ait analiz sonuçları Tablo 3'te görülmektedir.

Tablo 3

*Matematik Testinin Gizil Monotonluk Varsayımına Ait Analiz Sonuçları*

	H	#AÇ	#GMİ	#GMİ/AÇ	maxGMİ	TOP	#TOP/AÇ	zmax	#zsig	krit
1	0,38	0	0	0	0	0	0	0	0	0
2	0,36	0	0	0	0	0	0	0	0	0
3	0,42	0	0	0	0	0	0	0	0	0
4	0,37	0	0	0	0	0	0	0	0	0
5	0,36	0	0	0	0	0	0	0	0	0
6	0,23	0	0	0	0	0	0	0	0	0
7	0,33	0	0	0	0	0	0	0	0	0
8	0,33	0	0	0	0	0	0	0	0	0
9	0,27	0	0	0	0	0	0	0	0	0
10	0,23	0	0	0	0	0	0	0	0	0
11	0,41	0	0	0	0	0	0	0	0	0
12	0,33	0	0	0	0	0	0	0	0	0
13	0,38	0	0	0	0	0	0	0	0	0
14	0,44	0	0	0	0	0	0	0	0	0
15	0,36	0	0	0	0	0	0	0	0	0
16	0,28	0	0	0	0	0	0	0	0	0
17	0,35	0	0	0	0	0	0	0	0	0
18	0,18	0	0	0	0	0	0	0	0	0
19	0,24	0	0	0	0	0	0	0	0	0
20	0,31	0	0	0	0	0	0	0	0	0

Tabloda görüldüğü gibi matematik testinde yer alan tüm maddeler gizil monotonluk varsayımını karşılamaktadır. Bununla birlikte testte yer alan maddeler

ayırt edicilik açısından heterojen bir yapıya sahiptir. Analiz sonuçlarında görüldüğü gibi 6, 10, 18 ve 19 numaralı maddelerin ayırteediciliğı çok düşüktür. 9 ve 16 numaralı maddelerin ayırteedicilik düzeyleri ise sınıra yakındır. Diğer maddelerin ayırteedicilik düzeyleri ise referans değerinin üstündedir.

TBPMTK analizlerinde TEOG'da yer alan herhangi bir alt testin sonuçlarından yararlanabileceğı daha önce belirtilmişti. TBPOMTK analizlerinde ise Türkçe ve Matematik alt testlerinden birinin sonuçlarından yararlanılabileceğı görülmüştür. Bu durumda hem TBPMTK analizleri için hem de TBPOMTK analizleri için uygun olan ortak bir alt testin sonuçlarının kullanılmasına karar verilmiştir. Bu test Matematik alt testidir.

**Betimsel istatistik sonuçları.** Tek boyutlu parametrik ve parametrik olmayan MTK analizlerinde Matematik alt testinin sonuçlarından, çok boyutlu MTK analizlerinde birleşik teste ait sonuçlardan yararlanılmıştır. Matematik ve birleşik teste ait evren üzerinden elde edilen betimsel istatistik sonuçları Tablo 4'te görülmektedir.

Tablo 4

*Matematik Testine ve Birleşik Teste Ait Betimsel İstatistik Sonuçları*

	Matematik	Birleşik Test
N	63871	63871
Soru Sayısı	20	20
Std. Sapma	5,15	4,48
Ortalama	8,20	13,20
Çarpıklık Katsayısı	0,68	-0,31
Basıklık Katsayısı	-0,61	-0,79
Cronbach Alfa		
&	0,87	0,88
Tabakalı Alfa		

Tablo 4'e görüldüğü gibi veri analizinde kullanılan her iki test de 63871 kişiye uygulanmıştır ve her iki testte yer alan madde sayısı 20'dir. Matematik testine ait standart sapma değeri 5,15 birleşik teste ait standart sapma değeri ise 4,48 olup standart sapma değerleri birbirine yakındır. Bu durum her iki testten elde edilen puan çeşitliliğinin birbirine yakın olması şeklinde yorumlanabilir. Matematik

testinden elde edilen puan ortalaması 8,20 iken birleşik testten elde edilen puan ortalaması 13,20'dir. Tek boyutlu testten elde edilen sonuçların güvenilirliği için Cronbach alfa katsayısından yararlanılmıştır. İki boyutlu test için ise Tabakalı alfa katsayısı güvenilirlik kanıtı olarak kullanılmıştır. Tek boyutlu testin güvenilirliği 0,87 ve iki boyutlu testin güvenilirliği ise 0,88 olarak kestirilmiştir. Büyüköztürk ve arkadaşlarına (2010) göre 0,70 ve üzerindeki güvenilirlik katsayıları kabul edilebilir düzeyde olduğu için her iki testten elde edilen sonuçların güvenilir olduğu sonucuna ulaşılabilir. Her iki testten elde edilen çarpıklık ve basıklık katsayıları  $\pm 1$  aralığında yer almaktadır. Büyüköztürk ve arkadaşlarına (2010) göre çarpıklık ve basıklık katsayılarının  $\pm 1$  arasında yer alması puanların nomalden aşırı bir sapma göstermediğine işaret eder.

**Verilerin analizinde kullanılan model seçimi.** Tek ve çok boyutlu veri analizi için kullanılan testlerin belirlenmesinden sonra, tek ve çok boyutlu MTK kapsamında hangi modellerin kullanılacağı belirlenmesi gerekir. TBPMTK için matematik testinin sonuçlarına 2 PLM ile 3 PLM uygulanmıştır. 2 PLM için kestirilen parametre sayısı 40 ve elde edilen -2 LL değeri 1419674,370'tir. 3 PLM için ise kestirilen parametre sayısı 60 ve elde edilen -2 LL değeri 1702461,230'dur. 20 serbestlik derecesinde elde edilen fark değeri ise 282786,86'dır. Elde edilen bu sonuç 20 (60-40) serbestlik derecesindeki  $\chi^2$  kritik değeri (31,410) ile kıyaslandığında, anlamlı çıkmıştır. Dolayısıyla, -2 LL değeri düşük olan 2PLM'nin 3PLM'ye göre anlamlı farklılık yarattığı; diğer bir ifadeyle 2PLM'nin veriye daha iyi uyum sağlayan model olduğu söylenebilir. Bununla birlikte Embretson ve Reise (2000), çoktan seçmeli test maddesi ile çalışıldığı durumlarda 3PLM'nin, kişilik verileri ile çalışıldığı durumlarda ise 1PLM veya 2PLM'den birinin kullanılmasını önerir. Ancak modellere ait -2LL değerleri de göz önünde bulundurularak TBPMTK modeli olarak 2PLM'nin uygulanmasına karar verilmiştir.

TBPOMTK kapsamında ikili monotonluk modeli (IMM) tarafından açıklanabilen her veri seti daha zayıf bir MHM tarafından açıklanabilmektedir. MHM'nin bu kapsayıcı özelliğinden dolayı TBPOMTK analizleri için MHM'nin kullanılmasına karar verilmiştir.

ÇBMTK için birleşik testin sonuçlarına genişletilmiş M2PLM ile M3PLM uygulanmıştır. M2PLM için kestirilen parametre sayısı 40 ve elde edilen -2 LL değeri

1214446,23'dur. M3PLM için ise kestirilen parametre sayısı 60 ve elde edilen -2 LL değeri 1214490,81'dir. 20 serbestlik derecesinde elde edilen fark değeri ise 44,58'dir. Elde edilen bu sonuç 20 (60-40) serbestlik derecesindeki  $\chi^2$  kritik değeri (31,410) ile kıyaslandığında, anlamlı çıkmıştır. Bu durumda analizler için tercih edilen model -2 LL değeri daha küçük olan genişletilmiş M2PLM'dir.

**Verilerin analizinde kullanılan yazılımlar.** Tek ve çok boyutlu veri analizi için kullanılan modellerin belirlenmesinden sonra, analizlerin gerçekleştirilmesi için hangi yazılımlardan yararlanılacağı belirlenmiştir. Verilerin analizinde TBPOMTK analizlerinde R 3.0.2. yazılımındaki MOKKEN paketi, tek ve çok boyutlu MTK analizinde ise FlexMIRT 3.5 yazılımı kullanılmıştır (Cai, 2017). Tek boyutlu MTK parametre kestirimi 2 PLM'ye göre yapılmış; bireye ait yetenek kestirimlerinde EAP kestirim yöntemi kullanılmıştır. Çok boyutlu MTK analizleri genişletilmiş 2 PLM'ye göre kestirilmiştir. Bireye ait yetenek kestirimleri TBPMTK'de olduğu gibi ÇBMTK'de de EAP kestirim yöntemi üzerinden yapılmıştır. Hem TBPMTK hem de ÇBMTK analizlerinde hata değerleri EM algoritmasının hata değerleri kestirilerek belirlenmiştir (Cai, 2008).

**Veri analizi sonuçlarını yorumlamak için kullanılan göstergeler.** TBPOMTK, TBPMTK ve ÇBMTK için madde ayırtediciliği ve madde güçlüğü için farklı göstergelerden yararlanır. Daha önce belirtildiği gibi TBPOMTK için maddenin ayırtedicilikgücünü gösteren Hi parametresinin 0,30'dan küçük olması, maddenin ayırtedicilik bakımından zayıf olduğunu gösterir (Sijtsma ve Molenaar, 2002). TBPOMTK için madde güçlüğü'nün göstergesi olarak klasik güçlük parametresi olan p değerlerinden yararlanılmıştır. Bu durumun nedeni TBPOMTK için değişmeyen madde sıralaması klasik anlamda bir güçlük yorumunu beraberinde getirmektedir. Bu durum Sijtsma (2005)'in Weschler Zeka Ölçeği ile ilgili verdiği örnek üzerinden daha anlaşılır olabilir. Çocuklar için uygulanan Weschler Zeka Ölçeği değişmeyen madde sıralaması kuralına göre başlatılır ve sonlandırılır. Belirli bir yaş grubundaki çocuklar kendileri için "kolay" sayılacak bir madde ile teste başlar daha sonra başladıkları maddelere göre daha zor bir maddeye yönlendirilirler ve daha sonra ardışık maddelerde başarısız oldukları zaman uygulama sonlandırılır. Sonlandırma kararının temelinde çocukların bundan sonraki maddelerde de başarısız olacakları varsayımı bulunmaktadır. Örnekten de anlaşıldığı üzere



TBPOMTK'de madde güçlüğü bireylere bağlı bir parametredir. Bireyler ise madde güçlüğüne bağlı olarak başarılıdır ya da değildir. Bu yönüyle TBPOMTK her ne kadar bir MTK uygulaması olsa da klasik kuram ile madde tepki kuramının arasında kalmıştır. TBPOMTK uygulamalarındaki güçlük parametresi klasik kuramdaki güçlük parametresinden hesaplama süreçleri açısından biraz farklıdır. TBPOMTK uygulamaları için klasik güçlük parametresi hesaplanırken klasik kuramdan farklı olarak standart sapma değerleri de sürece dâhil edilmektedir (Andries van der Ark, 2007). Ancak yorumlanması açısından klasik kuramdaki güçlük parametresi ile benzer şekilde yorumlanmaktadır. Sonuç olarak daha önce belirtildiği gibi TBPOMTK için madde güçlüğünün göstergesi olarak klasik güçlük parametresi olan  $p$  değerlerinden yararlanılmıştır.

TBPMTK için ise ayırt ediciliğin göstergesi olarak  $a$  parametresi, madde güçlüğünün göstergesi olarak  $b$  parametresi kullanılır. Teorik olarak  $a$  ve  $b$  parametreleri  $\pm\infty$  aralığında değerler alırlar. Ancak pratikte  $a$  parametresinin  $[0,2]$  (Hambleton ve Swaminathan, 1985),  $b$  parametresinin  $[-3,3]$  aralığında değerler aldığı görülmüştür. TBPMTK'ye göre kestirilen  $a$  ve  $b$  parametrelerinin kestirilmesi için (2.16) denkleminde yararlanılmıştır.

Son olarak ÇBMTK için, TBPMTK'de olduğu gibi ayırt ediciliğin göstergesi olarak  $a$  parametresi kullanılır. Burada küçük bir fark vardır: ÇBMTK'de testin her bir boyutu için ayrı bir ayırtedicilik parametresi kullanılmaktadır. Bu çalışmada kullanılan birleşik test iki boyutlu olduğu için iki tane ayırtedicilik parametresi kestirilmiştir. Bunlar  $a_1$  ve  $a_2$  parametreleridir. ÇBMTK için madde güçlüğünün göstergesi olarak  $d$  parametresi kullanılır.  $d$  parametresi TBPMTK'deki  $b$  parametresi gibi yorumlanmaktadır. TBPMTK'ye göre kestirilen  $a_1$ ,  $a_2$  ve  $d$  parametrelerinin kestirilmesi için (2.24) denkleminde yararlanılmıştır.

Bu çalışmada madde puanlarından değil toplam test puanlarından yararlanılmıştır. Bu nedenle toplam test puanları için ortalama güçlük ve ayırtedicilik katsayıları hesaplanmıştır. Her üç kuram için de testin ortalama güçlüğü ve ayırtedicilik düzeyleri evrenden birer kez çekilen her bir örneklem için ayrı ayrı hesaplanmıştır. Her bir örneklem için kestirilen parametre ortalamalarının evren değerden ne kadar farklı olduğu incelenmiştir. Ancak bu fark istatistiksel olarak test edilmemiştir. Yapılan yorumlar sadece büyüklük-küçüklük ilişkisi çerçevesinde yapılmıştır.

Çalışmanın odağında olan parametre değişmezliğinin incelenmesi amacıyla, parametre değişmezliğinin göstergesi olarak standart hata ortalamalarından yararlanılmıştır. Ancak bulgular yorumlanırken parametre kestirimlerine ait SHO'ların yanı sıra parametre ortalamalarının da örneklem büyüklüğünden nasıl etkilendiği incelenmiştir. Bu amaçla her bir örneklem için kestirilen parametre ortalamalarına ait SHO'ların evren değerden ne kadar farklı olduğu incelenmiştir. Tıpkı parametre ortalamalarında olduğu gibi bu fark istatistiksel olarak test edilmemiştir. Yapılan yorumlar sadece büyüklük-küçüklük ilişkisi çerçevesinde yapılmıştır.

TBPOMTK için testin ortalama gücüne ait SHO, evrenden birer kez çekilen her bir örneklem için ayrı ayrı hesaplanmıştır.  $\sigma$  Standart hatayı,  $k$  madde sayısını,  $j$  herhangi bir maddeyi,  $S_j$   $j$ . Maddeye ait standart sapmayı ve  $n$  kişi sayısını göstermek üzere, testte yer alan her bir madde için  $p$  katsayılarına ait standart hata değeri (3.1) denklemi ile hesaplanmıştır.

$$(\sigma) = \frac{\sum_{j=1}^k \frac{S_j}{\sqrt{n}}}{k} \quad (3.1)$$

TBPMTK için ortalama  $a$  ve  $b$  parametreleri ile ÇBMTK için ortalama  $a_1$ ,  $a_2$  ve  $d$  parametrelerine ait SHO, Cai (2008) tarafından Flex MIRT 3.5 yazılımı için önerilen "Ortalama Standart Hata Algoritması" prosedürlerine göre hesaplanmıştır. TBPOMTK'de olduğu gibi TBPMTK'ye ve ÇBMTK'ye göre kestirilen testin ortalama gücüne ve ayırteçilik gücüne ait SHO, evrenden birer kez çekilen her bir örneklem için ayrı ayrı hesaplanmıştır.

Parametre değişmezliğinin yanı sıra örneklem büyüklüğünün etkisinin araştırıldığı bir diğer bağımsız değişkenin her bir kurama göre kestirilen yetenek kestirimleri arasında ilişki olduğu daha önce belirtilmişti. Tek ve çok boyutlu MTK'ye göre kestirilen yetenek kestirimleri EAP yöntemine göre elde edilmiştir. Toplam puana dayalı olarak elde edilen yetenek kestirimleri Tong ve Kolen (2010) tarafından ifade edilen EAP kestirim prosedürlerine göre elde edilmiştir.

Kestirilen yetenek parametreleri arasındaki ilişkiyi incelemek amacıyla iki farklı korelasyon türünden yararlanılmıştır. Bunlardan birincisi Pearson Momentler Çarpımı Korelasyon Katsayısı'dır (PMÇKK). Kestirilen yetenek parametreleri eşit

aralık ölçek düzeyinde olduğu için PMÇKK'nin kullanılmasına karar verilmiştir. Ancak yetenek aynı zamanda bireylerin sıralanmasını da beraberinde getirdiğinden ve özellikle küçük olan örneklerde yetenek değerinden çok yetenek sırasının benzer olmasının incelenmesi daha anlamlı olacağından Speraman'ın sıra farkları korelasyonundan da yararlanılmıştır.

Son olarak örneklem büyüklüğünün etkisinin incelendiği bağımsız değişken her bir kurama göre elde edilen güvenilirlik düzeyleridir. TBPOMTK için H katsayısı güvenilirlik katsayısı olarak kullanılırken, TBPMTK için marjinal güvenilirlikten yararlanılmıştır. ÇBMTK için ise genişletilmiş klasik güvenilirlikten yararlanılmıştır. D boyut sayısını, d herhangi bir boyutu,  $X_{id}=\{X_{i1}, X_{i2}, \dots, X_{iD}\}$  i kişinin tüm boyutlardan aldığı gözlenen puanını,  $T_{id}=\{T_{i1}, T_{i2}, \dots, T_{iD}\}$  i kişinin tüm boyutlardan aldığı gerçek puanını göstermek üzere  $\sum_{d=1}^D \sigma^2(T_{id})$  gerçek puan varyansını ve  $\sum_{d=1}^D \sigma^2(X_{id})$  gözlenen puan varyansını göstermek üzere, genişletilmiş klasik güvenilirlik indeksi (3.2) denklemi üzerinden elde edilmiştir.

$$\rho_T = \frac{\sum_{d=1}^D \sigma^2(T_{id})}{\sum_{d=1}^D \sigma^2(X_{id})} = \frac{T_T}{T} \quad (3.2)$$

Çalışmaya ait bulguların elde edilmesi genel olarak iki boyut altında toplanabilir. Bunlardan ilki tek boyutlu olan Matematik testi tek boyutlu parametrik ve parametrik olmayan MTK altında modellenmiş, bu modellemeye göre madde ve yetenek parametreleri ile güvenilirlik katsayıları kestirilmiştir. İkinci boyutta ise iki boyutlu olan birleşik testin tek boyutlu olduğu kabulünden yola çıkılarak tek boyutlu parametrik ve parametrik olmayan MTK'ye göre tek boyutlu bir model altında analiz edilmiştir. Buna ek olarak iki boyutlu olan birleşik test doğasına uygun olarak ÇBMTK'ye göre modellenerek analiz edilmiş ve bu analizlerin sonucunda madde ve yetenek parametreleri ile güvenilirlik düzeyleri kestirilmiştir.

### **Araştırmanın İç ve Dış Geçerliliği**

**Araştırmanın iç geçerliliği.** Araştırmanın iç geçerliliği bağımlı değişkendeki değişimlerin, bağımsız değişkenlerle açıklanma derecesi ile ilgilidir. Bu çalışmada madde ve yetenek kestirimleri ile bu kestirime ait güvenilirlik düzeylerindeki

değişim, örneklem büyüklüğü ve MTK'nin farklı uygulamaları tarafından açıklanabildiği için araştırmanın iç geçerliği sağlanmıştır.

**Araştırmanın dış geçerliği.** Araştırmanın dış geçerliği bulguların genellenebilirlik derecesi ile ilgilidir. Bu çalışmadan elde edilen bulguların genellenebilirliği kullanılan örneklem büyüklükleri, MTK uygulamaları ve kullanılan testlerin konu alanı ile sınırlı olduğu için araştırmanın dış geçerliğini bu çerçevede incelemek gerekir. Dolayısıyla kullanılan örneklem büyüklükleri, kullanılan MTK uygulamaları ve testlere ait konu alanı çerçevesinde araştırma sonuçları genellenebilir.

## Bölüm 4

### Bulgular Ve Yorumlar

Bu bölümde, alt problem sırasına göre verilmiş araştırma bulguları ve bu bulgularla ilgili değerlendirmeler yer almaktadır.

#### 1. Alt Problemin Çözümüne İlişkin Bulgular

“Tek boyutlu testlerde parametrik ve parametrik olmayan MTK’ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrenden kestirildiği durumda nasıldır?” şeklinde oluşturulan birinci alt problemin çözümü için matematik testinin sonuçlarından TBPMTK ve TBPOMTK kapsamında madde parametreleri ortalaması ve bu parametrelerin ortalamasına ait SHO kestirilmiştir. Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlarda ve evren üzerinden TBPMTK ile TBPOMTK’ye göre kestirilen madde parametrelerine ait ortalama değerleri ile standart hata ortalamaları (SHO) Tablo 5’te görülmektedir.

Tablo 5

*TBPMTK ve TBPOMTK’ye Göre Kestirilen Matematik Testi’ne Ait Madde Parametreleri ve Standart Hata Ortalamaları*

Örneklem Büüklüğü	TBPOMTK			TBPMTK				
	$H_i$	$H_i(\text{SHO})$	$p$	$p(\text{SHO})$	$a$	$a(\text{SHO})$	$b$	$b(\text{SHO})$
50	0,32	0,09	0,48	0,109	1,07	0,32	0,09	0,37
100	0,39	0,07	0,45	0,109	1,31	0,28	0,24	0,23
200	0,33	0,05	0,43	0,107	1,32	0,22	0,34	0,18
500	0,33	0,03	0,43	0,107	1,41	0,15	0,34	0,10
1000	0,33	0,02	0,43	0,107	1,48	0,11	0,30	0,07
5000	0,33	0,01	0,43	0,107	1,50	0,05	0,29	0,03
Evren	0,33	0,00	0,43	0,107	1,52	0,01	0,30	0,01

TBPOMTK’ye göre kestirilen tek boyutlu teste ait madde parametrelerinin ilki ayırtediciliğin göstergesi olan ortalama  $H_i$  parametreleridir. Tablo 5’te görüldüğü gibi  $H_i$  ortalamaları incelendiğinde evren değerinin 0,33 olduğu görülmektedir. Evren değerine göreli olarak en uzak  $H_i$  ortalaması 100 örneklem büyüklüğünden

kestirilmiştir ( $H_i=0,39$ ). Evren değere en uzak  $H_i$  ortalaması bile evren değerden çok farklı değildir. Bu nedenle örneklemekten kestirilen  $H_i$  ortalamalarının büyüklük olarak evren değere yakın olduğu sonucuna ulaşılabılır. Bununla birlikte 200 örneklem büyüklüğünden itibaren ise evren değeri yansıtacak düzeyde kararlı bir yapıya sahip  $H_i$ 'lerin kestirildiği tabloda görülmektedir ( $H_i=0,33$ ).

$H_i$  ortalamalarının yanı sıra  $H_i$  ortalamasına ait SHO'nun evren değeri 0'a çok yakındır. Evren değeri en az yansıtan örneklem büyüklüğünün en küçük örneklem olan 50 olduğu ( $SHO=0,09$ ) ve örneklem büyüklüğü arttıkça  $H_i$  parametresi ortalamasına ait SHO'ların evren değerine yaklaştığı yine tabloda görülmektedir. Bunun yanı sıra örneklem büyüklüğü kaç olursa olsun SHO'lar büyüklük olarak evren değere yakın büyüklükte kestirilmiştir. Bu durumda evren değere yakın SHO'lar ile ortalama  $H_i$  parametresi kestirmek için büyük örneklemelerin kullanılmasının zorunlu olmadığı sonucuna ulaşılabılır.

TBPOMTK'ye göre kestirilen tek boyutlu teste ait madde parametrelerinin bir diğeri ise güçlük göstergesi olan ortalama  $p$  parametreleridir.  $p$  ortalamalarına ait evren değerinin 0,43 olduğu Tablo 5'te görülmektedir. Evren değerine en uzak  $p$  ortalaması en küçük örneklem büyüklüğünden kestirilmiştir ( $p=0,48$ ). Ancak görüldüğü gibi evren değere en uzak  $p$  ortalaması bile evren değerden çok farklı değildir.

Ortalama  $p$  parametresine ait SHO'nun evren değeri ise 0,107 olarak kestirilmiştir. Evren değere görece olarak en uzak SHO ise 50 ve 100 örneklem büyüklüklerinden kestirilmiştir ( $SHO=0,109$ ). SHO 200 örneklem büyüklüğünde ise en düşük değerini almıştır ( $SHO=0,107$ ) ve bu değer evren değere eşit olduğu söylenebilir. Tıpkı  $H_i$  ortalamalarına ait SHO'larda olduğu gibi örneklemekten kestirilen  $p$  ortalamalarına ait SHO'lar ile evren değeri arasında büyük farklılıkların olmadığı belirlenmiştir. Bununla birlikte 200 örneklem büyüklüğü ile evren arasındaki örneklemelerde de SHO büyüklüklerinde bir değişim olmamıştır.

Sonuç olarak TBPOMTK'ye göre kestirilen madde parametresi ortalamaları ve madde parametresi ortalamalarına ait SHO'ların 200 örneklem büyüklüğünden itibaren evreni yansıtacak düzeyde kararlı bir yapıya sahip olduğu görülmüştür.

Buna göre TBPOMTK için 200 örneklem büyüklüğü itibariyle parametre değişmezliğinin sağlandığı sonucuna ulaşılmıştır.

TBPMTK'ye göre kestirilen tek boyutlu teste ait madde parametrelerinin ilki ayırtediciliğin göstergesi olan ortalama  $a$  parametreleridir. Örneklemelerden kestirilen  $a$  parametresi ortalamasının örneklem büyüklüğüne bağlı olarak 1,07 ile 1,52 arasında değerler aldığı Tablo 5'te görülmektedir.  $a$  parametresi ortalamasının evren değeri 1,52 olarak kestirilmiştir.  $a$  parametresi ortalaması evren değerine en uzak değerini en küçük örneklem büyüklüğü olan 50'de almıştır ve örneklem büyüklüğü arttıkça  $a$  parametresi ortalamasına ait değerlerin de artma eğiliminde olduğu ve evren değerine yaklaştığı yine tabloda görülmektedir. 1000 örneklem büyüklüğünden itibaren  $a$  parametresi ortalamaları evren değerine çok yakın büyüklükte kestirilmiştir.

TBPMTK için  $a$  parametresi ortalamasına ait SHO'nun evren değeri ise 0,01 olarak kestirilmiştir. Evren değerine en uzak SHO'nun 50 örneklem büyüklüğünden kestirildiği tabloda görülmektedir ( $SHO=0,32$ ). Örneklem büyüklüğünün artmasıyla  $a$  parametresi ortalamasına ait SHO azalma eğilimi göstererek evren değere yaklaşmıştır. Tabloda görüldüğü gibi 500 örneklem büyüklüğü itibariyle evren değere çok daha yakın bir SHO ile  $a$  parametresinin kestirildiği sonucuna ulaşılabilir.

TBPMTK'ye göre kestirilen tek boyutlu teste ait madde parametrelerinin bir diğeri ise güçlük göstergesi olan ortalama  $b$  parametresidir.  $b$  parametresi ortalamasına ait evren değer 0,30 olarak kestirilmiştir. Evren değerine en uzak  $b$  parametresi ortalaması 50 örneklem büyüklüğünden elde edilmiştir ( $b=0,09$ ). Evren değeri en iyi yansıtan örneklem büyüklüğü ise 1000'dir. Elde edilen sonuçlara göre  $b$  parametresi ortalamasının, örneklem büyüklüğünün değişmesi ile düzenli bir artma ya da azalma eğiliminde olduğunu söylemek mümkün değildir. Ancak 1000 örneklem büyüklüğü itibariyle  $b$  parametresi ortalamasına ait değişimlerin daha az olduğu söylenebilir. Bu durumda 1000 örneklem büyüklüğü itibariyle  $b$  parametresi ortalamalarının evrene yakın büyüklükte olduğu sonucuna ulaşılabilir.

TBPMTK için  $b$  parametresi ortalamasına ait SHO'nun evren değeri ise 0,01 olarak kestirilmiştir. Evren değerine en uzak SHO en küçük örneklem büyüklüğü olan 50 üzerinden kestirilmiştir ( $SHO= 0,37$ ).  $b$  parametresi ortalamasına ait

SHO'nun örneklem büyüklüğünün artmasına bağlı olarak istikrarlı bir şekilde evren değerine yaklaştığı Tablo 5'te görülmektedir. Tabloda görüldüğü gibi 500 örneklem büyüklüğünden itibaren de evren değerine daha yakın büyüklükte bir SHO ile b parametresinin kestirildiği sonucuna ulaşılabılır.

TBPMTK üzerinden elde edilen madde parametresi kestirimlerinin incelenmesinin sonucunda, a ve b parametreleri ile a ve b parametrelerine ait SHO'lar sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

## 2. Alt Problemin Çözümüne İlişkin Bulgular

“Çok boyutlu testlerde tek ve çok boyutlu MTK'ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrenden kestirildiği durumda nasıldır?” şeklinde oluşturulan ikinci alt problemin çözümü için birleşik testin sonuçlarından ÇBMTK ile TBPMTK kapsamında madde parametreleri ortalaması ve bu parametrelerin ortalamasına ait SHO kestirilmiştir. Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlarda ve evren üzerinden ÇBMTK ve TBPMTK'ye göre kestirilen madde parametrelerine ait ortalama değerleri ile standart hata ortalamaları (SHO) Tablo 6'da görülmektedir.

Tablo 6

*Tek ve Çok Boyutlu MTK'ye Göre Kestirilen Birleşik Teste Ait Madde Parametreleri ve Standart Hata Ortalamaları*

Örneklem Büyüküğü	ÇBMTK				TBPMTK					
	a <sub>1</sub>	a <sub>1</sub> (SHO)	a <sub>2</sub>	a <sub>2</sub> (SHO)	d	d (SHO)	a	a (SHO)	b	b (SHO)
50	0,98	1,47	0,51	0,75	1,12	0,99	1,23	0,54	-0,82	0,48
100	0,60	0,70	0,44	0,45	0,77	0,53	1,24	0,75	-0,74	0,47
200	0,68	0,57	0,42	0,30	1,02	0,43	1,48	0,52	-0,79	0,36
500	0,68	0,43	0,40	0,20	0,87	0,36	1,68	0,69	-0,55	0,19
1000	0,67	0,25	0,41	0,14	0,84	0,20	1,76	0,16	-0,48	0,09
5000	0,67	0,11	0,42	0,06	0,88	0,08	1,94	0,07	-0,49	0,03
Evren	0,41	0,01	0,43	0,01	-0,18	0,01	1,52	0,01	0,30	0,01



ÇBMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin ilki ayırtediciliğin göstergesi olan ortalama  $a_1$  ve  $a_2$  parametreleridir. ÇBMTK'den kestirilen  $a_1$  ortalamasının evren değeri 0,41 olarak kestirilmiştir.  $a_1$  ortalamasının evren değerine en uzak kestirim en küçük örneklem büyüklüğü olan 50 üzerinden elde edilmiştir.  $a_1$  ortalamasının evren değerini iyi yansıtan bir örneklem büyüklüğü ise bulunmamaktadır.  $a_2$  ortalamasının evren değeri ise 0,43 olarak kestirilmiştir.  $a_2$  ortalamasının evren değerine en uzak kestirim en küçük örneklem büyüklüğünden elde edilmiş ve 100 örneklem büyüklüğünden itibaren ise evren değerine yakın büyüklükte  $a_2$  ortalaması kestirimleri elde edilmiştir. Elde edilen sonuçlara göre  $a_1$  ve  $a_2$  ortalamalarının, örneklem büyüklüğünün değişimine bağlı olarak düzenli bir artma ya da azalma eğiliminin olduğunu söylemek mümkün değildir.

ÇBMTK için ayırtedicilik parametrelerinin ortalamasına ait SHO'nun evren değerinin ise 0,01 olarak kestirildiği tabloda görülmektedir. Evren değerine en uzak SHO kestirimi en küçük örneklem büyüklüğü üzerinden elde edilmiştir ( $SHO_{a_1}=1,47$  ;  $SHO_{a_2}=0,75$ ). Örneklem büyüklüğünün artmasıyla her iki ayırtedicilik parametresi ortalamasına ait SHO da azalma eğilimi göstermiştir ve evren değere yaklaşmıştır  $a_1$  ortalamasının 5000 örneklemden itibaren ve  $a_2$  ortalamasının 1000 örneklemden itibaren evren değere yakın büyüklükte bir SHO ile kestirildiği Tablo 6'da görülmektedir. Bununla birlikte her iki ayırtedicilik parametresi ortalamasına ait SHO da örneklemlerden evrene doğru sürekli bir değişim göstermektedir.

Sonuç olarak ÇBMTK'den kestirilen ayırtedicilik parametrelerine ait ortalama değerleri de SHO'ları da sürekli bir değişim gösterdikleri için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

ÇBMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin bir diğeri ise güçlük göstergesi olan ortalama  $d$  parametreleridir. ÇBMTK'den kestirilen  $d$  parametresi ortalamasının evren değeri ( $d = -0,18$ ) olarak kestirilmiştir. Evren değerine en uzak  $d$  parametresi ortalaması en küçük örneklem büyüklüğünden elde edilmiştir ( $d=1,12$ ). Örneklemelerden elde edilen  $d$  parametresi ortalamaları incelendiğinde evren değerden farklı oldukları görülmektedir. Elde edilen sonuçlara göre  $d$  parametresi ortalamasının örneklem büyüklüğünün değişimine bağlı olarak düzenli bir artma ya da azalma eğiliminin olmadığı da görülmektedir. Ancak 500 ve

5000 örneklem büyüklükleri arasındaki değişimin diğer örneklem büyüklüklerine göre daha az olduğu Tablo 6'da görülmektedir.

ÇBMTK için  $d$  parametresinin ortalamasına ait SHO'nun evren değeri ise 0,01 olarak kestirilmiştir. Evren değerine en uzak SHO kestiriminin en küçük örneklem büyüklüğünden elde edildiği ve örneklem büyüklüğü arttıkça SHO kestirimlerinin de evren değerine yaklaştığı tabloda görülmektedir.  $d$  parametresi ortalamasına ait SHO'ların örneklem boyunca farklı değerler aldığı yine Tablo 6'da görülmektedir.

Tıpkı ayırtedicilik parametrelerinin ortalamalarında olduğu gibi  $d$  parametresi ortalaması ile  $d$  parametresi ortalamasına ait SHO da örneklemlerden evrene doğru sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanmadığı sonucuna ulaşılmıştır.

TBPMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin ilki ayırtediciliğin göstergesi olan ortalama  $a$  parametreleridir. Örneklemekten kestirilen  $a$  parametresi ortalamasının örneklem büyüklüğüne bağlı olarak 1,23 ile 1,94 arasında değerler aldığı Tablo 6'da görülmektedir.  $a$  parametresi ortalamasının evren değeri 1,52 olup evren değerini en iyi yansıtan örneklem büyüklüğü 200'dür. Daha genel bir ifade ile bu çalışmada kullanılan orta büyüklükteki örneklemelerde evren değere daha yakın kestirimler elde edilirken, uçtaki (en büyük ve en küçük) örneklemelerde evren değere daha uzak kestirimler elde edilmiştir. Ayrıca  $a$  parametresi ortalamasının evren değerinin tek boyutlu testten elde edilen evren değer ile aynı olduğunu da belirtmek gerekir. Buradaki karşılaştırmayı anlamlı hale getiren nokta, karşılaştırılan değerlerin aynı evren üzerinden fakat iki farklı boyutluluğa sahip yapılardan elde edilmesidir. Bu noktadan hareketle ulaşılan sonuç tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden kestirilen  $a$  parametresine ait evren değerlerin ortalamasının değişmediğidir.

TBPMTK için  $a$  parametresi ortalamasına ait SHO'nun evren değeri 0,01 olarak kestirilmiştir. Evren değere büyüklük olarak en uzak SHO 100 örneklem büyüklüğünden elde edilmiştir (SHO=0,75).  $a$  parametresi ortalamasına ait SHO'nun örneklem büyüklüğüne bağlı olarak düzenli bir artma ya da azalma eğilimi gösterdiği söylenemez.  $a$  parametresi ortalamasına ait SHO ardışık örneklem büyüklüklerinin

bazılarında evren değere daha uzak bir değer alırken bazılarında daha yakın bir değer almıştır. Çok boyutlu veri yapısının tek boyutlu bir model altında modellenmesi bu düzensizliğin başlıca sebebidir. a parametresi ortalamasına ait SHO'ların 1000 örneklem büyüklüğüne kadar evreni yansıtmayan değerler aldığı yine tabloda görülmektedir. 500 örneklem büyüklüğünden itibaren ise istikrarlı bir şekilde azalarak evren değere yaklaşmıştır. Bunun yanı sıra a parametresi ortalamasına ait SHO'lar en küçük örneklem büyüklüğünden en büyük örneklem büyüklüğüne kadar sürekli bir değişim göstermektedir.

TBPMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin bir diğeri ise güçlüğün göstergesi olan ortalama b parametreleridir. b parametresi ortalamasının evren değeri 0,30 olarak kestirilmiştir. Evren değerine büyüklük olarak en uzak kestirim örneklem büyüklüğünün 50 olduğu durumda elde edilmiştir (b=-0,82). Bununla birlikte b parametresi ortalamasının evren değerini iyi yansıtan bir örneklem büyüklüğü bulunmamaktadır. a parametresi için ifade edilen "tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden kestirilen a parametresine ait evren değerlerin değişmediği" sonucu b parametresi için de geçerlidir. Bir başka ifadeyle kullanılan veri seti ister tek boyutlu olsun ister çok boyutlu olsun kestirilen b parametrelerine ait evren değerler birbiri ile aynıdır. Elde edilen sonuçlara göre b parametresi ortalamasının örneklem büyüklüğünün değişimine bağlı olarak net bir artma ya da azalma eğiliminin olmadığı görülmektedir.

TBPMTK için b parametresi ortalamasına ait SHO'nun evren değeri 0,01 olup, evren değerine büyüklük olarak en uzak b parametresi ortalamasına ait SHO en küçük örneklem büyüklüğü olan 50 üzerinden kestirilmiştir (SHO= 0,48). Örneklem büyüklüğü arttıkça kestirilen SHO'lar da evren değerine yaklaşmıştır. 500 örneklem büyüklüğü itibariyle evren değerine çok yakın büyüklükte bir SHO ile b parametresinin kestirildiği de Tablo 6'da görülmektedir (SHO=0,19). Tıpkı a parametresi ortalamasında olduğu gibi b parametresi ortalamasına ait SHO'lar da farklı örneklem büyüklüklerinde farklı değerler almıştır.

Sonuç olarak, TBPMTK'den kestirilen madde parametrelerinin ve madde parametrelerine ait SHO'ların incelenmesinin sonucunda, a ve b parametreleri ile a

ve b parametrelerine ait SHO'lar sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

### 3. Alt Problemin Çözümüne İlişkin Bulgular

“Çok boyutlu testlerde parametrik ve parametrik olmayan MTK'ye göre elde edilen madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrenden kestirildiği durumda nasıldır?” şeklinde oluşturulan üçüncü alt problemin çözümü için birleşik testin sonuçlarından TBPMTK ve TBPOMTK'ye göre madde parametreleri ortalaması ve bu parametrelerin ortalamasına ait SHO kestirilmiştir. Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlarda ve evren üzerinden TBPMTK ve TBPOMTK'ye göre kestirilen madde parametrelerine ait ortalama değerleri ile standart hata ortalamaları (SHO) Tablo 7'de görülmektedir.

Tablo 7

*TBPMTK ve TBPOMTK'ye Göre Kestirilen Birleşik Teste Ait Madde Parametreleri ve Standart Hata Ortalamaları*

Örneklem Büyüküklüğü	TBPMTK			TBPOMTK				
	a	a (SHO)	b	b (SHO)	H <sub>i</sub>	H <sub>i</sub> (SHO)	p	p (SHO)
50	1,23	0,54	-0,82	0,48	0,46	0,10	0,65	0,097
100	1,24	0,75	-0,74	0,47	0,38	0,07	0,65	0,096
200	1,48	0,52	-0,79	0,36	0,39	0,05	0,68	0,096
500	1,68	0,69	-0,55	0,19	0,38	0,03	0,65	0,096
1000	1,76	0,16	-0,48	0,09	0,40	0,02	0,65	0,096
5000	1,94	0,07	-0,49	0,03	0,40	0,01	0,66	0,095
Evren	1,52	0,01	0,30	0,01	0,33	0,00	0,43	0,107

Birleşik testin TBPMTK'ye göre analiz edilmesi ile elde edilen sonuçlar ikinci alt problemin çözümünde detaylı bir şekilde ifade edilmiştir. Birleşik testin TBPOMTK'ye göre analiz edilmesiyle elde edilen ilk parametre H<sub>i</sub> parametresidir.

H<sub>i</sub> ortalamasının evren değerinin 0,33 olduğu tabloda görülmektedir (H<sub>i</sub>=0,33). Evren değere en uzak H<sub>i</sub> ortalaması 50 örneklem büyüklüğünden

kestirilmiştir ( $H_i=0,46$ ). Bu durumda tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden kestirilen  $H_i$  parametresine ait evren değerlerin değişmediği sonucuna ulaşılabılır.  $H_i$  ortalaması açısından evreni en iyi yansıtan örneklem büyüklükleri 100 ve 500 örneklem büyüklükleridir.

$H_i$  ortalamasına ait SHO için 0'a çok yakın büyüklükte bir evren değer kestirilmiştir ( $SHO= 0,00$ ). Evren değere büyüklük olarak en uzak SHO en küçük örneklem büyüklüğü olan 50 üzerinden kestirilmiştir ( $SHO=0,10$ ).  $H_i$  ortalamasına ait SHO'ların örneklem büyüklüğü arttıkça istikrarlı bir şekilde azalarak evren değerine yaklaştığı görülmektedir. Bunun yanı sıra örneklemelerden elde edilen SHO'ların evren değerinden büyük farklılıklar göstermediği belirlenmiştir. Bu durumda çok boyutlu bir veri üzerinden küçük örneklem büyüklükleri ile çalışıldığı zaman, evren değere yakın SHO'lar ile ayırteçilik parametresinin kestirilmesinin mümkün olduğu söylenebilir. Ayrıca tek ve çok boyutlu testler TBPOMTK'ye göre analiz edildiğinde elde edilen  $H_i$  parametrelerine ait SHO'ların birbirlerine çok yakın değerler aldıkları görülmüştür. Bu durumda tek boyutluluğun ihlal edildiği durumda  $H_i$  ortalamalarının benzer SHO değerleri ile kestirildiği sonucuna ulaşılabılır. Bununla birlikte SHO'lar örneklem boyunca büyüklük olarak sürekli bir değişim göstermektedir.

Tablo 7'de görüldüğü gibi  $H_i$  ortalamaları ve SHO'ları örneklem boyunca birbirleri ile farklılık gösterdikleri için  $H_i$  ortalaması için de değişmezliğin sağlanamadığı sonucuna ulaşılabılır. Tek boyutluluğun sağlandığı durumda ise küçük bir örneklem büyüklüğünden itibaren değişmezliğin sağlandığı birinci alt problemin bulgularında ifade edilmiştir. Tek boyutlu modellerin kullanılması için sağlanması gereken en önemli varsayım olan tekboyutluluğun ihlali, TBPOMTK sonuçlarını değişmezlik açısından etkilemiştir.

TBPOMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin bir diğeri ise güçlüğün göstergesi olan ortalama  $p$  parametreleridir.  $p$  ortalamalarına göre testin örneklemelerde görel olarak daha kolay olduğu da Tablo 7'de görülmektedir. Bununla birlikte  $p$  ortalamalarının evren değeri ile büyüklük olarak evrene en uzak değeri arasında büyük farklılıkların olmadığı da görülmektedir.  $p$  ortalamasının evren değeri 0,43 olarak kestirilmiştir ( $p=0,43$ ). Büyüklük olarak

evrene en uzak p ortalaması ise 200 örneklem büyüklüğünden kestirilmiştir. Tıpkı Hi ortalamasının evren değerinde olduğu gibi olduğu gibi p ortalamasına ait evren değerinin de tek boyutlu testin TBPOMTK'ye göre analiz edildiğinde kestirilen evren değer ile aynı olduğu görülmüştür. Bu durumda yine tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden elde edilen ortalama p parametresine ait evren değerlerin değişmediği sonucuna ulaşılabilir.

Tablo 7'de görüldüğü gibi p ortalamalarına ait SHO'nun evren değeri 0,107 olarak elde edilmiştir (SHO=0,107). Evren değere en uzak SHO değeri ise 5000 örneklem üzerinden kestirilmiştir (SHO=0,095). Tek boyutlu test TBPOMTK'ye göre analiz edildiğinde evrenden elde edilen p ortalamalarının aynı SHO ile kestirilmesi dikkat çeken bir bulgudur. Buna göre tek boyutluluğun sağlandığı ve ihlal edildiği durumda p ortalamalarına ait SHO'ların aynı olduğu sonucuna ulaşılabilir. p ortalamalarına ait SHO'ların evren değeri ile evrene en uzak değeri arasında büyük farklılıkların olmadığı Tablo 7'de görülmektedir. Bu nedenle örneklem büyüklüğünün değişmesinin p ortalamalarının kestirildiği SHO'lar üzerinde büyük farklılıklara neden olmadığı söylenebilir. Bu durumda tek boyutluluğun sağlanmadığı ve büyük örneklemere ulaşamadığı takdirde evren değere yakın büyüklükteki SHO'lar ile ortalama p parametresinin kestirilmesinin mümkün olduğu sonucuna ulaşılabilir. p ortalamasına ait SHO'lar en büyük ve en küçük örneklem büyüklüklerinde farklı değerler almış, diğer örneklem büyüklüklerinde ise değişmemiştir.

Ayrıca ortalama p değerleri ve SHO'ları en büyük ve en küçük örneklem büyüklüklerinde farklı değerler aldığı için p ortalaması için de değişmezliğin sağlanmadığı sonucuna ulaşılmıştır. Tek boyutluluğun sağlandığı durumda ise küçük bir örneklem büyüklüğünden itibaren değişmezliğin sağlandığı birinci alt problemin bulgularında ifade edilmiştir. Hi ortalaması için de benzer sonuçlar ortaya konmuştur.

TBPMTK ile TBPOMTK'den elde edilen sonuçlar karşılaştırılacak olursa, Hi ortalamalarının a parametresi ortalamalarına göre küçük örneklem büyüklüklerinde bile evren değerine yakın SHO'lar ile kestirildiği tablo 7'de görülmektedir. Tablo 7'den elde edilen sonuçlara göre, a parametresi ortalamasını evren değerine yakın bir SHO ile kestirebilmek için en az 1000 hatta 5000 örneklem büyüklükleri ile

çalışmak gerekmektedir. Daha küçük bir örneklemden evren değerine yakın bir SHO ile ortalama ayırteçilik parametreleri kestirilebildiği için TBPOMTK'nin TBPMTK'ye göre daha avantajlı olduğu sonucuna ulaşılabilir. Tek boyutlu test parametrik ve parametrik olmayan MTK'ye göre analiz edildiğinde de benzer bir sonuca ulaşılmıştır. O halde tek boyutluluk varsayımı karşılanırsa da karşılanmasa da evren değerine yakın büyüklükte bir SHO ile ortalama ayırteçilik parametresi kestirmek için, TBPMTK için büyük örneklemelerin kullanılması gerekmektedir. TBPOMTK için ise böyle bir sınırlama yoktur. Benzer yorumlar parametrik ve parametrik olmayan MTK'den elde edilen ortalama güçlük parametreleri için de geçerlidir. Evren değerine yakın büyüklükte bir SHO ile ortalama b parametresinin kestirilmesi için en az 1000 örneklem ile çalışmak gerekmektedir. Ancak evren değerine yakın bir SHO ile ortalama p parametresinin kestirilmesi için küçük örneklem büyüklükleri ile de çalışılabileceği Tablo 7'de görülmektedir.

#### 4. Alt Problemin Çözümüne İlişkin Bulgular

*“Tek boyutlu testlerde parametrik ve parametrik olmayan MTK'ye göre hesaplanan yetenek kestirimleri arasında, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıl bir ilişki vardır?”* şeklinde oluşturulan dördüncü alt problemin çözümü için matematik testinin sonuçlarından TBPMTK ve TBPOMTK'ye göre sınavı alan bireylere ait yetenek parametreleri kestirilmiştir. Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlar ile evren üzerinden TBPMTK ve TBPOMTK'ye göre kestirilen yetenek parametreleri arasındaki ilişki düzeyleri Tablo 8'de görülmektedir.

Tablo 8

#### *TBPOMTK ve TBPMTK'ye Göre Kestirilen Matematik Testine Ait Yetenek Parametreleri Arasındaki İlişki*

Örneklem Büyüklüğü		PMÇKK TBPOMTK	Spearman TBPOMTK
50	TBPMTK	0,978	0,994
100	TBPMTK	0,978	0,992
200	TBPMTK	0,955	0,981
500	TBPMTK	0,951	0,982
1000	TBPMTK	0,956	0,986
5000	TBPMTK	0,950	0,990
Evren	TBPMTK	0,955	0,991

Tablo 8'de görüldüğü gibi örneklem büyüklüğü ve kullanılan korelasyon türü farketmeksizin, parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki korelasyon katsayıları yüksektir. Bu katsayılara göre parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek düzeyleri arasında oldukça yüksek bir ilişkinin olduğu söylenebilir.

Katsayılar ve örneklem büyüklüğü düzeyinde yorum yapılacak olursa parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki **PMÇKK**, örneklem büyüklüğünün 50 ve 100 olduğu durumlarda en büyük değerini almıştır ( $r_{TBPOMTK \times TBPMTK} = 0,978$ ). Bir başka ifadeyle, parametrik ve parametrik olmayan MTK'ye göre birbirine en fazla benzeyen yetenek parametreleri 50 ve 100 örneklem büyüklüklerinden elde edilmiştir. Bununla birlikte yetenek parametreleri arasındaki ilişkinin evren değeri 0,955'tir. Bu açıdan en küçük iki örneklem büyüklüğünden elde edilen bu PMÇKK değerinin evren değerinden farklı olduğu tablo 8'de görülmektedir. Örneklem büyüklüğünün artmasıyla yetenek parametreleri arasındaki PMÇKK'nin düzenli olarak olmasa da azalma eğilimi olduğu ve evren değerine yakın büyüklükte değerler aldığı görülmektedir. Buradan çıkarılacak sonuç örneklem büyüklüğünün artmasıyla iki kuramdan elde edilen yetenek düzeylerinin birbirleri ile olan benzerliği azalmıştır; ancak evreni daha iyi temsil eden yetenek kestirimleri elde edilmiştir. Evreni en iyi şekilde yansıtan örneklem büyüklüğünün ise 200 olduğu Tablo 8'de görülmektedir.

Parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki **Spearman sıra farkları korelasyon katsayısı**, örneklem büyüklüğü 50 olduğu zaman en büyük değerini almıştır ( $r_{TBPOMTK \times TBPMTK} = 0,994$ ). PMÇKK'de olduğu gibi iki kuramdan birbirine en fazla benzeyen yetenek düzeyleri en küçük örneklem büyüklüğünden elde edilmiştir. Bununla birlikte örneklem büyüklüğünün artmasının elde edilen korelasyon katsayıları üzerinde önemli değişikliklere neden olmadığı Tablo 8'de görülmektedir. Spearman sıra farkları korelasyon katsayısının evren değeri 0,991 olup örneklemelerden elde edilen korelasyon büyüklüklerinden çok farklı değildir. Bu durumda parametrik ve parametrik olmayan MTK'ye göre elde edilen yetenek kestirimleri örneklemeler boyunca birbirine çok benzerdir ve örneklemelerin Spearman sıra farkları korelasyon katsayısına göre evreni temsil gücü oldukça iyidir.



Sonuç olarak parametrik ve parametrik olmayan MTK'ye göre elde edilen yetenek kestirimlerine ilişkin en yüksek ilişki düzeyinin en küçük örneklem büyüklüğünde olduğu söylenebilir. Ancak evreni en iyi yansıtan örneklem büyüklüğünden elde edilen korelasyon büyüklüğü ile en az temsil eden korelasyon büyüklüğü arasında büyük farklılıkların olmadığını belirtmek gerekir. Ayrıca kullanılan Spearmann-Brown sıra farkları korelasyon katsayısına ait ilişki düzeylerinin PMÇKK'ye göre daha yüksek olduğu sonucuna ulaşılmıştır. Veri yapısına bağlı olarak bu durumla karşılaşmak beklenen bir sonuçtur.

## 5. Alt Problemin Çözümüne İlişkin Bulgular

“Çok boyutlu testlerde tek ve çok boyutlu MTK'ye göre hesaplanan yetenek kestirimleri arasında, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıl bir ilişki vardır?” şeklinde oluşturulan dördüncü alt problemin çözümü için birleşik testin sonuçlarından ÇBMTK ve TBPMTK'ye göre sınavı alan bireylere ait yetenek parametreleri kestirilmiştir. Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlar ile evren üzerinden ÇBMTK ve TBPMTK'ye göre kestirilen yetenek parametreleri arasındaki ilişki düzeyleri Tablo 9'da görülmektedir.

Tablo 9

### *TBPMTK ve ÇBMTK'ye Göre Kestirilen Birleşik Teste Ait Yetenek Puanları Arasındaki İlişki*

Örneklem Büüklüğü		PMÇKK		Spearman	
		ÇBMTK 1	ÇBMTK 2	ÇBMTK 1	ÇBMTK 2
50	TBPMTK	0,964	0,956	0,980	0,937
100	TBPMTK	0,965	0,972	0,977	0,944
200	TBPMTK	0,966	0,972	0,984	0,971
500	TBPMTK	0,971	0,956	0,987	0,947
1000	TBPMTK	0,966	0,964	0,979	0,960
5000	TBPMTK	0,969	0,962	0,985	0,954
Evren	TBPMTK	0,978	0,982	0,964	0,972

Tablo 9'da görüldüğü gibi kullanılan korelasyon türü farketmeksizin evrenin bütün örneklemi için tek ve çok boyutlu MTK'ye göre kestirilen birleşik teste ait

yetenek parametreleri arasındaki korelasyon katsayıları çok yüksektir. Örneklemelerden elde edilen katsayılara göre yetenek düzeyleri arasında mükemmele yakın ilişki düzeyleri olduğu görülmektedir.

Korelasyon türlerine göre evren değerleri için yorum yapılacak olursa 1. ve 2. boyut üzerinden kestirilen yetenek parametreleri ile TBPMTK'ye göre elde edilen yetenek parametreleri arasındaki **PMÇKK**'nin evren değeri 0,978 ve 0,982'dir ( $r_{\text{ÇBMTK}(1)\times\text{TBPMTK}} = 0,978$  ,  $r_{\text{ÇBMTK}(2)\times\text{TBPMTK}} = 0,982$ ). Benzer şekilde **Spearman sıra farkları korelasyon katsayısının** evren değerleri ise 0,964 ve 0,972'dir ( $r_{\text{ÇBMTK}(1)\times\text{TBPMTK}} = 0,964$  ,  $r_{\text{ÇBMTK}(2)\times\text{TBPMTK}} = 0,972$ ).

Korelasyon türlerine göre ve örneklem büyüklüğü düzeyinde yorum yapılacak olursa çok boyutlu MTK'nin birinci boyutundan elde edilen ve TBPMTK'ye göre kestirilen yetenek parametreleri arasındaki **PMÇKK ve Spearman sıra farkları korelasyon katsayısı** örneklem büyüklüğünün 500 olduğu durumda en büyük değerini almıştır ( $r_{\text{ÇBMTK}(1)\times\text{TBPMTK}} = 0,971$  ,  $r_{\text{ÇBMTK}(1)\times\text{TBPMTK}} = 0,987$ ). İkinci boyuttan elde edilen yetenek parametreleri ile TBPMTK'ye göre kestirilen yetenek parametreleri arasındaki PMÇKK 100 ve 200 örneklem büyüklüklerinde en yüksek değerini alırken ( $r_{\text{ÇBMTK}(2)\times\text{TBPMTK}}=0,972$  ), Spearman sıra farkları korelasyon katsayısı ise 200 örneklem büyüklüğünde almıştır ( $r_{\text{ÇBMTK}(2)\times\text{TBPMTK}} = 0,971$ ). Bir başka ifadeyle, çok boyutlu MTK'nin birinci boyutundan elde edilen ve TBPMTK'ye göre kestirilen ve birbirine en fazla benzeyen yetenek parametreleri her iki korelasyon türü için de 500 örneklem büyüklüğünden elde edilmiştir. İkinci boyuttan elde edilen ve TBPMTK'ye göre kestirilen ve birbirine en fazla benzeyen yetenek parametreleri ise PMÇKK'ye göre 100 ve 200, Spearman sıra farkları korelasyon katsayısına göre 200 örneklem büyüklüğünden elde edilmiştir. Bununla birlikte PMÇKK'ye göre evreni en iyi temsil eden örneklem en yüksek ilişki düzeyinin elde edildiği örneklemelerdir. Bu bağlamda tek ve çok boyutlu MTK'den elde edilen yetenek parametrelerinin birbirlerine olan bezerliği arttıkça evreni temsil gücünün de arttığı sonucuna ulaşılabilir. Bu sonuç ikinci boyuttan ve TBPMTK'den elde edilen parametreler arasındaki spearman sıra farkları korelasyon katsayısı için de geçerlidir. Ancak Spearman sıra farkları korelasyon katsayısına göre birinci boyuttan ve TBPMTK'den elde edilen parametreler arasındaki korelasyon katsayısı ne kadar küçükse evreni okadar iyi temsil etmektedir. Buyüzdten evreni en iyi temsil eden örneklem büyüklüğünün 100 olduğu söylenebilir.

Sonuç olarak tek ve çok boyutlu MTK'ye göre elde edilen yetenek kestirimleri arasında yüksek korelasyonlar elde edilmiştir. Evreni en iyi yansıtan örneklem büyüklüğünden elde edilen korelasyon büyüklüğü ile en az temsil eden korelasyon büyüklüğü arasında büyük farklılıkların olmadığı tabloda görülmektedir. Ayrıca kullanılan Spearman-Brown sıra farkları korelasyon katsayısına ait ilişki düzeylerinin PMÇKK'ye göre daha yüksek olduğu sonucuna ulaşılmıştır. Veri yapısına bağlı olarak bu durumla karşılaşmak beklenen bir sonuçtur.

## 6. Alt Problemin Çözümüne Yönelik İlişkin Bulgular

“Çok boyutlu testlerde parametrik ve parametrik olmayan MTK'ye göre hesaplanan yetenek kestirimleri arasında, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıl bir ilişki vardır?” şeklinde oluşturulan altıncı alt problemin çözümü için birleşik testin sonuçlarından TBPMTK ve TBPOMTK'ye göre sınavı alan bireylere ait yetenek parametreleri kestirilmiştir. Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlar ile evren üzerinden TBPMTK ve TBPOMTK'ye göre kestirilen yetenek parametreleri arasındaki ilişki düzeyleri Tablo 10'da görülmektedir.

Tablo 10

### *TBPMTK ve TBPOMTK'ye Göre Kestirilen Birleşik Teste Ait Yetenek Puanları Arasındaki İlişki*

Örneklem Büyüklüğü		PMÇKK	Spearman-Brown
		TBPOMTK	TBPOMTK
50	TBPMTK	0,993	0,996
100	TBPMTK	0,989	0,994
200	TBPMTK	0,976	0,992
500	TBPMTK	0,968	0,993
1000	TBPMTK	0,964	0,992
5000	TBPMTK	0,954	0,994
Evren	TBPMTK	0,955	0,991

Tablo 10'da görüldüğü gibi örneklem büyüklüğü ve kullanılan korelasyon türü farketmeksizin, parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki korelasyon katsayıları yüksektir. Bu katsayılara göre parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri

arasında oldukça yüksek bir ilişkinin olduğu söylenebilir. Ayrıca parametrik ve parametrik olmayan MTK'ye göre matematik testi ile birleşik testten elde edilen yetenek düzeyleri arasında büyük farklılıkların olmadığı görülmektedir.

Katsayılar ve örneklem büyüklüğü düzeyinde yorum yapılacak olursa parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki **PMÇKK**, örneklem büyüklüğünün 50 olduğu durumda en büyük değerini almıştır ( $r_{TBPOMTK \times TBPMTK} = 0,993$ ). Bir başka ifadeyle, parametrik ve parametrik olmayan MTK'ye göre birbirine en fazla benzeyen yetenek parametreleri 50 örneklem büyüklüğünden elde edilmiştir. Tek boyutlu olan matematik testinin parametrik ve parametrik olmayan MTK'ye göre kestiriminden elde edilen yetenek düzeyleri arasındaki **PMÇKK** için de benzer bir sonuç elde edilmiştir. Örneklem büyüklüğünün artmasına bağlı olarak iki kuramdan elde edilen yetenek düzeyleri arasındaki ilişkinin düzenli olarak azaldığı görülmüştür. Yetenek parametreleri arasındaki ilişkinin evren değeri ise 0,955'tir. Bu açıdan örneklemelerden elde edilen **PMÇKK** değerlerinin evren değerinden görece olarak farklı olduğu söylenebilir. Buradan çıkarılacak sonuç örneklem büyüklüğünün artmasıyla iki kuramdan elde edilen yetenek düzeylerinin birbirleri ile olan benzerliği az da olsa azalmıştır. Ancak evreni daha iyi temsil eden yetenek kestirimleri elde edilmiştir. Tek boyutlu olan matematik testinin parametrik ve parametrik olmayan MTK'ye göre kestiriminden elde edilen yetenek düzeyleri arasındaki **PMÇKK** için de benzer bir sonuç elde edilmiştir. Evreni en iyi şekilde yansıtan örneklem büyüklüğü ise 5000'dir.

Parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki **Spearman sıra farkları korelasyon katsayısı**, örneklem büyüklüğü 50 olduğu zaman en büyük değerini almıştır ( $r_{TBPOMTK \times TBPMTK} = 0,996$ ). **PMÇKK**'de olduğu gibi iki kuramdan birbirine en fazla benzeyen yetenek düzeyleri en küçük örneklem büyüklüğünden elde edilmiştir. Bununla birlikte örneklem büyüklüğünün artmasının elde edilen korelasyon katsayıları üzerinde önemli değişikliklere neden olmadığı Tablo 10'da görülmektedir. Spearman sıra farkları korelasyon katsayısının evren değeri 0,991 olup örneklemelerden elde edilen korelasyon büyüklüklerinden çok farklı değildir. Bu durumda parametrik ve parametrik olmayan MTK'ye göre elde edilen yetenek kestirimleri örneklem boyunca birbirine çok benzerdir ve örneklemelerin Spearman sıra farkları korelasyon

katsayısına göre evreni temsil gücü oldukça iyidir. Tek boyutlu olan matematik testinin parametrik ve parametrik olmayan MTK'ye göre kestiriminden elde edilen yetenek düzeyleri arasındaki Spearman sıra farkları korelasyon katsayısı için de benzer sonuçlar elde edilmiştir.

Sonuç olarak parametrik ve parametrik olmayan MTK'ye göre elde edilen yetenek kestirimlerine göre en yüksek ilişki düzeyinin en küçük örneklem büyüklüğünde olduğu söylenebilir. Ayrıca evreni en iyi yansıtan örneklem büyüklüğünden elde edilen korelasyon büyüklüğü ile en az temsil eden korelasyon büyüklüğü arasında büyük farklılıkların olmadığını belirtmek gerekir. Bununla birlikte kullanılan farklı korelasyon türlerine göre Spearmann sıra farkları korelasyon katsayısına ait ilişki düzeylerinin PMÇKK'ye göre daha yüksek olduğu sonucuna ulaşılmıştır. Veri yapısına bağlı olarak bu durumla karşılaşmak beklenen bir sonuçtur. Tek boyutlu testten elde edilen yetenek kestirimleri arasındaki ilişki düzeylerinin karşılaştırılması ile de benzer sonuçlara ulaşılmıştır. Bu durumda tek boyutluluk varsayımının ihlal edilmesinin iki kuramdan elde edilen yetenek kestirimlerini benzer şekilde ekilediği sonucuna ulaşılabilir. Bir başka ifadeyle tek boyutluluk varsayımı karşılanırsa da ihlal edilse de iki kuramdan benzer yetenek kestirimleri elde edilmiştir.

## **7. Alt Problemin Çözümüne İlişkin Bulgular**

*“Tek ve çok boyutlu testlerde TBPMTK, TBPOMTK ve ÇBMTK'ye göre hesaplanan güvenilirlik düzeyleri örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduğunda ve evrende nasıldır?”* şeklinde oluşturulan yedinci alt problemin çözümü için tek boyutlu olan matematik testi tek boyutlu parametrik ve parametrik olmayan MTK'ye göre modellenerek güvenilirlik analizi yapılmıştır. Bununla birlikte iki boyutlu olan birleşik testin tek boyutlu gibi düşünülmesiyle, tek boyutlu parametrik ve parametrik olmayan MTK'ye göre güvenilirlik analizi yapılmıştır. Ayrıca iki boyutlu olan birleşik testin güvenilirliği ÇBMTK kapsamında da elde edilmiştir. Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlar ile evren üzerinden TBPMTK ve TBPOMTK'ye göre kestirilen matematik testine ait güvenilirlik katsayıları tablo 11'de görülmektedir.

Tablo 11

*TBPMTK ve TBPOMTK 'ye Göre Matematik Testi için Kestirilen Güvenirlik Katsayıları*

Örneklem Büyüklüğü	Marjinal	H
	Güvenirlik TBPMTK	TBPOMTK
50	0,81	0,87
100	0,84	0,89
200	0,84	0,87
500	0,85	0,87
1000	0,86	0,87
5000	0,86	0,87
Evren	0,86	0,87

Tablo 11'de tek boyutlu olan matematik testine ait TBPMTK ve TBPOMTK'ye göre kestirilen güvenirlik düzeylerinin yüksek olduğu görülmektedir. TBPMTK'ye göre kestirilen marjinal güvenirlik en düşük değerini örneklem büyüklüğünün 50 olduğu durumda almıştır ( $r_{\text{marjinal}}=0,81$ ). 50 örneklem büyüklüğünden 1000 örneklem büyüklüğüne kadar örneklem büyüklüğü arttıkça marjinal güvenirlik de istikrarlı bir şekilde artmıştır. 1000 örneklem büyüklüğü itibariyle ise evreni yansıtacak düzeyde kararlı bir yapıya geçmiştir ( $r_{\text{marjinal}}=0,86$ ). Bir başka ifadeyle 1000 örneklem büyüklüğünden sonra TBPMTK için güvenirlik katsayıları açısından değişmezliğin sağlandığı tablo 11'de görülmektedir.

TBPOMTK için elde edilen H katsayıları incelendiğinde H katsayısının evren değerinin 0,87 olduğu Tablo 11'de görülmektedir. Bu durumda sadece 100 örneklem büyüklüğünden elde edilen H katsayısının evren değerden farklı olduğu söylenebilir. Bir başka ifadeyle en küçük örneklem büyüklüğünden 200 örneklem büyüklüğüne kadar elde edilen H katsayıları az da olsa değişim göstermiştir. 200 örneklem büyüklüğünden itibaren ise değişmezliğin sağlandığı yine Tablo 11'de görülmektedir.

Örneklem büyüklüğünün 50, 100, 200, 500, 1000 ve 5000 olduğu durumlar ile evren üzerinden ÇBMTK, TBPMTK ve TBPOMTK'ye göre kestirilen birleşik teste ait güvenirlik katsayıları Tablo 12'de görülmektedir.

Tablo 12

*Birleşik Test için TBPMTK, TBPOMTK ve ÇBMTK'ye Göre Kestirilen Güvenirlik Katsayıları*

Örneklem Büüklüğü	Klasik Güvenirlik		Marjinal Güvenirlik	H
	ÇBMTK 1	ÇBMTK 2	TBPMTK	TBPOMTK
50	0.77	0.74	0.81	0,87
100	0.80	0.81	0.84	0,89
200	0.77	0.75	0.84	0,87
500	0.77	0.77	0.85	0,87
1000	0.76	0.77	0.86	0,87
5000	0.76	0.76	0.86	0,87
Evren	0.76	0.77	0.86	0,87

Tablo 12'de görüldüğü gibi ÇBMTK'ye göre her bir boyut için farklı bir güvenirlik katsayısı kestirilmiştir. Birinci ve ikinci boyuttan elde edilen güvenirlik katsayılarının örneklem büyüklüğünün değişimine bağlı olarak net bir artma ya da azalma eğilimi göstermediği tabloda görülmektedir. Bununla birlikte örneklem büyüklüğünün her iki boyuta ait güvenirlik katsayıları üzerinde önemli değişimlere neden olmadığı da tabloda görülmektedir. Birinci boyuta ait güvenirlik katsayısının evren değeri 0,76'dır. Buna göre birinci boyuta ait güvenirlik katsayısının 1000 örneklem büyüklüğünden itibaren evreni yansıtacak düzeyde kararlı bir yapıya sahip olduğu söylenebilir. İkinci boyuta ait güvenirlik katsayısının evren değeri ise 0,77'dir. Hernekadar bazı örneklemelerden (500 ve 1000 örneklem büyüklükleri) elde edilen güvenirlik katsayıları evren değeri ile aynı olsa da ikinci boyuttan kestirilen güvenirlik katsayısı için değişmezliğin sağlandığı bir örneklem büyüklüğünün olmadığı Tablo 12'de görülmektedir.

Tablo 12'de görüldüğü gibi TBPMTK'ye göre kestirilen marjinal güvenirlik en düşük değerini örneklem büyüklüğünün 50 olduğu durumda almıştır ( $r_{\text{marjinal}}=0,81$ ). 50 örneklem büyüklüğünden 1000 örneklem büyüklüğüne kadar örneklem büyüklüğü arttıkça marjinal güvenirlik de istikrarlı bir şekilde artmıştır. Marjinal güvenirliğe ait evren değeri 0,86 olduğu tabloda görülmektedir. Bu durumda 1000 örneklem büyüklüğünden itibaren TBPMTK için değişmezliğin sağlandığı sonucuna ulaşılabilir.

Son olarak TBPOMTK için elde edilen H katsayıları incelendiğinde H katsayısının evren değerinin 0,87 olduğu Tablo 12’de görülmektedir. Bu durumda TBPMTK için 200 örneklem büyüklüğünden itibaren değişmezliğin sağlandığı sonucuna ulaşılabılır.

Elde edilen güvenilirlik katsayılarının örneklem büyüklüklerine göre fazla değişim göstermediği üç kuram için de ulaşılan önemli bir sonuçtur. Ayrıca Tablo 11 ve Tablo 12’de görülen H ve marjinal güvenilirlik katsayıları kendi aralarında karşılaştırıldığı zaman, her bir örneklem büyüklüğü için kestirilen H katsayısı ve marjinal güvenilirlik katsayılarının bire bir aynı olduğu görülmektedir. Buradan hareketle tek boyutluluk varsayımı karşılanırsa da karşılanmasa da parametrik ve parametrik olmayan tek boyutlu MTK’ye göre kestirilen güvenilirlik katsayıları değişmemiştir, sonucuna ulaşılabılır.



## Bölüm 5

### Sonuç, Tartışma Ve Öneriler

Birinci alt problemin bulgularından elde edilen sonuçlar doğrultusunda tek boyutlu testin TBPOMTK'ye göre analiz edilmesiyle ortalama  $p$  ve ortalama  $p'$ 'ye ait SHO ile ortalama  $H_i$  ve ortalama  $H_i'$ 'ye ait SHO'nun örneklem büyüklüğüne bağlı olarak büyük değişimler göstermediği sonucuna ulaşılmıştır. Bununla birlikte 200 örneklem büyüklüğü itibariyle madde parametresi değişmezliği sağlanmıştır. Büyük örneklem büyüklüklerine ulaşamadığı durumda ortalama madde parametresi kestirimi amacıyla küçük örneklemelerden elde edilen veriler için TBPOMTK uygulamasından yararlanılabilir. Bu durum, MTK'nin özellikle okul uygulamalarında eğitimin farklı kademelerinde uygulanan sınavlara ait madde parametresi kestirimlerine imkân sağlaması açısından büyük bir önem taşımaktadır.

Tek boyutlu testin TBPMTK'ye göre analiz edilmesiyle TBPMTK için örneklem büyüklüğü arttıkça  $a$  parametresi ortalamasına ait SHO'nun azaldığı ve evren değerine yaklaştığı görülmüştür. Thissen ve Wainer (1982) parametre kestirimi için 10.000 ve daha fazla örnekleme ihtiyaç olduğunu, Goldman ve Raju (1986) ise  $a$  parametresinin doğru kestirimi için en az 1000 kişilik örnekleme ihtiyaç duyulduğunu ifade etmişlerdir. Örneklem büyüklüğü 5000'i geçtiği halde  $a$  parametresine ait ortalama değer ve  $a$  parametresi ortalamasına ait standart hata ortalamasının değişmeye devam etmesi, Thissen ve Wainer (1982) bulgularını destekler niteliktedir.

Tek boyutlu testin TBPMTK'ye göre analiz edilmesiyle elde edilen  $b$  parametresi ortalamasına ait SHO örneklem büyüklüğü arttıkça azalma eğilimi göstermiş ve evren değerine yaklaşmıştır.  $b$  parametresi ortalaması ise örneklem büyüklüğüne bağlı olarak düzenli bir artma ya da azalma eğilimi göstermemiştir ki Sünbül (2011) bu durumu "örneklem büyüklüğünün  $b$  parametresi üzerindeki önemsizliği" şeklinde ifade etmiştir. Ancak örneklem büyüklüğü azaldıkça az miktarda da olsa  $b$  parametresine ait SHO'nun arttığını da ortaya koymuştur. Ayrıca 1000 örneklem büyüklüğü itibariyle  $b$  parametresi ortalamasına ait değişimlerin daha az olduğu söylenebilir. Hulin, Lissak ve Drasgow (1982) 2 PLM için 1000'den daha büyük örneklemelerde kestirilen parametrelerin önemli değişikliklerin

olmadığını ortaya koymuştur. Bu nedenle elde edilen bulgular Hulin, Lissak ve Drasgow (1982)'nin bulgularıyla tutarlılık göstermektedir.

Sonuç olarak küçük örneklem büyüklüklerinden elde edilen madde parametrelerinin evreni yansıtacak düzeyde kararlı olması, bir başka ifadeyle küçük bir örneklem büyüklüğünden itibaren parametre değişmezliğinin sağlanması TBPOMTK'nin TBPMTK'ye göre önemli avantajlara sahip olduğunun bir kanıtıdır. Bu nedenle özellikle okul uygulamalarında TBPOMTK uygulamaları TBPMTK'ye tercih edilebilir.

İkinci alt problemin bulgularından elde edilen sonuçlar doğrultusunda birleşik testin ÇBMTK'ye göre analiz edilmesiyle elde edilen  $a_1$  ve  $a_2$  parametreleri ortalamasının, örneklem büyüklüğüne bağlı olarak düzenli bir artma ya da azalma eğilimi göstermediği sonucuna ulaşılmıştır. Ackermann (2005) ise örneklem büyüklüğü arttıkça maddelerin ayırtecilik gücünün arttığı sonucuna ulaşmıştır. Bu açıdan çalışmadan elde edilen sonuç, Ackermann (2005)'in elde ettiği sonuç ile farklılık göstermektedir. Ayrıca 100 örneklem büyüklüğü itibarıyla elde edilen  $a_2$  parametresi ortalamalarının evren değerine çok yakın olduğu görülmüştür. Ancak  $a_1$  parametresi ortalamasının evren değerini iyi yansıtan bir örneklem büyüklüğü bulunmamaktadır. Her iki ayırtecilik parametresi ortalaması ile ortalamalara ait SHO da örneklemelerden evrene doğru sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

ÇBMTK için  $d$  parametresinin ortalamasına ait SHO'nun, örneklem büyüklüğünün artmasıyla azalma eğilimi gösterdiği belirlenmiştir. SHO'nun en küçük değerinin evren değerine eşit olduğu ortaya konmuştur. SHO'ların yanı sıra örneklemelerden kestirilen  $d$  parametresi ortalamaları incelendiğinde evren değerini yansıtan bir örneklem büyüklüğünün olmadığı görülmektedir. Elde edilen sonuçlara göre  $d$  parametresi ortalamasının örneklem büyüklüğünün değişimine bağlı olarak net bir artma ya da azalma eğiliminin olmadığı görülmektedir.  $d$  parametresi ortalaması ile  $d$  parametresi ortalamasına ait SHO da örneklemelerden evrene doğru sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

İki boyutlu olan birleşik testin TBPMTK altında modellenmesi ile elde edilen sonuçlar incelendiğinde ayırtecilikparametresi ortalamasının ÇBMTK'den elde edilen değerlere göre artış gösterdiği belirlenmiştir. Birleşik test ÇBMTK'ye göre analiz edildiğinde her bir boyut için elde edilen a parametrelerinin ortalamasının ikisi de (0,1) aralığında değerler almıştır. TBPMTK'ye göre kestirilen ortalama ayırtecilikparametreleri ise 1'in üzerinde değerler almıştır. Elde edilen bu sonuç Ansley ve Forsyth (1985)'in, iki kuramdan birbirine yakın büyüklükte a parametresi kestirimini elde ettiği çalışmasında ortaya koyduğu sonuçlardan farklılık göstermektedir.

TBPMTK için a ve b parametresi ortalamasına ait SHO'nun, örneklem büyüklüğünün artmasına bağlı olarak istikrarlı bir şekilde azaldığı ve evren değerine yaklaştığı görülmüştür. a parametresi ortalamasının 1000 örneklem büyüklüğü itibariyle, b parametresi ortalamasının ise 500 örneklem büyüklüğü itibariyle daha düşük ve evren değerine daha yakın SHO'lar ile kestirildiği sonucuna ulaşılmıştır. Ayrıca a ve b parametresi ortalamasına ait SHO'ların en düşük değerinin evren değerine eşit olduğu ortaya konmuştur. a ve b parametresi ortalamaları ve a ve b ortalamalarına ait SHO'lar bütün örneklerde farklı değerler aldığı için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

Sonuç olarak çok boyutlu bir veri ister TBPMTK altında modellenmiş ister ÇBMTK altında modellenmiş evren değerine yakın bir büyüklükte SHO ile ortalama bir parametre kestirimi yapmak için en az 5000 örneklem büyüklüğünün kullanılması gerekir. 5000 ya da daha büyük örneklem kullanılması durumunda kestirilen SHO evren değerine yaklaşmış olur. Ancak her ne kadar parametre ortalamalarına ait SHO evren değerine yaklaşırsa da çok boyutlu bir verinin tek boyutlu modele göre analiz edilmesi ile elde edilen parametre ortalamaları ile çok boyutlu modele göre analiz edildiğinde elde edilen parametre ortalamaları örneklem büyüklüğüne bağlı olmaksızın birbirinden farklı çıkmıştır. Elde edilen bu sonuç tek boyutluluk varsayımı sağlanmadığı zaman parametre kestirimlerine ait sonuçların farklılaştığını ifade eden Drasgow ve Parsons (1983)'ün bulgularıyla tutarlılık göstermektedir.

Burada değişmezliğin sağlanamamasının ve parametre ortalamalarının örneklem büyüklüğüne bağlı olarak net bir artma ya da azalma göstermemesinin nedeni bu çalışmada örneklem seçiminde replikasyon yapılmaması olabilir.

Replikasyon yapıldığında her bir örnekleme çok sayıda aynı birey bulunabilir ve bu bireyler parametre değişmezliğini şışirebilir bir başka ifadeyle değişmezliğin sağlanmasına otomatik neden olurlar veya kestirimlerin yanlı olmasına neden olabilirler. Bu nedenle bu çalışmada replikasyon yapılmamıştır.

Üçüncü alt problemin bulgularından elde edilen sonuçlar doğrultusunda birleşik testin TBPOMTK'ye göre analiz edilmesiyle elde edilen Hi ortalamasına ait SHO'ların örneklem büyüklüğü arttıkça istikrarlı bir şekilde azalarak evren değerine yaklaştığı sonucuna ulaşılmıştır. Koğar (2014) de çalışmasında örneklem büyüklüğü arttıkça Hi katsayılarına ait standart hatanın azaldığını ortaya koymuştur. Elde edilen bu sonuç Koğar (2014)'ün bulgularıyla tutarlılık göstermektedir. Bunun yanı sıra örneklem büyüklüğü kaç olursa olsun SHO'lar büyüklük olarak evrene çok yakın değerler almıştır. Bu durumda tek boyutluluğun ihlal edilmesinin Hi ve p ortalamalarının kestirildiği SHO değerlerinin üzerinde önemli bir etkisinin olmadığı sonucuna ulaşılmıştır. Bu durumda büyük örneklemere ulaşamadığı takdirde evren değere yakın SHO'lar ile madde parametresi ortalamasının kestirilmesinin mümkün olduğu sonucuna ulaşılmıştır. SHO'ların yanı sıra tek boyutluluğun ihlal edildiği durumda elde edilen Hi ve p ortalamasının evren değeri ile tek boyutluluğun sağlandığı durumda elde edilen evren değeri aynıdır. Ancak örneklemelerden elde edilen Hi ve p ortalamaları birbirinden farklıdır.

Tek boyutluluk varsayımının ihlal edilmesinin parametre kestirimleri üzerindeki en önemli etkisi, tek boyutluluk varsayımı karşılandığı durumda parametre değişmezliği sağlanırken, varsayım karşılanmadığında parametre değişmezliğinin de sağlanamaması olmuştur. Bu nedenle bu çalışmada Koğar (2014)'ün örneklem büyüklüğü arttıkça p değerleri için değişmezliğin sağlandığı bulgusuna ulaşamamıştır. Koğar (2014)'ün elde ettiği bu sonuç, 1. alt problemde elde edilen bulgularda ifade edildiği gibi tek boyutlu veriye ait p ortalamaları için elde edilmiştir. Bir başka ifadeyle tek boyutlu veriye ait p ortalamaları için örneklem büyüklüğü arttıkça parametre değişmezliği sağlanmıştır.

TBPMTK ile TBPOMTK'den elde edilen sonuçlar karşılaştırıldığında TBPOMTK'den kestirilen parametre ortalamalarının evren değerine yakın bir SHO ile kestirilebildiği sonucuna ulaşılmıştır. TBPMTK'den elde edilen parametre ortalamalarının evren değerine yakın bir SHO ile kestirilebilmesi için ise en az 1000

örneklem büyüklüğü ile çalışmak gerekmektedir. Daha küçük bir örneklemden evren değerine yakın bir SHO ile madde parametreleri kestirilebildiği için TBPOMTK'nin TBPMTK'ye göre daha avantajlı olduğu sonucuna ulaşılmıştır. Tek boyutluluk varsayımın karşılanıp karşılanmamasının bu sonucu değiştirmedeği dikkat çeken bir bulgudur.

Dördüncü alt problemin bulgularından elde edilen sonuçlar doğrultusunda parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki PMÇKK'ye göre birbirine en fazla benzeyen yetenek parametrelerinin 50 ve 100 örneklem büyüklüklerinden elde edildiği görülmüştür. Örneklem büyüklüğünün artmasının iki kuramdan elde edilen yetenek düzeylerinin benzerliğini azaltacak yönde bir etkisi olmuştur. Bununla birlikte örneklem büyüklüğünün artmasıyla evreni daha iyi temsil eden yetenek kestirimleri elde edilmiştir. Parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek düzeyleri arasındaki Spearman sıra farkları korelasyon katsayısına göre birbirine benzer yetenek düzeylerinin elde edildiği görülmüş ve örneklemelerin evreni oldukça iyi temsil ettiği sonucuna ulaşılmıştır.

Beşinci alt problemin bulgularından elde edilen sonuçlar doğrultusunda örneklemelerden ve her iki kuramdan elde edilen yetenek düzeyleri arasında korelasyon türü ve örneklem büyüklüğü farketmeksizin mükemmel yakın ilişki düzeyleri olduğu görülmüştür. Bu durumda tek boyutluluk varsayımı karşılanmadığında çok boyutlu bir testin tek ve çok boyutlu MTK'ye göre analiz edilmesiyle örneklemelerden elde edilen yetenek parametrelerinin birbirine benzer oldukları görülmüştür. Her iki kurama göre evrenden kestirilen yetenek parametreleri arasında da tıpkı örneklemelerde olduğu gibi mükemmel yakın bir ilişki vardır. Her ne kadar Chang (1992) çok boyutlu MTK'den elde edilen yetenek kestirimleri tek boyutlu MTK'den elde edilen yetenek kestirimlerine göre daha büyük bir ölçme kesinliğine sahip olduğunu ifade etmiş olsa da tek boyutlu model altında çok boyutlu bir veriye ait yetenek parametrelerinin çok boyutluluk kural ve esaslarına göre elde edilen parametrelerle yüksek ilişki düzeylerine sahip olduğu ortaya konmuştur. Ayrıca her iki kuramdan da küçük örneklemeler ile bile evren parametrelerine benzer yetenek parametrelerinin kestirilebildiği görülmüştür.

Altıncı alt problemin bulgularından elde edilen sonuçlar doğrultusunda parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki PMÇKK'nin, örneklem büyüklüğünün 50 olduğu durumda en büyük değerini aldığı ve dolayısıyla birbirine en fazla benzeyen yetenek parametrelerinin 50 örneklem büyüklüğünden elde edildiği görülmüştür. Örneklem büyüklüğünün artmasına bağlı olarak iki kuramdan elde edilen yetenek düzeyleri arasındaki ilişkinin düzenli olarak azaldığı ve evreni daha iyi temsil eden yetenek kestirimlerinin elde edildiği görülmüştür. Parametrik ve parametrik olmayan MTK'ye göre kestirilen yetenek parametreleri arasındaki Spearman sıra farkları korelasyon katsayısı için birbirine en fazla benzeyen yetenek düzeylerinin en küçük örneklem büyüklüğünden elde edildiği sonucuna ulaşılmıştır. Ancak örneklem büyüklüğünün değişmesinin elde edilen korelasyon katsayıları üzerinde büyük değişikliklere neden olmadığı görülmüştür. Bir başka ifadeyle parametrik ve parametrik olmayan MTK'ye göre elde edilen yetenek kestirimleri açısından örneklemelerin Spearman sıra farkları korelasyon katsayısına göre evreni temsil gücünün oldukça iyi olduğu sonucuna ulaşılmıştır. Dördüncü ve altıncı alt problemlerden elde edilen bulguların karşılaştırılması ile ulaşılan en önemli sonuç tek boyutluluk varsayımının ihlal edilmesinin iki kuramdan elde edilen yetenek kestirimlerini benzer şekilde etkilemiş olmasıdır. Bir başka ifadeyle tek boyutluluk varsayımı karşılanırsa da ihlal edilse de iki kuramdan benzer yetenek kestirimleri elde edildiği sonucuna ulaşılmıştır.

Yedinci alt problem kapsamında birleşik testin sonuçlarından elde edilen bulgular doğrultusunda örneklem büyüklüğünün ÇBMTK'ye göre elde edilen birinci ve ikinci boyuta ait güvenilirlik katsayıları üzerinde önemli değişimlere neden olmadığı sonucuna ulaşılmıştır. Bununla birlikte birinci boyuta ait güvenilirlik katsayısının 1000 örneklem büyüklüğünden itibaren evreni yansıtacak düzeyde kararlı bir yapıya sahip olduğu görülmüştür. İkinci boyuttan elde edilen güvenilirlik katsayısı için ise değişmezliğin sağlanamadığı sonucuna ulaşılmıştır.

Tek ve çok boyutlu testin TBPMTK'ye göre analiz edilmesiyle elde edilen marjinal güvenilirlik katsayıları, 50 örneklem büyüklüğünden 1000 örneklem büyüklüğüne kadar örneklem büyüklüğü arttıkça istikrarlı bir şekilde artma eğilimi göstermiştir. 1000 örneklem büyüklüğünden itibaren ise TBPMTK için değişmezliğin sağlandığı sonucuna ulaşılmıştır. Tek ve çok boyutlu testin TBPOMTK'ye göre analiz edilmesiyle elde edilen H katsayıları için ise 200 örneklem büyüklüğünden

itibaren deęişmezlięin saęlandığı sonucuna ulaşılmıştır. Ayrıca her bir örneklem büyüklüğü için tek ve çok boyutlu testler üzerinden kestirilen H katsayısı ve marjinal güvenirlik katsayılarının bire bir aynı olduęu görülmüştür. Buna göre tek boyutluluk varsayımının karşılanıp karşılanmamasının parametrik ve parametrik olmayan tek boyutlu MTK'ye göre kestirilen güvenirlik katsayıları üzerinde önemli bir etki yaratmadığı sonucuna ulaşılmıştır.

### **Araştırma Sonuçlarına İlişkin Öneriler**

Tek boyutlu bir veriye ait ortalama madde parametrelerini, tek boyutlu parametrik modeller üzerinden evren değerine yakın büyüklükte kestirmek için en az 1000 örneklem büyüklüğü ile çalışmak gerekir. Ancak büyük örneklemere ulaşamadığı durumlarda küçük örneklem büyüklükleri ile de evren değeri temsil edebilen madde parametrelerinin kestirilmesi için tek boyutlu parametrik olmayan modellerden yararlanılabilir.

İki boyutlu bir veri seti, ister çok boyutlu bir kuram altında modellensin, ister tek boyutlu bir kuram altında modellensin iki boyutlu veriye ait madde parametrelerini evren değerine yakın bir büyüklükte kestirmek için en az 1000'in ve hatta 5000'in üzerinde örneklemle çalışılması önerilir.

İki boyutlu bir veriye ait madde parametrelerini, tek boyutlu parametrik modeller üzerinden evren değerine yakın bir büyüklükte kestirmek için en az 1000 örneklem büyüklüğü ile çalışmak gerekir. Ancak büyük örneklemere ulaşamadığı durumlarda küçük örneklem büyüklükleri ile de evren değerine yakın büyüklüklerde maddde parametrelerinin kestirilmesi için parametrik olmayan modellerden yararlanılabilir.

TBPOMTK ve TBPMTK'ye göre tek boyutluluğun saęlandığı ve ihlal edildiği durum ile aynı SHO'ya sahip ve aynı büyüklükte parametre kestirimi yapmak için evren üzerinden madde parametrelerinin kestirilmesi önerilebilir.

Tek boyutlu bir veriye ait yetenek düzeyleri kullanılan model parametrik olsa da olmasa da en küçük örneklemde elde edildiği durumda bile evreni yansıtabilecek düzeyde kestirilmiştir. Bu nedenle büyük örneklemere ulaşamadığı zaman tek

boyutlu bir veri üzerinden yetenek kestirimi yapmak amaçlanıyorsa tek boyutlu parametrik olmayan modellerden yararlanılabilir.

Tek boyutluluğun ihlal edildiği ve verinin TBPMTK altında modellenmesi gerektiği durumda ÇBMTK'ye yakın yetenek parametresi kestirimleri için küçük örneklem büyüklükleri ile de çalışılabilir.

### **Araştırmacılara Yönelik Öneriler**

Bu çalışmada örneklem büyüklüğünün etkisini incelemek için 50, 100, 200, 500, 1000 ve 5000 örneklem büyüklükleri ile çalışılmıştır. Benzer bir çalışma farklı örneklem büyüklükleri ile de gerçekleştirilebilir.

Bu çalışmada TBPMTK için 2PLM, TBPOMTK için MHM ve ÇBMTK için M2PLM kullanılmıştır. Sonuçları incelenen kuramlara ait olan farklı modeller ile benzer bir çalışma yapılabilir.

Bu çalışmada kullanılan modeller üzerinden kestirilen parametre değerleri tüm teste ait ortalama parametre değerleridir. Benzer bir çalışma bir testte yer alan maddelere ait parametre kestirimleri için de gerçekleştirilebilir.

Bu çalışmada örneklem büyüklüğünün etkisini araştırmak amacıyla parametre kestirimlerine ait standart hata ortalaması kullanılmıştır. Başka bir çalışmada örneklem büyüklüğünün etkisini araştırmak için farklı göstergelerden de yararlanılabilir.

Bu çalışmada şans başarısının etkisi göz önünde bulundurulmamıştır. Şans başarısından arındırılmış gerçek puanlar üzerinden de başka bir çalışma yapılabilir.

Bu çalışmada örneklem büyüklüğünün parametre değişmezliğinin üzerindeki etkisi araştırılmıştır. Farklı bağımsız değişkenler ile de benzer bir çalışma yapılabilir.

Bu çalışmada farklı kuramlardan elde edilen yetenek düzeyleri arasındaki ilişkiyi belirlemek için PMÇKK ve Spearman Sıra Farkları korelasyon katsayısından



yararlanılmıştır. Başka bir çalışmada farklı korelasyon katsayılarından da yararlanılabilir.

Bu çalışmada her bir kurama göre elde edilen analiz sonuçlarının güvenilirliğini belirlemek için Marjinal güvenilirlik, klasik güvenilirlik ve H katsayısından yararlanılmıştır. Bir başka çalışmada farklı güvenilirlik katsayılarından da yararlanılabilir.

Bu çalışmada örneklem seçiminde her bir örneklem büyüklüğü yalnızca bir kez seçilmiştir. Başka bir çalışmada replikasyon yoluyla örneklem seçerek örneklem büyüklüğünün parametre değişmezliği üzerindeki etkisi araştırılabilir.

## Kaynakça

- Ackerman, T.A. (1989). *Unidimensional IRT Calibration of Compensatory and Non-Compensatory Multidimensional Items*. Applied Psychological Measurement, 13.
- Ackerman, T.A. (1992). *Assessing Construct Validty Using Multidimensional Item Response Theory*. Paper Presented at the Annual Meeting of American Educational Research Association. San Fransisco, CA, USA.
- Ackerman T.A. (1994). *Using Multidimensional Item Response Theory to Understand What Items and Tests are Measuring*. Applied Measurement In Education, 7 (4).
- Ackerman, T. A. (2005). Multidimensional item response theory modeling. A. Maydeu-Olivares, J.J. McArdle, (Ed.), *Comtemporary psychometrics* (3-27). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Andries van der Ark, L. (2007). *Mokken Scale Analysis in R*. Journal of Statistical Software. 20(11).
- Ansley, T.N. and Forsyth, R.A. (1985). *An Examination of the characteristics of Unidimensional IRT Estimates Derived From Two Dimensional Data*. Applied Psychological Measurement, 9(1), 37-48.
- Baykul, Y. ve Turut, F. (1992). Ölçekleme Teknikleri. ÖSYM Yayınları, Ankara.
- Bıkmaz-Bilgen, Ö. ve Doğan, N. (2017). *Çok Kategorili Parametrik ve Parametrik Olmayan Madde Tepki Kuramı Modellerinin Karşılaştırılması*. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 8(3);354-372.
- Bulut, O. (2013). *Between-person and Within-person Subscore Reliability: Comparison of Unidimensional and Multidimensional IRT Models*. Unpublished Doctoral Dissertation.The University of Minesota.

- Bulut, O. (2015). *Applying Item Response Theory Models to Entrance Examination for Graduate Studies: Practical Issues and Insights*. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2); 313-330.
- Cai L. (2017). *flexMIRT®: Flexible multilevel multidimensional item analysis and test scoring (version 3.51) [Computer software]*. Chapel Hill, NC: Vector Psychometric Group.
- Cai, L. (2008). *SEM Of Another Flavour: Two New Applications Of The Supplemented EM Algorithm*. *British Journal of Mathematical and Statistical Psychology*, 61, 309–329.
- Chang, Y.W. (1992). *A Comparison of Unidimensional and Multidimensional IRT Approaches to Test Information in A Test Battery*. Unpublished Doctoral Dissertation. University of Minnesota.
- Choi J, Peters M, Mueller R.O. (2010). *Correlation Analysis of Ordinal Data: From Pearson's R to Bayesian Polycoric Correlation*, *Asia Pacific Educ.* 11:459-466, S:460
- Coenders G., Saris W.e.(1995). *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments, (Chapter: 7)*. Publisher: Eötvös University Press, Budapest, Editors: Willem E. Saris, Akos Münnich, pp.125-144
- Crocker, L. ve Algina, J. (1986). *Intruduction to Classical and Modern Test Theory*. New York: CBS College Pub Co.
- Çakıcı Eser, D. (2015). *Çok Boyutlu Madde Tepki Kuramının Farklı Modellerinden Çeşitli Koşullar Altında Kestirilen Parametrelerin İncelenmesi*. Doktora Tezi. Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal Bilimler İçin Çok Değişkenli İstatistik Teknikleri*. Ankara: Pegem Akademi.
- Drasgow ve Parsons (1983). *Application Of Unidimensional Item Response Theory Models to Multidimensional Data*. *Applied Psychological Measurement*, 7, 189-199.

- Douglas, A. J. (2001). *Asymptotic Identifiability of Nonparametric Item Response Theory Models*. *Psychometrika*, Vol-66, No:4, 531-540.
- Embretson, S.E. ve Reise, S.P. (2000). *Item Response Theory For Psychologists*. Lawrence Erlbaum Associate, Inc.
- Gao, F. (1997). *DIMTEST Enhancements and Some Parametric IRT Asymptotics*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.
- Goldman, S.H., ve Raju, N. S. (1986). *Recovery of One- And Two-Parameter Logistic Item Parameters: An Empirical Study*. *Educational and psychological measurement*, 46(1),11-21.
- Gorsuch, R.L. (1974). *Factor Analysis (1. edition)*. Philadelphia, W.B. Saunders Company.
- Hambleton, R. K. ve Swaminathan, H. (1985). *Item response theory: principles and applications*. Boston: Academic Publishers Group.
- Hambleton, R. K. (1989). Principles and selected applications of item response theory. R.L. Linn(ed.), *Educational Measurement*, (s.147-200). Washington, DC: American Council of Education.
- Hambleton, R. K., Swaminathan, H. and Rogers, H. (1991). *Fundamentals of Item Response Theory*. Newbury Park CA: Sage.
- Harman, H. H. (1962). *Modern Factor Analysis (2. Baskı)*. University of Chicago Press, Chicago.
- Hulin, C. L., Lissak, R. I., ve Drasgow, F. (1982). *Recovery of two and three-parameter logistic item characteristic curves: A Monte Carlo study*. *Applied Psychological Measurement*, 6, 249-260.
- Kan, A. (2004). *Farklı Yetenek Ve Örneklem Gruplarından Kestirilen Madde Parametrelerinin Karşılaştırılması*. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi, Ankara.

- Kaptan, S. (1977). *Bilimsel araştırma Teknikleri*, Ankara: Tekişik Matbaası ve Rehber 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, Ankara.
- Köse, A. (2010). *Madde Tepki Kuramına Dayalı Tek Boyutlu ve Çok Boyutlu Modellerin Test Uzunluğu ve Örneklem Büyüklüğü Açısından Karşılaştırılması*. Yayınlanmamış Doktora Tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Linacre, J. M. (1994). *Sample Size and Item Calibrations Stability*. Rasch Measurement Transactions, 7, 4, 328.
- Lord, F. M. ve Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. New York: Addison-Wesley Pub. Co.
- Lord, F. M. (1968). *An analysis of the verbal scholastic aptitude test using Birnbaum's three-parameter logistic model*. Educational and Psychological Measurement, 28, 989-1020.
- McDonald, R.P. (1982). *Linear Versus Models in Item Response Theory*. Applied Psychological Measurement, 6, 379-396.
- Meara, K., Robin, F. and Sireci, S.G. (2000). *Using Multidimensional Scaling to Assess the Dimensionality of Dichotomous Item Data*. Multivariate Behavioral Research, 35 (2), 229–259.
- Meijer, R.R.; Sijtsma, K.; Smid, N.G. (1990). *Theoretical and Empirical Comparison of the Mokken and the Rasch approach to IRT*. Applied Psychological Measurement, 14(3), 283-298.
- Mokken, R. J. (1971). *A Theory and Procedure of Scale Analysis*. De Gruyter, Berlin, Germany.

- Molenaar, I. W. (2001). *Thirty years of nonparametric item response theory*. Applied Psychological Measurement. 25, 295–299.
- Molenaar, L.W. , Sijtsma, K. (1984). *Internal Consistency and Reliability in Mokken's Nonparametric Item Response Model*. Tijdschrift voor Onderwijsresearch 9(5). 257-268.
- Mor-Dirlik, E. (2017). *Parametrik ve parametrik olmayan madde tepki kuramı modellerinden çeşitli faktörlere göre elde edilen madde ve yetenek kestirimlerinin karşılaştırılması*. Yayınlanmamış Doktora Tezi. Eğitim Bilimleri Enstitüsü, Ankara Üniversitesi
- Murphy, R. K. ve Davidshofer, O. C. (2005). *Psychological Testing: Principles and Application* (6. bs.). New Jersey: Prentice-Hall.
- Özbek Baştuğ, Ö.Y. (2012). *Sosyal Bilgiler Testinde Çok Boyutluluk Analizi*. Kuram ve Uygulamada Eğitim Bilimleri- 12(1), 375-385
- Özer Özkan, Y. (2012). *Öğrenci Başarısının Belirlenmesi Sınavından (ÖBBS) Klasik Test Kuramı, Tek Boyutlu ve Çok Boyutlu Madde Tepki Kuramı İle Kestirilen Başarı Puanlarının Karşılaştırılması*. Yayınlanmamış Doktora Tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Öztuna D., Elhan A. H., Kurşun N. (2008). Sağlık Araştırmalarında Kullanılan İlişki Katsayıları, Türkiye Klinikleri Med Sci, 28:160-165
- Paas, L.J. , Sijtsma, K. (2007). *Nonparametric item response theory for investigating dimensionality of marketing scales: A SERVQUAL application*. Springer (19:157–170).
- Reeve B. (2006). *An introduction to modern measurement theory*. Rockville, MD: National Cancer Institute. <http://appliedresearch.cancer.gov/areas/cognitive/immt.pdf>
- Reckase, M.D.(2009). *Multidimensional Item Response Theory*. Springer Science&Business Media. USA.

- Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to nonparametric item response theory*. Thousand Oaks, CA: Sage.
- Sijtsma, K. (2005). *Nonparametric item response theory models*. In K. Kempf-Leonard (Ed.), *Encyclopedia of Social Measurement*. (pp. 875-882). New York: Elsevier.
- Sijtsma, K., & Meijer, R. R. (2007). *Nonparametric item response theory and related topics*. In C. R. Rao, & S. Sinharay (Eds.), *Handbook of statistics 26: Psychometrics*. (pp. 719-746). Amsterdam: Elsevier.
- Sinharay, S., Haberman, S. J., and Puhon, G. (2007). *Subscores Based on Classical Test Theory: to Report or Not to Report*. *Educational Measurement: Issues and Practice*, 26 (4), 21–28.
- Smith, J. (2009). *Some Issues in Item Response Theory: Dimensionality Assessment and Models for Guessing*. Unpublished Doctoral Dissertation. University of South California.
- Stochl, J. (2007). *Nonparametric Extension of Item Response Theory Models and Its Usefulness for Assessment of Dimensionality of Motor Yests*. *Acta Universitatis Carolinae*, 42(1), 75-94.
- Stochl, J., Jones, P. B., & Croudace, T. J. (2012). *Mokken Scale Analysis of Mental Health And Well-Being Questionnaire Item Responses: A Non-Parametric IRT Method In Empirical Research For Applied Health Researchers*. *BMC Med Res Methodol*, 12:74.
- Sünbül, Ö. (2011). *Çeşitli Boyutluluk Özelliklerine Sahip Yapılarda, Madde Parametrelerinin Değişmezliğinin Klasik Test Teorisi, Tek Boyutlu Madde Tepki Kuramı ve Çok Boyutlu Madde Tepki Kuramı Çerçevesinde İncelenmesi*. Yayınlanmamış doktora tezi, Mersin Üniversitesi Sosyal Bilimler Enstitüsü, Mersin.
- Şahin, A. (2012). *Madde Tepki Kuramı'nda Test Uzunluğu ve Örneklem Büyüklüğünün Model Veri Uyumu, Madde Parametreleri ve Standart Hata*

*Değerlerine Etkisinin İncelenmesi.* Eğitim Bilimleri Enstitüsü, Hacettepe Üniversitesi.

Tabachnick, G. B. ve Fidell, S. L. (2001). *Using Multivariate Statistics*, Fourth Edition, Allyn and Bacon, A Pearson Education Company.

Tekin, H. (1993). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Yargı Yayınları.

Thissen, D., ve Wainer, H. (1982). *Some Standart Errors in İtem Response Theory*. Psychometrika, 47, 397-412

Traub, R.E (1983). *A Priori Consideration In Choosing An Item Response Model*. In R.K. Hambleton (Ed.), *Applicatons of Item Response Theroy*. Vancouver BC. Educational Research Institute of British Columbia.

Tong, Y. ve Kolen, J.M. (2010). *IRT Proficiency Estimators and Their Impact*. Paper Presented at the Annual Conference of the National Council on Measurement in Education, Denver, CO.

Turgut, M. F. (1987). *Eğitimde Ölçme ve Değerlendirme Metodları*. Ankara: Saydam Matbaacılık.

Walker, C.M. ve Beretvas, S.N. (2003). *Comparing Multidimensional and Unidimensional Proficiency Classifications: Multidimensional IRT As a Diagnostic Aid*. Journal of Educational Measurement, 40 (3), 255-275.



## EK-A: Türkçe Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu

PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS  
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices  
Number of random correlation matrices: 500  
Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	47.4*	10.1	11.9
2	5.9	9.4	11.0
3	5.9	8.9	10.2
4	5.1	8.3	9.4
5	5.0	7.8	8.7
6	4.4	7.3	8.0
7	3.7	6.7	7.5
8	3.6	6.3	7.0
9	3.3	5.8	6.5
10	2.9	5.3	6.0
11	2.8	4.8	5.5
12	2.6	4.2	5.0
13	2.0	3.7	4.6
14	1.8	3.2	4.1
15	1.5	2.7	3.6
16	0.9	2.2	3.1
17	0.7	1.7	2.6
18	0.5	1.1	2.0
19	0.2	0.6	1.2
20	0.0	0.0	0.0

\* Advised number of dimensions: 1

```
> summary(check.monotonicity(tr))
  ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
v1 0.27 21 0 0 0 0 0 0 0 0
v2 0.19 21 0 0 0 0 0 0 0 0
v3 0.29 21 0 0 0 0 0 0 0 0
v4 0.19 21 0 0 0 0 0 0 0 0
v5 0.36 21 0 0 0 0 0 0 0 0
v6 0.30 21 0 0 0 0 0 0 0 0
v7 0.23 21 0 0 0 0 0 0 0 0
v8 0.31 21 0 0 0 0 0 0 0 0
v9 0.46 21 0 0 0 0 0 0 0 0
v10 0.27 21 0 0 0 0 0 0 0 0
v11 0.32 21 0 0 0 0 0 0 0 0
v12 0.35 21 0 0 0 0 0 0 0 0
v13 0.31 21 0 0 0 0 0 0 0 0
v14 0.25 21 0 0 0 0 0 0 0 0
v15 0.23 21 0 0 0 0 0 0 0 0
v16 0.27 21 0 0 0 0 0 0 0 0
v17 0.32 21 0 0 0 0 0 0 0 0
v18 0.29 21 0 0 0 0 0 0 0 0
v19 0.28 21 0 0 0 0 0 0 0 0
v20 0.31 21 0 0 0 0 0 0 0 0
```

## EK-B: Matematik Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi

### Sonucu

-----  
PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS  
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices  
Number of random correlation matrices: 500  
Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	46.9*	10.1	12.1
2	6.4	9.5	11.1
3	4.6	8.9	10.2
4	4.4	8.3	9.4
5	4.0	7.8	8.8
6	3.6	7.3	8.1
7	3.3	6.8	7.5
8	3.2	6.3	6.9
9	2.8	5.8	6.4
10	2.8	5.2	6.0
11	2.6	4.7	5.5
12	2.5	4.2	5.1
13	2.4	3.7	4.6
14	2.3	3.2	4.1
15	2.2	2.7	3.7
16	2.1	2.1	3.1
17	1.8	1.6	2.6
18	1.4	1.1	1.8
19	0.7	0.6	1.3
20	0.0	0.0	0.0

\* Advised number of dimensions: 1  
-----

```
> summary(check.monotonicity(mat))
  ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
v1  0.38  28  0      0      0  0      0      0      0      0
v2  0.36  21  0      0      0  0      0      0      0      0
v3  0.42  28  0      0      0  0      0      0      0      0
v4  0.37  21  0      0      0  0      0      0      0      0
v5  0.36  21  0      0      0  0      0      0      0      0
v6  0.23  28  0      0      0  0      0      0      0      0
v7  0.33  21  0      0      0  0      0      0      0      0
v8  0.33  21  0      0      0  0      0      0      0      0
v9  0.27  28  0      0      0  0      0      0      0      0
v10 0.23  28  0      0      0  0      0      0      0      0
v11 0.41  28  0      0      0  0      0      0      0      0
v12 0.33  28  0      0      0  0      0      0      0      0
v13 0.38  28  0      0      0  0      0      0      0      0
v14 0.44  28  0      0      0  0      0      0      0      0
v15 0.36  21  0      0      0  0      0      0      0      0
v16 0.28  28  0      0      0  0      0      0      0      0
v17 0.35  21  0      0      0  0      0      0      0      0
v18 0.18  28  0      0      0  0      0      0      0      0
v19 0.24  28  0      0      0  0      0      0      0      0
v20 0.31  21  0      0      0  0      0      0      0      0
```

## EK-C: T.C. İnkılap Tarihi ve Atatürkçülük Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu

-----  
 PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS  
 (Timmerman & Lorenzo-Seva, 2011)

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices  
 Number of random correlation matrices: 500  
 Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	47.8*	10.0	11.0
2	5.6	9.3	10.2
3	4.8	8.7	9.5
4	4.6	8.2	8.9
5	4.4	7.7	8.3
6	4.2	7.2	7.8
7	3.9	6.7	7.2
8	3.7	6.2	6.7
9	3.2	5.7	6.2
10	2.9	5.2	5.7
11	2.7	4.7	5.2
12	2.5	4.3	4.8
13	2.3	3.8	4.4
14	2.1	3.3	3.9
15	1.8	2.8	3.4
16	1.3	2.3	2.9
17	1.1	1.8	2.4
18	0.6	1.3	1.9
19	0.3	0.7	1.3
20	0.0	0.0	0.0

\* Advised number of dimensions: 1

-----

```
> summary(check.monotonicity(it))
  ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
v1  0.37 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v2  0.39 28  2  0.07  0.04 0.07  0.0026 6.38  2 66
v3  0.30 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v4  0.33 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v5  0.31 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v6  0.37 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v7  0.29 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v8  0.28 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v9  0.22 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v10 0.30 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v11 0.32 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v12 0.33 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v13 0.27 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v14 0.28 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v15 0.22 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v16 0.30 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v17 0.35 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v18 0.39 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v19 0.37 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
v20 0.35 28  0  0.00  0.00 0.00  0.0000 0.00  0  0
```

## EK-Ç: İngilizce Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu

PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS  
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices  
Number of random correlation matrices: 500  
Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	46.1*	9.9	11.0
2	7.4	9.3	10.2
3	5.4	8.7	9.5
4	4.6	8.2	8.9
5	4.3	7.7	8.3
6	3.8	7.2	7.8
7	3.3	6.7	7.3
8	3.3	6.2	6.7
9	3.1	5.8	6.2
10	3.0	5.2	5.7
11	2.5	4.7	5.3
12	2.5	4.3	4.8
13	2.2	3.8	4.3
14	2.1	3.3	3.9
15	2.0	2.8	3.4
16	1.7	2.3	2.8
17	1.5	1.8	2.4
18	0.7	1.3	1.9
19	0.5	0.7	1.2
20	0.0	0.0	0.0

\* Advised number of dimensions: 1

```
> summary(check.monotonicity(ing))
  ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
v1  0.38 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v2  0.36 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v3  0.35 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v4  0.35 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v5  0.32 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v6  0.34 28  1  0.04  0.03 0.03 0.0011 4.84  1  46
v7  0.33 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v8  0.09 28 13  0.46  0.08 0.74 0.0266 12.22 13 247
v9  0.24 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v10 0.27 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v11 0.37 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v12 0.42 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v13 0.22 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v14 0.29 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v15 0.41 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v16 0.37 21  0  0.00  0.00 0.00 0.0000 0.00  0  0
v17 0.24 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v18 0.38 28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v19 0.34 28  2  0.07  0.04 0.07 0.0026 5.46  2  64
v20 0.39 21  0  0.00  0.00 0.00 0.0000 0.00  0  0
```

## EK-D: Fen Bilgisi Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi

### Sonucu

PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS  
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices  
Number of random correlation matrices: 500  
Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	41.2*	10.0	12.4
2	7.1	9.0	10.8
3	6.1	8.6	10.0
4	4.9	8.1	9.3
5	4.2	7.6	8.6
6	4.1	7.1	8.0
7	4.0	6.7	7.4
8	3.7	6.2	6.9
9	3.3	5.7	6.3
10	3.2	5.2	6.0
11	3.2	4.8	5.5
12	2.8	4.3	5.0
13	2.8	3.8	4.6
14	2.6	3.3	4.2
15	2.2	2.9	3.9
16	2.1	2.4	3.5
17	1.6	1.9	2.9
18	0.9	1.5	2.5
19	0.2	0.9	2.1
20	0.0	0.0	0.0

\* Advised number of dimensions: 1

```
> summary(check.monotonicity(fb))
  ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
v1  0.33  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v2  0.29  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v3  0.30  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v4  0.30  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v5  0.33  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v6  0.28  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v7  0.34  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v8  0.35  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v9  0.22  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v10 0.28  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v11 0.34  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v12 0.34  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v13 0.34  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v14 0.30  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v15 0.31  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v16 0.32  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v17 0.21  21  7  0.33  0.1 0.46 0.0217 17.4  7 225
v18 0.29  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v19 0.16  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
v20 0.11  21  0  0.00  0.0 0.00 0.0000  0.0  0  0
```

## EK-E: Din Kültürü ve Ahlak Bilgisi Testi'ne Ait Faktör Analizi ile Gizil Monotonluk Testi Sonucu

PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS  
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices  
Number of random correlation matrices: 500  
Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	57.5*	10.0	11.2
2	5.3	9.2	10.1
3	4.0	8.6	9.3
4	3.6	8.1	8.8
5	3.5	7.6	8.2
6	3.1	7.1	7.6
7	3.0	6.7	7.1
8	2.9	6.2	6.6
9	2.5	5.7	6.1
10	2.4	5.3	5.7
11	2.3	4.8	5.3
12	2.1	4.3	4.8
13	1.8	3.8	4.3
14	1.7	3.4	3.9
15	1.4	2.9	3.4
16	1.1	2.4	3.0
17	0.9	1.9	2.5
18	0.6	1.3	2.0
19	0.2	0.8	1.3
20	0.0	0.0	0.0

\* Advised number of dimensions: 1

```
> summary(check.monotonicity(dk))
  ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
v1  0.38  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v2  0.36  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v3  0.35  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v4  0.35  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v5  0.32  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v6  0.34  28  1  0.04  0.03 0.03 0.0011 4.84  1  46
v7  0.33  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v8  0.09  28 13  0.46  0.08 0.74 0.0266 12.22 13 247
v9  0.24  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v10 0.27  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v11 0.37  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v12 0.42  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v13 0.22  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v14 0.29  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v15 0.41  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v16 0.37  21  0  0.00  0.00 0.00 0.0000 0.00  0  0
v17 0.24  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v18 0.38  28  0  0.00  0.00 0.00 0.0000 0.00  0  0
v19 0.34  28  2  0.07  0.04 0.07 0.0026 5.46  2  64
v20 0.39  21  0  0.00  0.00 0.00 0.0000 0.00  0  0
```

## EK-F: Birleşik Test'e Ait Faktör Analizi Sonucu

-----  
PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS  
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

Correlation matrices analyzed: Polychoric correlation matrices  
Number of random correlation matrices: 500  
Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu, 1992)

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	52.6*	10.1	11.8
2	10.7*	9.5	11.0
3	4.2	8.9	10.2
4	3.9	8.4	9.5
5	3.7	7.8	8.8
6	3.3	7.3	8.2
7	3.3	6.8	7.6
8	3.1	6.3	7.0
9	2.7	5.8	6.5
10	2.6	5.2	6.0
11	2.4	4.7	5.5
12	2.1	4.2	5.1
13	1.6	3.7	4.6
14	1.2	3.1	4.1
15	1.1	2.6	3.6
16	0.6	2.1	3.1
17	0.5	1.6	2.6
18	0.3	1.1	2.0
19	0.1	0.6	1.3
20	0.0	0.0	0.0

\* Advised number of dimensions when mean is considered: 2

-----



**EK-G: Birleşik Test'ten Elde Edilen %20'lik Dilime Ait Maddeler Arası**

**Korelasyon Matrisi**

	1	2	3	4	5	6	7	8	9	10
1	1	-0,004	-0,003	-0,003	-0,002	-0,001	-0,003	-0,006	-0,005	-0,003
2	-0,004	1	-0,008	-0,008	-0,006	-0,004	-0,008	-0,017	-0,014	-0,009
3	-0,003	-0,008	1	-0,005	-0,004	-0,003	-0,005	-0,010	-0,008	-0,005
4	-0,003	-0,008	-0,005	1	-0,004	-0,014	-0,005	-0,010	-0,008	-0,005
5	-0,002	-0,006	-0,004	-0,004	1	-0,005	-0,004	-0,008	-0,007	-0,004
6	-0,001	-0,004	-0,003	-0,014	-0,005	1	-0,004	-0,012	-0,01	-0,007
7	-0,003	-0,008	-0,005	-0,005	-0,004	-0,004	1	-0,010	-0,008	-0,005
8	-0,006	-0,017	-0,010	-0,010	-0,008	-0,012	-0,010	1	-0,018	-0,012
9	-0,005	-0,014	-0,008	-0,008	-0,007	-0,01	-0,008	-0,018	1	-0,010
10	-0,003	-0,009	-0,005	-0,005	-0,004	-0,007	-0,005	-0,012	-0,010	1
11	-0,014	0,008	-0,023	-0,023	-0,019	-0,014	-0,023	-0,051	-0,041	0,038
12	-0,008	-0,022	-0,014	-0,014	-0,011	-0,012	-0,014	-0,030	-0,024	-0,016
13	-0,023	-0,033	0,013	-0,040	-0,033	-0,01	-0,040	-0,038	-0,039	-0,046
14	-0,033	-0,093	-0,010	-0,057	-0,046	-0,021	0,037	-0,058	-0,045	0,015
15	-0,018	-0,012	-0,031	0,031	-0,025	-0,015	-0,031	-0,067	-0,053	-0,035
16	-0,011	-0,030	-0,018	-0,018	-0,015	-0,016	-0,018	-0,040	-0,032	-0,021
17	-0,006	-0,018	-0,011	-0,011	-0,009	-0,004	-0,011	-0,023	-0,019	-0,012
18	-0,012	-0,034	-0,021	-0,021	-0,017	-0,012	-0,021	-0,007	0,059	-0,024
19	-0,007	-0,020	-0,012	-0,012	-0,010	-0,017	-0,012	-0,027	-0,022	-0,014
20	0,023	-0,034	-0,067	-0,014	0,032	-0,028	-0,067	-0,064	-0,086	-0,093

	11	12	13	14	15	16	17	18	19	20
1	-0,014	-0,008	-0,023	-0,033	-0,018	-0,011	-0,006	-0,012	-0,007	0,023
2	0,008	-0,022	-0,033	-0,093	-0,012	-0,030	-0,018	-0,034	-0,020	-0,034
3	-0,023	-0,014	0,013	-0,010	-0,031	-0,018	-0,011	-0,021	-0,012	-0,067
4	-0,023	-0,014	-0,040	-0,057	0,031	-0,018	-0,011	-0,021	-0,012	-0,014
5	-0,019	-0,011	-0,033	-0,046	-0,025	-0,015	-0,009	-0,017	-0,010	0,032
6	-0,014	-0,012	-0,01	-0,021	-0,015	-0,016	-0,004	-0,012	-0,017	-0,028
7	-0,023	-0,014	-0,040	0,037	-0,031	-0,018	-0,011	-0,021	-0,012	-0,067
8	-0,051	-0,030	-0,038	-0,058	-0,067	-0,040	-0,023	-0,007	-0,027	-0,064
9	-0,041	-0,024	-0,039	-0,045	-0,053	-0,032	-0,019	0,059	-0,022	-0,086
10	0,038	-0,016	-0,046	0,015	-0,035	-0,021	-0,012	-0,024	-0,014	-0,093
11	1	-0,013	-0,08	-0,176	-0,11	-0,049	-0,053	-0,067	-0,061	-0,177
12	-0,013	1	-0,096	-0,08	-0,066	-0,019	-0,031	-0,060	-0,036	-0,079
13	-0,08	-0,096	1	-0,231	-0,171	-0,038	-0,067	-0,138	-0,084	-0,273
14	-0,176	-0,08	-0,231	1	-0,178	-0,181	-0,107	-0,206	-0,149	-0,191
15	-0,11	-0,066	-0,171	-0,178	1	-0,101	-0,041	-0,075	-0,056	-0,18
16	-0,049	-0,019	-0,038	-0,181	-0,101	1	-0,041	-0,013	-0,048	-0,113
17	-0,053	-0,031	-0,067	-0,107	-0,041	-0,041	1	-0,010	-0,028	-0,008
18	-0,067	-0,060	-0,138	-0,206	-0,075	-0,013	-0,010	1	-0,055	-0,089
19	-0,061	-0,036	-0,084	-0,149	-0,056	-0,048	-0,028	-0,055	1	0,061
20	-0,177	-0,079	-0,273	-0,191	-0,18	-0,113	-0,008	-0,089	0,061	1



**EK-H: Birleşik Test'ten Elde Edilen %80'lik Dilime Ait Maddeler Arası**

**Korelasyon Matrisi**

	1	2	3	4	5	6	7	8	9	10
1	1	-0,018	-0,039	0,081	-0,048	-0,013	-0,053	-0,068	-0,039	-0,008
2	-0,018	1	-0,015	-0,055	-0,016	0,008	0,003	-0,061	-0,088	0,101
3	-0,039	-0,015	1	-0,003	-0,103	0,025	-0,009	-0,072	0,019	0,076
4	0,081	-0,055	-0,003	1	0,000	-0,013	0,034	-0,061	-0,048	-0,034
5	-0,048	-0,016	-0,103	0,000	1	0,065	0,046	-0,058	-0,004	0,046
6	-0,013	0,008	0,025	-0,013	0,065	1	0,064	-0,016	0,043	0,014
7	-0,053	0,003	-0,009	0,034	0,046	0,064	1	-0,008	-0,016	0,052
8	-0,068	-0,061	-0,072	-0,061	-0,058	-0,016	-0,008	1	0,000	0,028
9	-0,039	-0,088	0,019	-0,048	-0,004	0,043	-0,016	0,000	1	0,020
10	-0,008	0,101	0,076	-0,034	0,046	0,014	0,052	0,028	0,020	1
11	-0,111	-0,124	-0,105	-0,145	-0,007	-0,149	-0,15	-0,094	-0,113	-0,143
12	-0,146	-0,093	-0,164	-0,014	-0,050	-0,129	-0,125	-0,052	-0,158	-0,125
13	-0,072	-0,097	-0,105	0,029	-0,063	-0,069	-0,093	-0,161	-0,164	-0,133
14	-0,024	-0,12	-0,117	-0,112	-0,085	-0,112	-0,1	-0,032	-0,103	-0,113
15	-0,084	-0,048	-0,067	-0,049	-0,145	-0,103	-0,077	-0,021	-0,063	-0,133
16	-0,012	-0,135	-0,060	-0,123	-0,066	-0,070	-0,109	-0,107	-0,082	-0,102
17	-0,033	-0,088	-0,096	-0,09	-0,117	-0,09	-0,027	-0,146	-0,059	-0,116
18	-0,026	-0,22	-0,113	-0,027	-0,048	-0,086	-0,172	-0,031	-0,148	-0,143
19	-0,086	-0,075	-0,149	-0,095	-0,08	-0,078	-0,136	-0,172	-0,089	-0,2
20	-0,163	-0,045	-0,049	-0,145	-0,101	-0,14	-0,167	-0,048	-0,092	-0,219

	11	12	13	14	15	16	17	18	19	20
1	-0,111	-0,146	-0,072	-0,024	-0,084	-0,012	-0,033	-0,026	-0,086	-0,163
2	-0,124	-0,093	-0,097	-0,12	-0,048	-0,135	-0,088	-0,22	-0,075	-0,045
3	-0,105	-0,164	-0,105	-0,117	-0,067	-0,060	-0,096	-0,113	-0,149	-0,049
4	-0,145	-0,014	0,029	-0,112	-0,049	-0,123	-0,09	-0,027	-0,095	-0,145
5	-0,007	-0,050	-0,063	-0,085	-0,145	-0,066	-0,117	-0,048	-0,08	-0,101
6	-0,149	-0,129	-0,069	-0,112	-0,103	-0,070	-0,09	-0,086	-0,078	-0,14
7	-0,15	-0,125	-0,093	-0,1	-0,077	-0,109	-0,027	-0,172	-0,136	-0,167
8	-0,094	-0,052	-0,161	-0,032	-0,021	-0,107	-0,146	-0,031	-0,172	-0,048
9	-0,113	-0,158	-0,164	-0,103	-0,063	-0,082	-0,059	-0,148	-0,089	-0,092
10	-0,143	-0,125	-0,133	-0,113	-0,133	-0,102	-0,116	-0,143	-0,2	-0,219
11	1	0,066	-0,052	-0,062	0,072	0,054	0,041	0,047	0,015	0,084
12	0,066	1	-0,024	-0,011	-0,055	-0,057	-0,050	0,021	0,032	0,043
13	-0,052	-0,024	1	0,046	0,009	-0,043	0,023	-0,035	0,051	0,050
14	-0,062	-0,011	0,046	1	-0,008	0,023	-0,042	0,030	0,009	0,055
15	0,072	-0,055	0,009	-0,008	1	-0,022	-0,004	0,030	-0,027	0,031
16	0,054	-0,057	-0,043	0,023	-0,022	1	-0,059	0,1	0,020	-0,036
17	0,041	-0,050	0,023	-0,042	-0,004	-0,059	1	-0,119	0,010	-0,012
18	0,047	0,021	-0,035	0,030	0,030	0,1	-0,119	1	0,038	-0,004
19	0,015	0,032	0,051	0,009	-0,027	0,020	0,010	0,038	1	0,017
20	0,084	0,043	0,050	0,055	0,031	-0,036	-0,012	-0,004	0,017	1

## EK-I: TBPMTK İçin Yazılan Kodlar

<Project>

Title = "2pl Calibration";

Description = "Cigdem";

<Options>

Mode = Calibration;

Etol = 0.005;

NQuadrature=41,5.5;

SavePRM=Yes;

SaveSCO=Yes;

Score=EAP;

NormalMetric3PL = Yes;

SlopeThreshold = Yes;

Quadrature = 49, 6.0;

SE = SEM;

MaxE = 1000;

MaxM = 250;

<Groups>

%Group%

File = "01.MAT50.DAT";

Varnames = V1-V20;

Select = V1-V20;

N = 50;

Ncats(V1-V20) = 2;

Model(V1-V20) = Graded(2);

<Constraints>

Prior (V1-V20), Slope : LogNormal(0.0,0.5);

## EK-İ: ÇBMTK İçin Yazılan Kodlar

<Project>

Title = "Multidimensional Calibration";

Description = "Cigdem";

<Options>

Mode = Calibration;

Quadrature = 21,5.0;

SE = SEM;

NumDec = 3;

SavePRM = Yes;

Score = EAP;

SaveSCO = YES;

Processor = 4;

FactorLoadings = Yes;

NewThreadModel = Yes;

MaxE = 500;

MaxM =100;

Etol = 1e-4;

Mtol = 1e-9;

SEMtol =1e-3;

SStol =1e-4;

<Groups>

%Group1%

File = "01.MAT50.DAT";

Varnames = v1-v20;

N = 50;

Dimensions = 2;

Ncats(v1-v20) = 2;

Model(v1-v20) = Graded(2);

<Constraints>

Fix (v1-v20), Slope;

Free (v1-v10), Slope(1);

Free (v11-v20), Slope(2);

Value Cov(2,1), 0.704;

Coeff Group1, (v1-v20), Slope, 1.702;

Coeff Group1, (v1-v20), Intercept, 1.702;

## EK-J: TBPOMTK İçin Yazılan Kodlar

```
setwd("C:/Users/ryhn/Desktop/Cigdem/Mokken_result")
```

```
library(mokken)
```

```
Veri50 <- read.table("MAT50.DAT", header=F, sep=",")
```

```
Veri100 <- read.table("MAT100.DAT", header=F, sep=",")
```

```
Veri200 <- read.table("MAT200.DAT", header=F, sep=",")
```

```
Veri500 <- read.table("MAT500.DAT", header=F, sep=",")
```

```
Veri1000 <- read.table("MAT1000.DAT", header=F, sep=",")
```

```
Veri5000 <- read.table("MAT5000.DAT", header=F, sep=",")
```

```
Evren <- read.table("EVREN.DAT", header=F, sep=",")
```

```
coefH(Veri50)
```

```
coefH(Veri100)
```

```
coefH(Veri200)
```

```
coefH(Veri500)
```

```
coefH(Veri1000)
```

```
coefH(Veri5000)
```

```
coefH(Evren)
```

```
mean(sapply(Veri50, mean, na.rm = TRUE))
```

```
mean(sapply(Veri100, mean, na.rm = TRUE))
```

```
mean(sapply(Veri200, mean, na.rm = TRUE))
```

```
mean(sapply(Veri500, mean, na.rm = TRUE))
```

```
mean(sapply(Veri1000, mean, na.rm = TRUE))
```

```
mean(sapply(Veri5000, mean, na.rm = TRUE))
```

```
mean(sapply(Evren, mean, na.rm = TRUE))
```

```
mean(sapply(Veri50, sd, na.rm = TRUE)/sqrt(length(Veri50)))
```

```
mean(sapply(Veri100, sd, na.rm = TRUE)/sqrt(length(Veri100)))
mean(sapply(Veri200, sd, na.rm = TRUE)/sqrt(length(Veri200)))
mean(sapply(Veri500, sd, na.rm = TRUE)/sqrt(length(Veri500)))
mean(sapply(Veri1000, sd, na.rm = TRUE)/sqrt(length(Veri1000)))
mean(sapply(Veri5000, sd, na.rm = TRUE)/sqrt(length(Veri5000)))
mean(sapply(Evren, sd, na.rm = TRUE)/sqrt(length(Evren)))
```

```
library(KernSmoothIRT)
```

```
keymat <- c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)
```

```
KernelVeri50 <- ksIRT(responses = Veri50, key = keymat, format = 1)
KernelVeri100 <- ksIRT(responses = Veri100, key = keymat, format = 1)
KernelVeri200 <- ksIRT(responses = Veri200, key = keymat, format = 1)
KernelVeri500 <- ksIRT(responses = Veri500, key = keymat, format = 1)
KernelVeri1000 <- ksIRT(responses = Veri1000, key = keymat, format = 1)
KernelVeri5000 <- ksIRT(responses = Veri5000, key = keymat, format = 1)
KernelEvren <- ksIRT(responses = Evren, key = keymat, format = 1)
```

```
subjthetaML(KernelVeri50)
subjthetaML(KernelVeri100)
subjthetaML(KernelVeri200)
subjthetaML(KernelVeri500)
subjthetaML(KernelVeri1000)
subjthetaML(KernelVeri5000)
subjthetaML(KernelEvren)
```

```
write.table(subjthetaML(KernelVeri50), "ThetaVeri50.csv", row.names = FALSE,
col.names=FALSE)
write.table(subjthetaML(KernelVeri100), "ThetaVeri100.csv", row.names = FALSE,
col.names=FALSE)
write.table(subjthetaML(KernelVeri200), "ThetaVeri200.csv", row.names = FALSE,
col.names=FALSE)
```

```
write.table(subjthetaML(KernelVeri500), "ThetaVeri500.csv", row.names = FALSE,  
col.names=FALSE)  
write.table(subjthetaML(KernelVeri1000), "ThetaVeri1000.csv", row.names =  
FALSE, col.names=FALSE)  
write.table(subjthetaML(KernelVeri5000), "ThetaVeri5000.csv", row.names =  
FALSE, col.names=FALSE)  
write.table(subjthetaML(KernelVEvren), "ThetaEvren.csv", row.names = FALSE,  
col.names=FALSE)
```

```
#Theta
```

```
setwd("C:/Users/ryhn/Desktop/Cigdem")
```

```
Theta50 <- read.table("Theta50.csv", header=F, sep="," , col.names=c("NonPar",  
"Par", "MIRT1", "MIRT2"))  
Theta100 <- read.table("Theta100.csv", header=F, sep="," ,  
col.names=c("NonPar", "Par", "MIRT1", "MIRT2"))  
Theta200 <- read.table("Theta200.csv", header=F, sep="," ,  
col.names=c("NonPar", "Par", "MIRT1", "MIRT2"))  
Theta500 <- read.table("Theta500.csv", header=F, sep="," ,  
col.names=c("NonPar", "Par", "MIRT1", "MIRT2"))  
Theta1000 <- read.table("Theta1000.csv", header=F, sep="," ,  
col.names=c("NonPar", "Par", "MIRT1", "MIRT2"))  
Theta5000 <- read.table("Theta5000.csv", header=F, sep="," ,  
col.names=c("NonPar", "Par", "MIRT1", "MIRT2"))  
ThetaEvren <- read.table("ThetaEvren.csv", header=F, sep="," ,  
col.names=c("NonPar", "Par", "MIRT1", "MIRT2"))
```

```
cor(Theta50)
cor(Theta50, method="spearman")
```

```
cor(Theta100)
cor(Theta100, method="spearman")
```

```
cor(Theta200)
cor(Theta200, method="spearman")
```

```
cor(Theta500)
cor(Theta500, method="spearman")
```

```
cor(Theta1000)
cor(Theta1000, method="spearman")
```

```
cor(Theta5000)
cor(Theta5000, method="spearman")
```

```
cor(ThetaEvren)
cor(ThetaEvren, method="spearman")
```



## EK-K: Etik Kurul İzin Muafiyeti Formu

Form 40

### Tez Çalışması Etik Kurul İzin Muafiyeti Formu

24 / 12 / 2015

Hacettepe Üniversitesi  
Eğitim Bilimleri Enstitüsü  
Eğitim Bilimleri Anabilim Dalı Başkanlığı'na

<b>Tez Başlığı / Konusu:</b>	<b>FARKLI BOYUTLULUK VE ÖRNEKLEM ÖZELLİKLERİNDE MADDE TEPKİ KURAMI UYGULAMALARININ PARAMETRE DEĞİŞMEZLİĞİNİN İNCELENMESİ</b> Bu çalışmada tek boyutlu parametrik ve parametrik olmayan madde tepki kuramı ile çok boyutlu madde tepki kuramı kapsamında gerçek verilerde parametre değişmezliğinin incelenmesi amaçlanmıştır. Bu amaçla TEOG sınavının farklı boyutlardaki alt testlerinden elde edilen başarı puanlarının, tek boyutlu parametrik ve parametrik olmayan MTK ile çok boyutlu MTK kapsamında farklılaşıp farklılaşmadığı incelenecektir. Araştırmada karşılaştırmanın yapılması için kullanılacak olan indeksler cevap aranan alt problemler doğrultusunda belirlenecektir. Bireylere ait yetenek kestirimleri için her bir kuram kapsamında elde edilecek standart hata ortalamaları, bireylere ait yetenek düzeyleri arasındaki ilişki büyüklüğü için araştırılan üç kuram kapsamında kesitirilen yetenek parametreleri ile toplam test puanı arasındaki ilişkiyi ortaya koyacak olan pearson momentler çarpımı korelasyon katsayısı ile spearman-brown sıra farkları korelasyon katsayısı, maddelere ait parametrelerin karşılaştırılması için üç kuram kapsamında kesitirilecek olan standart hata ortalamaları ve her bir kuramın ne derecede güvenilir sonuçlar verdiğini belirlemek amacıyla kesitirilen güvenilirlik katsayıları incelenecektir.
------------------------------	--

Yukarıda başlığı/konusu gösterilen tez çalışmam:

1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır.
2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmasını gerektirmemektedir.
3. Beden bütünlüğüne müdahale içermemektedir.
4. Gözlemsel ve betimsel araştırma (anket, ölçek/skala çalışmaları, dosya taramaları, veri kaynakları taraması, sistem-model geliştirme çalışmaları) niteliğinde değildir.

Hacettepe Üniversitesi Etik Kurullar ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

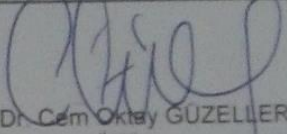
Gereğini saygılarımla arz ederim.

Çiğdem REYHANLIOĞLU KEÇEOĞLU  
(Öğrencinin Adı Soyadı, İmzası)

#### Öğrenci Bilgileri

<b>Adı Soyadı</b>	Çiğdem REYHANLIOĞLU KEÇEOĞLU		
<b>Öğrenci No</b>	N11244388		
<b>Anabilim Dalı</b>	Eğitim Bilimleri		
<b>Programı</b>	Eğitimde Ölçme ve Değerlendirme		
<b>Statüsü</b>	<input type="checkbox"/> Yüksek Lisans	<input checked="" type="checkbox"/> Doktora	<input type="checkbox"/> Bütünleşik Dr.

Danışman Görüşü ve Onayı

  
Doç. Dr. Cem Oktay GÜZELLER  
(İmza)  
(Danışmanın unvanı, Adı ve Soyadı)

#### EK-L: Etik Beyanı

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

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

beyan ederim.

08/06/2018



(İmza)

Çiğdem REYHANLIOĞLU KEÇEOĞLU



## EK-M:Doktora Tez Çalışması Orijinallik Raporu

08/06/2018

HACETTEPE ÜNİVERSİTESİ

Eğitim Bilimleri Enstitüsü

Eğitim Bilimleri Ana Bilim Dalı Başkanlığına,

Tez Başlığı : PARAMETRİK VE PARAMETRİK OLMAYAN MADDE TEPKİ KURAMINDA FARKLI ÖRNEKLEM BÜYÜKLÜKLERİNE VE BOYUTLULUKLARINA GÖRE PARAMETRE DEĞİŞMEZLİĞİNİN İNCELENMESİ

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

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
04/06 /2018	82	151,726	27/04 /2018	%9	972158331

Uygulanan filtreler:

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

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

Ad Soyadı: Çiğdem REYHANLIOĞLU KEÇEOĞLU

Öğrenci No.: N12244388

Ana Bilim Dalı: Eğitim Bilimleri

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü:  Y.Lisans  Doktora  Bütünleşik Dr.

  
İmza

DANIŞMAN ONAYI

UYGUNDUR.  
Prof. Dr. Nuri DOĞAN



## EK-N: Dissertation Originality Report

08/06/2018

HACETTEPE UNIVERSITY

Graduate School Of Educational Sciences

To The Department Of Educational Sciences

Thesis Title : EXAMINING OF THE PARAMETER INVARIANCE ACCORDING TO DIFFERENT SAMPLE SIZES AND DIMENSIONALITIES IN PARAMETRIC AND NONPARAMETRIC ITEM RESPONSE THEORY

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

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
04/06 /2018	82	151,726	27/04 /2018	%9	972158331

Filtering options applied:

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

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

I respectfully submit this for approval.

Name Lastname: Çiğdem REYHANLIOĞLU KEÇEOĞLU

Student No.: N12244388

Department: Educational Sciences

Program: Measurement and Evaluation In Education

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

  
Signature

### ADVISOR APPROVAL

APPROVED

Prof. Dr. Nuri DOĞAN





### **EK-O: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı**

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

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

- Tezimin/Raporumun tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.**

(Bu seçenekle teziniz arama motorlarında indekslenebilecek, daha sonra tezinizin erişim statüsünün değiştirilmesini talep etmeniz ve kütüphane bu talebinizi yerine getirirse bile, teziniz arama motorlarının ön belleklerinde kalmaya devam edebilecektir)

- Tezimin/Raporumun 19.06.2021 tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.**

(Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisi alınabilir).

- Tezimin/Raporumun ..... tarihine kadar erişime açılmasını istemiyorum ancak kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisinin alınmasını onaylıyorum.**

- Serbest Seçenek/Yazarın Seçimi: .....**

08/06/2018

  
(İmza)  
Çiğdem REYHANLIOĞLU KEÇEOĞLU

