



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

Eğitim Bilimleri Ana Bilim Dalı

Eğitimde Ölçme ve Değerlendirme Programı

ÇOK BOYUTLU TESTLERİN TEK BOYUTLU VE ÇOK BOYUTLU YÖNTEMLERE GÖRE  
EŞİTLENMESİ

Yaşar Mehmet ZOR

Doktora Tezi

Ankara, 2023

Liderlik, arařtırma, inovasyon, kaliteli eđitim ve deđiřim ile

*Daha ileriye... En iyiye...*



Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

ÇOK BOYUTLU TESTLERİN TEK BOYUTLU VE ÇOK BOYUTLU YÖNTEMLERE GÖRE  
EŞİTLENMESİ

EQUATING MULTIDIMENSIONAL TESTS ACCORDING TO UNIDIMENSIONAL AND  
MULTIDIMENSIONAL METHODS

Yaşar Mehmet ZOR

Doktora Tezi

Ankara, 2023

## Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

Yařar Mehmet ZOR'un hazırladıđı "Çok Boyutlu Testlerin Tek Boyutlu ve Çok Boyutlu Yöntemlere Göre Eřitlenmesi" 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

J¼ri Üyesi (Danıřman) Prof. Dr. H¼lyya KELECİOđLU

J¼ri Üyesi Prof. Dr. Nuri DOđAN

J¼ri Üyesi Prof. Dr. Dilara BAKAN  
KALAYCIOđLU

J¼ri Üyesi Dr. Öğr. Üyesi Ömer KUTLU

Bu tez Hacettepe Üniversitesi Lisans¼st¼ Eđitim, Öğretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 13 / 07 / 2023 tarihinde uygun gör¼lm¼ř ve Enstit¼ Yönetim Kurulunca ..... / ..... / ..... tarihi itibarıyla kabul edilmiřtir.

Prof. Dr. İsmail Hakkı MİRİCİ  
Eđitim Bilimleri Enstitüsü M¼d¼r¼

## Öz

Bu çalışmada çok boyutlu test formlarına çeşitli koşullar altında çok boyutlu ve tek boyutlu ölçek dönüştürme işlemleri yapıp, madde ve yetenek parametrelerinden elde edilen eşitleme hatalarının karşılaştırılması amaçlanmıştır. Değerlendirme ölçütü olarak eşitleme hatası (RMSE) kullanılmıştır. R yazılımında iki boyutlu ve basit yapıli test formları üretilmiştir. Madde ve yetenek parametrelerinin kestiriminde IRTPRO, çok boyutlu ölçek dönüştürmede Linkmirt, tek boyutlu ölçek dönüştürmede IRTEQ yazılımı kullanılmıştır. Çalışma koşulları olarak örneklem büyüklüğü (1000 ve 2000), boyutlar arası korelasyon (0.1; 0.5 ve 0.9), ortak madde oranı (%20 ve %40), gruplar arası yetenek dağılımı farkı (0.05 ve 0.5) ve parametre kestirim modeli (2PLM ve 3PLM) alınmıştır. Eşdeğer olmayan gruplar ortak test deseninde çok boyutlu ve tek boyutlu ölçek dönüştürme işlemleri yapılmış ve elde edilen RMSE değerleri karşılaştırılmıştır. Çok yönlü varyans analizi ve t testi ile ele alınan koşullara ve bunların etkileşimlerine göre RMSE değerleri arasındaki farkın istatistiksel olarak anlamlılığı incelenmiştir. Örneklem büyüklüğü 2000 ve ortak madde oranı %40 olduğunda daha düşük RMSE değerleri elde edilmiştir. Çok boyutlu ve tek boyutlu ölçek dönüştürmede boyutlar arası korelasyonun sırasıyla 0.1 ve 0.9 olduğu koşullarda daha düşük RMSE değerleri kestirilmiştir. Gruplar arası yetenek dağılımı farkı düşük olduğu koşullarda elde edilen RMSE değerleri daha düşüktür. Çok boyutlu ölçek dönüştürmede ortalama-ortalama yöntemi, tek boyutlu ölçek dönüştürmede Stocking-Lord yöntemi kullanıldığında daha düşük RMSE değerleri elde edilmiştir.

**Anahtar sözcükler:** madde tepki kuramı, çok boyutlu madde tepki kuramı, test eşitleme, ölçek dönüştürme, çok boyutlu ölçek dönüştürme

## Abstract

In this study, multidimensional and unidimensional scale transformation procedures were performed on multidimensional test forms under various conditions, and it was aimed to compare the equating errors obtained from item and ability parameters. Equating error (RMSE) was used as an evaluation criterion. Two-dimensional and simple structured test forms were produced in R software. IRTPRO software was used for estimation of item and ability parameters, Linkmirt software was used for multidimensional scale transformation, and IRTEQ software was used for unidimensional scale transformation. Sample size (1000 and 2000), correlation between dimensions (0.1; 0.5 and 0.9), common item ratio (20% and 40%), difference in ability distribution between groups (0.05 and 0.5) and parameter estimation model (2PLM and 3PLM) were taken as study conditions. Multidimensional and unidimensional scale transformation procedures were performed in the non-equivalent groups anchor test design and the RMSE values obtained were compared. The statistical significance of the difference between the RMSE values according to the conditions and their interactions was examined by means of multi-way analysis of variance and t-test. Lower RMSE values were obtained when the sample size was 2000 and the common item ratio was 40%. In multidimensional and unidimensional scale transformation, lower RMSE values were estimated when the correlation between dimensions was 0.1 and 0.9, respectively. RMSE values are lower when the difference in ability distribution between groups is low. Lower RMSE values were obtained when the mean-mean method was used in multidimensional scale transformation and Stocking-Lord method was used in unidimensional scale transformation.

**Keywords:** item response theory, multidimensional item response theory, test equating, scale transform, multidimensional scale transform

## Teşekkür

Doktora eğitimimde ve tez çalışmamda hoşgörüsü ve yardımlarıyla bana rehberlik eden, eğitimime ve gelişimime büyük katkısı olan değerli danışman hocam Prof. Dr. Hülya KELECİOĞLU'na,

Tezimin gelişimine getirdikleri yapıcı önerilerle destek veren, aklıma takılan sorularıma anlayışlı ve güler yüzlü tavırları ile yardımcı olan değerli tez izleme komitesi üyelerim Prof. Dr. Selahattin GELBAL'a, Dr. Öğr. Üyesi Ömer KUTLU'ya,

Doktora ders ve tez süreçlerimde destek ve katkılarıyla yanımda olan Prof. Dr. Nuri DOĞAN'a, Doç. Dr. Meltem YURTÇU'ya, Doç. Dr. İlhan KOYUNCU'ya, Dr. Öğr. Üyesi Gökhan KUMLU'ya, Dr. Öğr. Üyesi Ömer Faruk ŞEN'e, Dr. Öğr. Üyesi Başak ERDEM KARA'ya, Arş. Gör. Ertunç UKŞUL'a, Süleyman ÖNAL'a,

Manevi desteklerini her zaman arkamda hissettiğim ve bugünlere gelmem için sağladıkları koşulsuz desteklerinden ötürü kıymetli aileme ve dostlarıma teşekkür ederim.

**İçindekiler**

Kabul ve Onay.....	ii
Öz.....	iii
Abstract.....	iv
Teşekkür.....	v
Tablolar Dizini.....	viii
Şekiller Dizini.....	x
Simgeler ve Kısaltmalar Dizini.....	xi
Bölüm 1 Giriş.....	1
Problem Durumu.....	2
Araştırmanın Amacı ve Önemi.....	3
Araştırma Problemi.....	3
Sayıtlılar.....	4
Sınırlılıklar.....	4
Tanımlar.....	5
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	6
Madde Tepki Kuramı.....	6
Tek Boyutlu Madde Tepki Kuramı.....	9
Çok Boyutlu Madde Tepki Kuramı.....	11
Test Eşitleme.....	16
Test Eşitleme Koşulları.....	17
Test Eşitleme Desenleri.....	18
Kalibrasyon, Ölçek Dönüştürme ve Eşitleme.....	19
İlgili Araştırmalar.....	24
Tek Boyutlu Eşitleme Yapılan Çalışmalar.....	24
Çok Boyutlu Eşitleme Yapılan Çalışmalar.....	26
Bölüm 3 Yöntem.....	30



Araştırmanın Türü .....	30
Simülasyon Koşulları.....	30
Verilerin Üretilmesi .....	33
Ölçek Dönüştürme Süreci ve Verilerin Analizi .....	33
Değerlendirme Kriteri .....	34
Çalışmada Ele Alınan Koşullara Göre Elde Edilen RMSE Değerlerinin Karşılaştırılması .....	35
Bölüm 4 Bulgular, Yorumlar ve Tartışma.....	36
Alt Problem 1'e İlişkin Bulgular ve Yorumlar .....	36
Alt Problem 2'ye İlişkin Bulgular ve Yorumlar.....	49
Alt Problem 3'e İlişkin Bulgular ve Yorumlar .....	62
Bölüm 5 Sonuç ve Öneriler.....	81
Sonuçlar .....	81
Öneriler .....	86
Kaynaklar .....	89
EK-A: Araştırmada Yer Alan Koşullara İlişkin Elde Edilen RMSE Değerleri .....	xcv
EK-B: Araştırma Etik Komisyon İzin Muafiyeti Formu/ Araştırma Etik Komisyonu Onay Bildirimi .....	xcvii
EK-C: Etik Beyanı.....	xcviii
EK-Ç: Doktora Tez Çalışması Orijinallik Raporu .....	xcix
EK-D: Dissertation Originality Report .....	c
EK-E: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı.....	ci

## Tablolar Dizini

<b>Tablo 1</b> Çalışma Kapsamında Yer Alan Değişkenler ve Koşullar.....	31
<b>Tablo 2</b> Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri.....	37
<b>Tablo 3</b> Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri.....	40
<b>Tablo 4</b> Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri.....	43
<b>Tablo 5</b> Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri.....	46
<b>Tablo 6</b> Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	50
<b>Tablo 7</b> Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	53
<b>Tablo 8</b> Gruplar Arası Yetenek Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	56
<b>Tablo 9</b> Gruplar Arası Yetenek Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	59
<b>Tablo 10</b> Çok Boyutlu ve Tek Boyutlu Ölçek Dönüştürme Yöntemlerine Göre Kestirilen Ortalama RMSE Değerlerinin Karşılaştırılması .....	63
<b>Tablo 11</b> Çok boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre a parametresinden kestirilen ortalama RMSE değerleri .....	65
<b>Tablo 12</b> Çok boyutlu ölçek dönüştürmede a parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması .....	66
<b>Tablo 13</b> Çok boyutlu ölçek dönüştürmede a parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları .....	67
<b>Tablo 14</b> Çok boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre b parametresinden kestirilen ortalama RMSE değerleri .....	68
<b>Tablo 15</b> Çok boyutlu ölçek dönüştürmede b parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması .....	69
<b>Tablo 16</b> Çok boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre yetenek parametresinden kestirilen ortalama RMSE değerleri .....	70
<b>Tablo 17</b> Çok boyutlu ölçek dönüştürmede yetenek parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması .....	71
<b>Tablo 18</b> Tek boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre a parametresinden kestirilen ortalama RMSE değerleri .....	72

<b>Tablo 19</b> <i>Tek boyutlu ölçek dönüştürmede a parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması</i> .....	73
<b>Tablo 20</b> <i>Tek boyutlu ölçek dönüştürmede a parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları</i> .....	74
<b>Tablo 21</b> <i>Tek boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre b parametresinden kestirilen ortalama RMSE değerleri</i> .....	75
<b>Tablo 22</b> <i>Tek boyutlu ölçek dönüştürmede b parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması</i> .....	76
<b>Tablo 23</b> <i>Tek boyutlu ölçek dönüştürmede b parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları</i> .....	77
<b>Tablo 24</b> <i>Tek boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre yetenek parametresinden kestirilen ortalama RMSE değerleri</i> .....	78
<b>Tablo 25</b> <i>Tek boyutlu ölçek dönüştürmede yetenek parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması</i> .....	79
<b>Tablo 26</b> <i>Tek boyutlu ölçek dönüştürmede yetenek parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları</i> .....	80

## Şekiller Dizini

<b>Şekil 1</b> MTK'da Madde Karakteristik Eğrisi .....	7
<b>Şekil 2</b> ÇBMTK'da Madde Karakteristik Yüzeyi .....	12
<b>Şekil 3</b> İki yetenekten oluşan çok boyutlu basit yapılı madde tepki kuramı modeli .....	16
<b>Şekil 4</b> Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	38
<b>Şekil 5</b> Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	41
<b>Şekil 6</b> Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	45
<b>Şekil 7</b> Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	48
<b>Şekil 8</b> Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	51
<b>Şekil 9</b> Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	54
<b>Şekil 10</b> Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	58
<b>Şekil 11</b> Gruplar Arası Yetenek Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri .....	61

## Simgeler ve Kısaltmalar Dizini

**MTK:** Madde Tepki Kuramı

**ÇBMTK:** Çok Boyutlu Madde Tepki Kuramı

**1PLM:** 1 Parametrelili Lojistik Model

**2PLM:** 2 Parametrelili Lojistik Model

**3PLM:** 3 Parametrelili Lojistik Model

**OO:** Ortalama-Ortalama

**OS:** Ortalama-Sigma

**SL:** Stocking-Lord

**RMSE:** Hata Kareleri Ortalamasının Karekökü

## Bölüm 1

### Giriş

Eğitim ve psikoloji alanındaki testler, öğrenme eksiklerini belirleme, bir kuruma birey seçme ve yerleştirme ya da bireyleri işe alma gibi çeşitli amaçlarla kullanılmaktadır. Bireyler hakkında doğru ve adil kararların verilebilmesi için uygulanan test sonuçlarının geçerli ve güvenilir olması beklenir. Ulusal ve uluslararası düzeyde uygulanan pek çok geniş ölçekli test uygulamasında sınav katılımcılarına yıl içerisinde birden fazla test uygulaması yapılmaktadır. Bu uygulamalarda test güvenliğini sağlamak ve birden fazla kere maddeye maruz kalma durumlarını kontrol etmek amacıyla farklı test formları kullanılmaktadır (Kim, Lee ve Kolen, 2019). Bu uygulamalarda test maddeleri bloklara bölünmektedir ve bu bloklarda yer alan maddelerden, maddenin ait olduğu içerik ve madde formatı birbirine eşdeğer olacak şekilde, eşit sayıda maddeden oluşan farklı test formları elde edilmektedir (Xu, 2009). Farklı eşdeğer test formlarının kullanılmasının temel amacı, puanların karşılaştırılabilir olmasını ve birbirinin yerine kullanılabilmesini sağlamaktır. Test formları aynı özelliği ölçmek amacıyla benzer kapsama ve psikometrik özelliklere sahip olacak şekilde hazırlansa da, test formlarından elde edilen puanlara sayısal dönüştürme yapılarak puanların eşitlenmesi gerekir (Braun ve Holland, 1982). Yapılan istatistiksel ayarlama sonucunda test formlarından elde edilen puanlar aynı ölçeğe yerleştirilmiş olur ve böylelikle test puanları birbirinin yerine kullanılabilir (Kolen ve Brennan, 2014).

Teste ait farklı formlardan elde edilen puanlar genellikle aynı ölçek düzeyinde bulunmadığından, bu puanların karşılaştırılması hatalı sonuçlar elde edilmesine neden olacaktır. Test eşitleme ile farklı test formlarından elde edilen puanlar arasında bağlantı kurulur ve bu şekilde puanlar aynı ölçek üzerinde yer alır. İstatistiksel bir teknik olan test eşitleme, test formlarından elde edilen puanlar arasındaki ilişkiyi ortaya koymaktadır (Chu ve Kamata, 2003). Test formların birim sistemlerini birbirine eşitlemek olarak da tanımlanabilen test eşitleme sonucu formlardan elde edilen eşitlenmiş puanlar birbirlerinin

yerine kullanılabilir (Angoff, 1984; Hambleton ve Swaminathan, 1985; Kolen ve Brennan, 2014).

Yaygın olarak kullanılan eşitleme yöntemleri, daha çok tek boyutlu madde tepki kuramına dayalı olan yöntemlerdir. Tek boyutlu madde tepki kuramı, bir testte yer alan maddelerin yalnızca tek bir yeteneği ölçmesini gerektiren tek boyutluluk varsayımına dayanır. Ancak bu durum, birkaç içerik alanından oluşan testlerde ya da karma formatlı testler gibi birden fazla madde formatının kullanıldığı testlerde tek boyutluluk varsayımının karşılanmasını imkansız hale getirmektedir (Kim, Lee ve Kolen, 2020). Çok boyutlu madde tepki kuramı psikolojik ve eğitsel süreçlerin karmaşıklığı göz önüne alınarak geliştirilmiş ve çok boyutlu madde tepki kuramı modelleri çoklu yetenekler tarafından belirlenen sınav katılımcılarının konumunu bir maddeden belirli bir puan alma olasılığı ile ilişkilendiren matematiksel bir ifade sağlar (Reckase, 2009). Bu bağlamda, tek boyutluluk varsayımının karşılanmadığı durumlarda çok boyutlu madde tepki kuramına dayalı modellerin kullanılması gerekir. Test planını ve veri yapısını aslına uygun şekilde yansıtan modellerin gerçekliğin daha iyi bir resmini sunduğuna şüphe yoktur (Kim ve Lee, 2023).

### **Problem Durumu**

Eğitimde ve psikolojide kullanılan testlerin büyük çoğunluğu tek boyuttan oluşmamakta, maddeler kimi zaman birden çok boyutla ilişkili olmakta, baskın bir boyutun dışında test maddeleri ile ölçülen başka boyutlar da bulunmaktadır. Çok boyutlu veri yapılarına tek boyutlu eşitleme yöntemleri uygulandığında, ortaya çıkan eşitleme ilişkilerinin tek boyutluluk varsayımının ihlali nedeniyle büyük miktarda hata içereceği yorumu yapılabilir (Brossman, 2010). Tek bir özelliği ölçen bir test geliştirmenin pratikte olanaksız olması, bunun yanında testin ölçtüğü özelliklere ilişkin olabildiğince az bilgi kaybı yaşanması için çok boyutlu yöntemler geliştirilmiştir (Reckase, 2009). Bu bağlamda çok boyutlu veri yapıları için çok boyutlu eşitleme yöntemlerinin kullanılması gerekmektedir. Bu çalışmada çeşitli koşullar altında çok boyutlu veri yapılarına çok boyutlu ve tek boyutlu ölçek dönüştürme

yöntemleri uygulandığında elde edilecek eşitleme hatalarının değişiminin nasıl olduğu sorusuna yanıt aranmıştır.

### **Araştırmanın Amacı ve Önemi**

Eğitimde ve psikolojide kullanılan testlerin, genellikle birden fazla örtük özelliği ölçtüğü söylenebilir. Yapılan ölçek dönüştürme çalışmalarında genellikle verilerin tek boyutlu olduğu kabul edilerek tek boyutlu analiz süreçleri yürütülmektedir. Fakat testlere yapılan boyutluluk analizlerinde bazı test yapılarının tek boyutluluk varsayımını sağlamadığı görülmektedir. Çok boyutlu maddelerden oluşan bir test formunun verilerine tek boyutlu analiz süreçleri yürütülmesi bilgi kaybına neden olabilir ve bu durum testin sahip olduğu model veri uyumunu düşürücü yönde etki eder. Bununla birlikte çok boyutlu veri yapıları üzerinden yapılan çok boyutlu ölçek dönüştürme çalışmaları oldukça azdır. Bu bağlamda çok boyutlu veriler üzerinden çok boyutlu ve tek boyutlu ölçek dönüştürme yapılan bu çalışmanın sonuçlarının alan yazına katkı sağlayacağı düşünülmektedir.

### **Araştırma Problemi**

Çok boyutlu test formlarına eşdeğer olmayan gruplar ortak test deseninde çok boyutlu ölçek dönüştürme yöntemleri ve tek boyutlu ölçek dönüştürme yöntemleri kullanılarak eşitleme yapıldığında elde edilen eşitleme hataları nasıldır?

### **Alt Problemler**

1. Çok boyutlu test verileri eşdeğer olmayan gruplar ortak test deseninde çok boyutlu ölçek dönüştürme yöntemleriyle eşitlendiğinde madde ve yetenek parametrelerinden kestirilen eşitleme hatası (RMSE) değerleri;
  - a. Örneklem büyüklüğüne
  - b. Boyutlar arası korelasyona
  - c. Ortak madde oranına



- d. Gruplar arası yetenek dağılımı farkına
  - e. Parametre kestirim modeline göre nasıl değişmektedir?
2. Çok boyutlu test verileri eşdeğer olmayan gruplar ortak test deseninde tek boyutlu ölçek dönüştürme yöntemleriyle eşitlendiğinde madde ve yetenek parametrelerinden kestirilen eşitleme hatası (RMSE) değerleri;
    - a. Örneklem büyüklüğüne
    - b. Boyutlar arası korelasyona
    - c. Ortak madde oranına
    - d. Gruplar arası yetenek dağılımı farkına
    - e. Parametre kestirim modeline göre nasıl değişmektedir?
  3. Ele alınan koşullara göre yapılan çok boyutlu ve tek boyutlu ölçek dönüştürme işlemleri sonucu elde edilen RMSE değerlerinin koşullar arası etkileşimi ve anlamlılık düzeyi nasıldır?

### **Sayıtlılar**

Araştırma kapsamında üretilen veri yapılarının gerçek durumları yansıttığı varsayılmaktadır.

### **Sınırlılıklar**

1. Araştırma kapsamında kullanılan çok boyutlu veri setinde boyut sayısı 2 ile sınırlıdır.
2. Bu araştırma tek ve çok boyutlu ölçek dönüştürmeye ilişkin hataların incelendiği çeşitli koşullar (ölçek dönüştürme yöntemleri, örneklem büyüklüğü, boyutlar arası korelasyon, ortak madde oranı, gruplar arası yetenek dağılımı farkı ve parametre kestirim modeli) ile sınırlıdır.

**Tanımlar**

Basit Yapılı Test: Her bir maddenin kendi yetenek düzeyi ile ilişkili olduğu ve yalnızca o özelliği ölçtüğü çok boyutlu testlerdir.

## Bölüm 2

### Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

#### Madde Tepki Kuramı

Bir testi yanıtlayan bireyin sahip olduğu özelliği ortaya koymak için geliştirilen ilk kuram Klasik Test Kuramı (KTK)'dir. Bu kurama göre, testten alınan puan, gerçek puan ve hata puanının toplamına eşittir. Test ve madde istatistiklerinin testin uygulandığı gruba bağlı olması, bir test maddesiyle karşılaşan sınav katılımcısının nasıl performans göstereceğini belirlemek için hiç bir temel sağlamaması, bir dizi test maddesiyle ölçülen bir yetenek üzerinde sınav katılımcılarının karşılaştırmalarının, sınav katılımcılarına aynı test maddelerinin uygulandığı sınırlı durumlarda yapılması KTK'nın sınırlılıklarını oluşturur (Hambleton ve Swaminathan, 1985). KTK'nın sahip olduğu bu sınırlılıkları aşabilmek için Madde Tepki Kuramı (MTK) geliştirilmiştir. MTK, bir sınav katılımcısının  $\theta$  yetenek düzeyini bir maddeden belirli bir puan alma olasılığı ile ilişkilendiren bir istatistiksel model ailesidir (Lord, 1980). MTK, eğitimde ve psikolojide çeşitli amaçlarla kullanılan ve testler için avantajlı uygulamalara (ölçekleme, test eşitleme, madde yanlılığını belirleme, bilgisayar ortamındaki bireyselleştirilmiş test uygulamaları) sahip olan güçlü bir yöntemdir. MTK, bireyin yeteneği ile bir maddeyi doğru cevaplama olasılığı arasında bulunan ilişkiyi matematiksel olarak modellemeyi amaçlar (Cook ve Eignor, 1991). Bu kuramda madde güçlüğü ve yetenek aynı ölçekte yer alır ve bu durum madde ve yetenek parametrelerinin değişmezliğini sağlar. MTK'da madde parametreleri testin uygulandığı gruptan bağımsız olarak kestirilirken, yetenek parametreleri de madde parametrelerinden bağımsız olarak kestirilir (Crocker ve Algina, 1986; Lord, 1980).

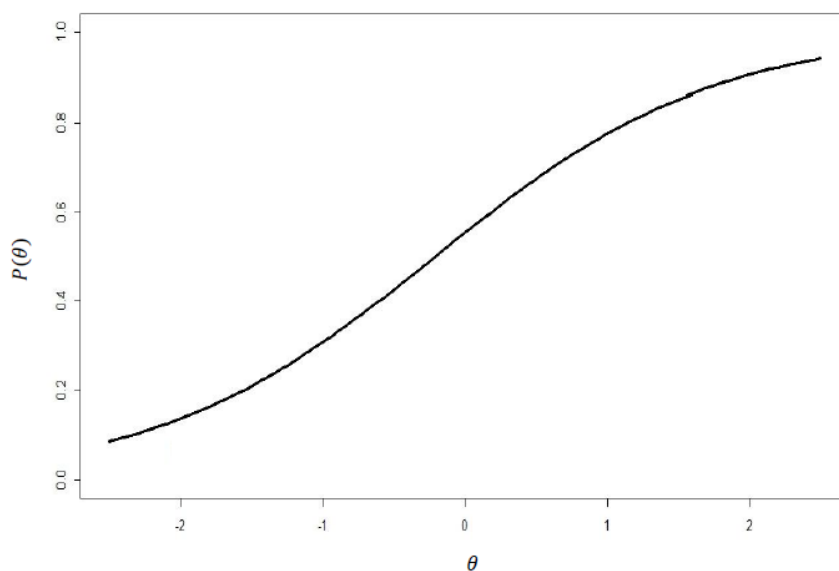
Test eşitleme ve ölçekleme için alanyazında yaygın olarak MTK'ya dayalı yöntemler kullanılmaktadır. MTK'ya dayalı test eşitleme yöntemlerinde, katılımcıların yetenekleri ile maddeyi doğru yanıtlama ihtimalleri arasındaki bağıntının matematiksel bir fonksiyon ile tanımlanması amaçlanmaktadır (Hambleton ve Swaminathan, 1985).

MTK'nın sahip olduđu varsayımlardan birisi tek boyutluluktur. Tek boyutluluk, bir testte yer alan maddelerin aynı örtük özelliđi ölçmeyi amaçlaması ve bireyin testten kestirilen yeteneđinin sadece o özelliđe ait olmasıdır. Bir diđer varsayım olan yerel bağımsızlık ise, sınava giren kişinin yetenek seviyesinin performansı ile ilgili tüm bilgileri içerdini ve dolayısıyla bir maddeyi doğru yanıt olasığının, yetenek düzeyi açıklandıktan sonra diđer maddeleri doğru yanıt olasığını etkilemediğini ifade eder (Hambleton, Swaminathan ve Rogers, 1991; Kim, 2018).

MTK'da, KTK'dan farklı olarak madde ve yetenek parametrelerinin deđişmezliđi mevcuttur. Bireye ait yetenek parametresi ( $\theta$ ) ile bireyin maddeyi doğru yanıt olasığını arasındaki ilişkiyi karakterize eden fonksiyon Madde Tepki Fonksiyonu (MTF) olarak isimlendirilir (Lord, 1980). MTF'nin görsel temsili olan ve bireyin maddeyi yanıt olasığını yetenek ( $\theta$ ) ölçüğü üzerinde gösteren eğriye ise Madde Karakteristik Eğrisi (MKE) adı verilir (Embretson ve Reise, 2000). Şekil 1'de MKE gösterilmiştir.

## Şekil 1

*MTK'da Madde Karakteristik Eğrisi*



Eđitim ve psikoloji alanındaki testler, puanlamaya karışan çok sayıda faktörden dolayı doğası geređi bir dereceye kadar çok boyutludur (Ackerman, Gierl & Walker, 2003).

Bu yüzden eğitimde ve psikolojide kullanılan testlerdeki maddeler arasındaki ilişkinin tek boyutlu modellerde açıklandığı gibi basit olmadığı açıktır. Örneğin bir matematik problemini ele aldığımızda, bir matematik problemini alan yanıtlayıcıdan önce okuduğunu anlama becerilerini kullanması, sonrasında da matematiksel akıl yürütme ve işlem becerilerini kullanarak problemi çözmesi beklenir. Bu durumda, soruya doğru yanıt verebilmek için hem okuduğunu anlama becerisi hem de matematiksel akıl yürütme ve işlem becerisi gerekmektedir. Bu gibi durumlarda tek boyutluluk varsayımı ihlal edilmektedir. Tek boyutlu modellerin sahip olduğu bu sınırlılığın önüne geçebilmek için çok boyutlu modeller geliştirilmiştir (Smith, 2009).

Testte yer alan maddelerin toplandığı baskın bir boyut varsa ve testin sahip olduğu boyutluluk kaynağı büyük bir çoğunlukla bu boyut tarafından açıklanıyorsa, tek boyutlu modellerin parametre kestirim doğruluğu artmaktadır. Bunun yanında veri yapısı birden fazla örtük özelliği ölçüyorsa, tek boyutlu analiz yöntemleri veri yapısındaki başat boyut ekseninde kestirim yapmakta (Gibbons, Immekus ve Bock, 2007), diğer boyutlar hakkındaki bilgi kaybolmakta ve elde edilen yetenek ve madde parametre kestirimleri yüksek oranda yanlış olabilmektedir (Kreiter, 1993; Reckase, 1985). Birden fazla bölümden oluşan testlerin çok boyutlu olma ihtimali yüksektir. Testteki maddelere verilen yanıtları birden fazla psikolojik özellik etkiliyorsa, tek boyutluluk varsayımının ihlal edildiği yorumu yapılabilir ve bu durumda tek boyutlu modellerin kullanılmaması gerekir (Kim ve Lee, 2023; Zhang, 2009). Bunun yerine çok boyutlu madde tepki kuramı (ÇBMTK) altında yer alan modeller kullanılabilir. Bu modellere çalışmanın ilerleyen kısımlarında değinilecektir.

İki kategorili (1-0) puanlanan veriler için MTK kapsamında sırasıyla bir parametrelili lojistik model (1PLM), iki parametrelili lojistik model (2PLM) ve üç parametrelili lojistik model (3PLM) kullanılmaktadır. Bu çalışma kapsamında iki kategorili puanlanan veriler kullanıldığından, tek boyutlu MTK ve çok boyutlu MTK modelleri ayrı başlıklar halinde açıklanmıştır.

## Tek Boyutlu Madde Tepki Kuramı

İki kategorili puanlanan bir testte yer alan maddelerin madde parametre kestirimi 1PLM, 2PLM ya da 3PLM'den birisi seçilerek gerçekleştirilir. Tek boyutlu MTK'da yer alan bu modellerde sadece tek bir özelliğin bireylerin yanıtlarını etkilediği kabul edilir.

**Bir Parametrelili Lojistik Model (1PLM).** Bu modelde maddenin doğru yanıtlanma olasılığı, bireyin yeteneği ile madde güçlüğü'nün bir fonksiyonudur. Madde güçlük indeksi ( $b$ ), testi alan bireylerin yetenek düzeyini açıklayan tek madde parametresidir. Madde güçlük indeksi, bireyin maddeyi doğru yanıtlanma olasılığı 0,5 iken yetenek ölçeğindeki değere karşılık gelir.  $b$  parametresi -3 ile +3 arasında değerler alırken, ayırt edicilik parametresi her madde için aynıdır ve şans parametresinin değeri 0'dır (DeMars, 2016). Rasch modeli ise 1PLM'nin özel bir hali olup, madde ayırt edicilik parametresi tüm maddeler için 1 olarak alınır (De Ayala, 2009). 1PLM'ye ilişkin matematiksel ifade eşitlik 1'de verilmiştir.

$$P_i(\theta) = \frac{e^{D_i(\theta-b_i)}}{1 + e^{D_i(\theta-b_i)}} \quad (1)$$

$P_i(\theta)$ :  $\theta$  yeteneğinde  $i$  maddesinin doğru cevaplanma olasılığı

$b_i$ :  $i$  maddesinin güçlük indeksi

$D$ : ölçek sabiti (1,7)

**İki Parametrelili Lojistik Model (2PLM).** Bu modelde 1PLM'den farklı olarak güçlük parametresine ek olarak ayırt edicilik parametresi ( $a$ ) vardır. Madde karakteristik eğrisinin eğimine karşılık gelen  $a$  parametresi  $-\infty$  ile  $+\infty$  aralığında değerler alsa da, negatif değerler yanıtlayıcılar için negatif ayırt ediciliğe karşılık geleceğinden, bu maddelerin testten çıkarılması gerekir.  $a$  parametresinin genellikle 0 ile 2 arasında değerler alabileceği kabul edilirken, şans parametresinin değeri ise 0'dır (De Ayala, 2009). Modelde  $a$  parametresinin

değerinin artması MKE'yi dikleştirir, bu durum maddeden elde edilen bilginin fazlaşmasını sağlar. 2PLM'ye ilişkin eşitlik aşağıda verilmiştir.

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad (2)$$

$P_i(\theta)$  :  $\theta$  yeteneğinde i maddesinin doğru yanıtlanma ihtimali

$b_i$  : i maddesine ait güçlük indeksi

$a_i$  : i maddesine ait ayırt edicilik indeksi

$D$  : ölçek sabiti (1,7)

**Üç Parametrelili Lojistik Model (3PLM).** Bu modelde güçlük ( $b$ ) ve ayırt edicilik ( $a$ ) parametrelerine ek olarak şans başarısı parametresi ( $c$ ) bulunmaktadır. 3PLM'ye ilişkin matematiksel ifade aşağıdaki şekildedir:

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{D_i(\theta-b_i)}}{1 + e^{D_i(\theta-b_i)}} \quad (3)$$

$P_i(\theta)$  :  $\theta$  yeteneğinde i maddesinin doğru yanıtlanma ihtimali

$b_i$  : i maddesine ait güçlük indeksi

$a_i$  : i maddesine ait ayırt edicilik indeksi

$c_i$  : i maddesine ait şans parametresi

$D$  : ölçek sabiti (1,7)

Bu modelde yer alan şans parametresi, yetenek düzeyinin en düşük değeri için maddenin doğru yanıtlanma ihtimalidir ve 0 ile 1 arasında değerler alır. Özellikle çoktan seçmeli maddelerden oluşan testlerde doğru yanıtın şansa bulunma olasılığı

bulduğundan, çoktan seçmeli testlerde bu model kullanılır (Hambleton ve Swaminathan, 1985). İki kategorili maddeler için geniş ölçekli sınavlarda şans başarısının da dikkate alınmasından ötürü daha çok 3PLM tercih edilmektedir (Huggins, 2012).

### **Çok Boyutlu Madde Tepki Kuramı**

Eğitimde ve psikolojide kullanılan testlerin çoğu zaman tek boyutlu bir yapıya sahip olduğu söylenemez. Verilerin çok boyutlu yapısını dikkate almamak bilgi kaybına yol açar, bu da testin yapı geçerliğini düşürücü yönde etki eder. Testlerin sahip olduğu boyutluluk yapısı incelenip buna göre uygun olan MTK modeli seçilmelidir. Tek boyutlu modellerde, veri yapısının kullanılan MTK modeli tarafından yeterince temsil edilebilmesi gerekir (Kim, 2018). Bunun sağlanamadığı durumlar için, tek boyutlu MTK modellerinin uzantısı olarak birden fazla boyutu ve yetenek düzeyini ölçen ÇBMTK modelleri geliştirilmiştir (Reckase, 2009). ÇBMTK modelleri telafisel (compensatory) model ve telafisel olmayan (non-compensatory) model olmak üzere ikiye ayrılır (Ackerman, 1989). Telafisel modeller bir boyutta bulunan üstün yeteneğin, diğer boyut ya da boyutlarda mevcut olan eksik yeteneği telafi ettiği durumlarda kullanılırlar. Telafisel olmayan modeller ise bir boyuta ilişkin yetenek düzeyindeki eksikliğin diğer boyut ya da boyutlarda bulunan yüksek yetenekle telafi edilemediği durumlarda kullanılır (Reckase, 2009).

ÇBMTK altında yer alan modeller, iki veya daha fazla boyut ile yanıtlayıcının belirli bir test maddesini doğru yanıtlama ihtimali arasındaki ilişkiyi matematiksel olarak açıklayan modellerdir (Reckase, 2009). Diğer bir ifade ile tek boyutlu MTK modellerinin çok boyutluluğa uyarlanmasıdır (Ackerman, Gierl ve Walker, 2003). Tek boyutlu MTK'da kullanılan madde karakteristik eğrisinin yerine ÇBMTK'da madde karakteristik yüzeyi (MKY) kullanılır. Şekil 2'de iki boyutlu bir maddeye ait MKY örneği verilmiştir (Reckase, 2009).

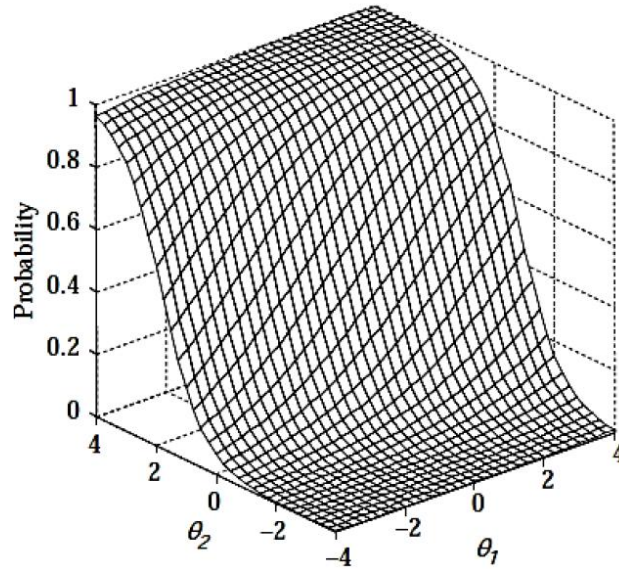
ÇBMTK'da olasılık fonksiyonu tanımlanırken kullanılan parametrelerin sayısına göre çeşitli modeller vardır. Bu çalışmada telafisel olan çok boyutlu iki parametrelilik lojistik model



ve üç parametrelı lojistik model kullanılarak analizler yürütülmüştür. Aşağıda bu modellere ilişkin ayrı başlıklar halinde bilgiler verilmiştir.

## Şekil 2

### ÇBMTK'da Madde Karakteristik Yüzeyi



**Çok Boyutlu İki Parametrelı Lojistik Model.** Bu model tek boyutlu 2PLM'nin çok boyutluluğa uyarlanmış halidir. İki kategorili verilerde tek boyutlu modellere benzer şekilde, çok boyutlu modellerde de maddeye verilen doğru yanıt 1 ve yanlış yanıt 0 olacak şekilde kodlanır. Eşitlik 4'te çok boyutlu 2PLM'ye ilişkin eşitlik verilmiştir (Reckase, 2009).

$$P(X_{ij} = 1 | \theta_j, a_i, d_i) = \frac{e^{a_i \theta_j + d_i}}{1 + e^{a_i \theta_j + d_i}} \quad (4)$$

$X_{ij}$  : i maddesine j bireyinin verdiği yanıt

$\theta_j$  : j bireyinin yetenek düzeyi

$a_i$  : i maddesine ait ayrıcalık

$d_i$  : i maddesinin kesim noktası

Eşitlik 4'te yer alan  $\theta_j$ , çok boyutlu testte bulunan yeteneklerin vektörü,  $a_i$  madde eğimlerinin vektörü,  $d_i$  ise madde güçlüğü ile ilişkili olup a ve b parametrelerine bağlı olarak hesaplanan kesme (intercept) parametresidir (Reckase, 2009).

**Çok Boyutlu Üç Parametrelili Lojistik Model.** Bu model tek boyutlu 3PLM'nin genelleştirilmiş halidir. Çok boyutlu 3PLM'ye ilişkin eşitlik aşağıda verilmiştir (Reckase, 2009).

$$P(X_{ij} = 1 | \theta_j, a_i, c_i, d_i) = c_i + (1 - c_i) \frac{e^{a_i \theta_j + d_i}}{1 + e^{a_i \theta_j + d_i}} \quad (5)$$

$X_{ij}$  : i maddesine j bireyinin verdiği yanıt (1-0)

$\theta_j$  : j bireyinin yeteneği

$a_i$  : i maddesinin ayırt ediciliği

$c_i$  : i maddesinin şans parametresi

$d_i$  : i maddesinin kesim noktası

ÇBMTK altında yer alan modellerde, her boyut için ayrı olacak şekilde ayırıcılık (a) parametresi varken, testin bütününe ilişkin tek bir tane güçlük (b) parametresi bulunur. Bu durum, telafisel modellerde bulunan bir özelliğe ilişkin bireyde bulunan yüksek yetenek düzeyinin başka bir özelliğe ilişkin düşük yeteneği tamamlaması özelliğinden dolayı maddenin yüksek yetenek düzeyi kullanılarak yanıtlanmasından kaynaklanmaktadır (Reckase, 2009).

Tek boyutlu MTK'da madde ayırt ediciliği MKE'nin en dik noktasında bulunan eğime karşılık gelir ve maddeye doğru yanıt verme olasılığının ne kadar dikleştiğini açıklar, madde güçlüğü ise MKE'nin en dik olduğu noktadaki yetenek değerine eşittir (DeMars, 2016). Testi alan  $\theta=b$  noktasındaki bireylerin yaklaşık %50'sinin maddeye doğru yanıt verdiği yorumu yapılabilir. ÇBMTK'da ise maddenin doğru yanıtlanma olasılığı birden fazla yeteneğe bağlıdır ve bu ilişki Madde Karakteristik Yüzeyi (MKY) ile gösterilir. ÇBMTK'daki madde

güçlüğü,  $\theta$  ölçeğinin başlangıç noktasının MKY'nin en yüksek noktasına olan uzaklığıdır. Bu uzaklık  $d$  parametresi ile gösterilir. MKE'nin tanımlanmasında sadece  $b$  parametresi yeterliyken, MKY'nin tanımlanmasında birden fazla parametre değeri gerekir (Reckase, 2009).

ÇBMTK'da birden fazla boyut bulunduğundan, her boyuta ait bir ayırt edicilik parametresi bulunduğu açıklanmıştır. Bu yüzden ÇBMTK'da bulunan madde ayırt ediciliği tek boyutlu MTK'dan farklıdır. Madde ayırt ediciliği, çok boyutlu ayırt edicilik (MDISC) ile temsil edilmektedir. Her bir boyuta ilişkin ayırt edicilik değeri  $a_{ik}$  olmak üzere, MDISC indeksine ilişkin vektörün uzunluğu eşitlik 6'da verildiği şekilde hesaplanmaktadır (Reckase, 2009).

$$MDISC = \sqrt{\sum_{n=1}^k a_{ik}^2} \quad (6)$$

ÇBMTK'da vektörün doğrultusu ve eksenlerle yaptığı açıya göre vektörün konumu belirlenebilir. Testte yer alan  $i$  maddesine ilişkin ayırt edicilik vektörünün yönü eşitlik 7'de verildiği gibi hesaplanır (Reckase, 2009).

$$a_i = \arccos\left(\frac{a_{i1}}{MDISC}\right) \quad (7)$$

Eğer maddeye ait ayırt edicilik vektörü  $\theta_1$  ve  $\theta_2$  eksenleri arasında 45 derecelik bir açıya sahipse, maddenin  $\theta_1$  ve  $\theta_2$  yeteneklerinin her ikisini de eşit olarak ölçtüğü yorumu yapılabilir. Ayırt edicilik vektörünün açısı  $\theta_1$  eksenine yakınsa, maddelerin ağırlıklı olarak  $\theta_1$  yeteneğini;  $\theta_2$  eksenine yakınsa ağırlıklı olarak  $\theta_2$  yeteneğini ölçtüğü yorumu yapılabilir (Zhang, 2009).

MTK'da yer alan  $b$  parametresi, ÇBMTK'da çok boyutlu güçlük indeksi (MDIFF) olarak ifade edilir. MDIFF, eşitlik 8'de verildiği şekilde hesaplanır. MDIFF indeksinin yüksek

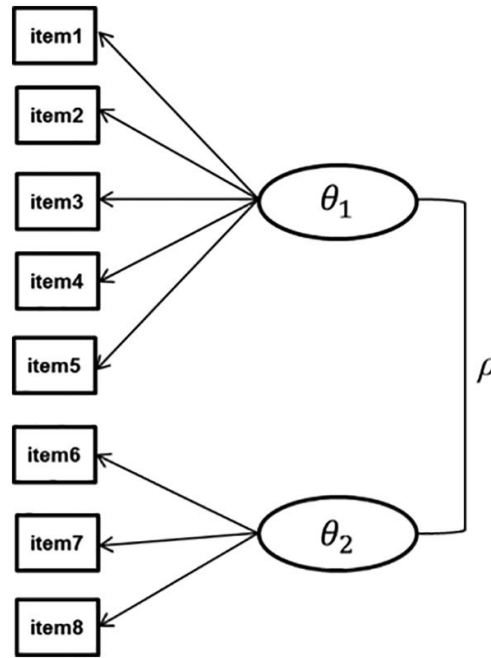
pozitif deęerleri zor maddelere, düşük deęerleri ise genellikle gözlenen yetenek seviyeleri için doęru yanıt olasılıęı yüksek olan maddeleri gösterir (Reckase, 2009).

$$MDIFF = \frac{-d_i}{MDISC} \quad (8)$$

ÇBMTK'da testin ölçtüęü örtük özellik sayısı arttıkça, testin birden fazla yeteneęi ölçtüęü yorumu yapılabilir. Bazı testlerde yer alan maddeler birden fazla yeteneęi ölçerken, bazı testlerdeki maddelerin her biri ayrı bir yetenek altında birleşebilir. Birden fazla yeteneęi ölçen ve birden çok yetenek ile ilişkili olan maddelerin olduęu testlere karmaşık yapılı, her bir maddenin farklı bir yetenek düzeyini belirttięi ve kendi yetenek düzeyi ile ilişkili olduęu testlere ise basit yapılı test denir (Zhang, 2012). Basit yapılı testte yer alan maddeler sadece bir boyutta yük verirken, karmaşık yapılı bir testte maddeler birden fazla boyutta çapraz olarak yer alır (Ackerman, Gierl ve Walker, 2003). Basit yapı gerçek veri yapılarında hiçbir zaman ortaya çıkmayabilir, bununla birlikte basit yapı, veriler için maddelerin ikincil yüklerinin sıfır etrafında dalgalandıęını varsayar (Kim ve Lee, 2023). Bu çalışmada iki boyutlu basit yapılı veri setleri üzerinden analizler gerçekleştirilmiştir. Şekil 3'te iki yetenekten oluşan basit yapılı çok boyutlu madde tepki kuramı modeli örneęi verilmiştir (Kim, Lee ve Kolen, 2020).

### Şekil 3

*İki yetenekten oluşan çok boyutlu basit yapılı madde tepki kuramı modeli*



### Test Eşitleme

Ulusal ve uluslararası düzeyde uygulanan pek çok geniş ölçekli test uygulamasında benzer kapsam ve güçlük düzeyinde olan farklı test formları kullanılır. Teste katılan bireyler arasında doğru bir değerlendirmenin yapılabilmesi için, testin farklı formlarından elde edilen puanların karşılaştırılabilir olması gerekir (Tian, 2011). Uygulanan testin farklı formlarından elde edilen puanların birbiri yerine kullanılabilmesi ancak ortak bir ölçeğe yerleştirilmeleri, yani eşitlenmeleri ile mümkündür. Test eşitleme, benzer özelliklere sahip farklı test formlarından elde edilen puanların ve farklılıkların düzenlenerek bu formlardan alınan puanların birbirinin yerine kullanılabilmesini sağlayan bir istatistiksel süreç (Kolen ve Brennan, 2014) ya da uygulanan test formlarının birim sistemlerinin birbirine dönüştürülmesi (Angoff, 1984) olarak tanımlanabilir. Eşitleme yapılacak olan testlerin sağlaması gereken koşullar aşağıda açıklanmıştır.

## Test Eşitleme Koşulları

Test formlarına test eşitleme yapılabilmesi için testlerin sağlaması gereken bazı koşullar vardır. Bu koşullar; eşitlik, gruplar arası değişmezlik, simetriklik, aynı yapıyı ölçme ve eşit güvenilirliğe sahip olma şeklinde sıralanabilir (Dorans ve Holland, 2000; Lord, 1980).

**Eşitlik.** İki test formu eşitlendiği durumda bireylere hangi test formunun (X-Y) uygulandığı bir fark yaratmayacaktır (Kolen ve Brennan, 2014). Eşitlik koşulunun sahip olduğu birtakım özellikler bulunmaktadır. Eğer testler farklı yetenek ya da becerileri ölçüyorsa, bunun yanında iki test formundan birisine ait puanların yetenek düzeyindeki koşullu frekans dağılımı diğer formun dönüştürülmüş koşullu frekans dağılımı ile aynı değilse testler eşitlenemez. Yine testler arasında bulunan güçlük farkı çok fazlaysa ve testler arasında paralellik bulunmuyorsa test eşitleme yapılamayacaktır (Lord, 1980).

**Gruplar Arası Değişmezlik.** Eşitleme ilişkisinin test formlarının uygulandığı gruptan bağımsız olması, bir başka deyişle eşitleme ilişkisinin tüm gruplar için aynı olmasıdır (Hambleton, Swaminathan ve Rogers, 1991; Lord, 1980).

**Simetriklik.** Test eşitlemede, test formlarından hangisinin referans ve hangisinin temel form olduğunun fark etmemesi durumudur (Hambleton, Swaminathan ve Rogers, 1991). Örneğin, bu özelliğe göre X Formundaki 26 ham puan, Y Formundaki 27 ham puana karşılık geliyorsa, Y Formundaki 27 ham puan da X formundaki 26 ham puana karşılık gelmelidir (Livingston, 2004; Lord, 1980).

**Aynı Yapıyı Ölçme.** Eşitlenecek iki test formunun aynı içerik ve istatistiksel özelliklere göre oluşturulmuş olması gerekir. Aksi takdirde, kullanılan istatistiksel prosedürler ne olursa olsun, ölçülen özellik ve içerik açısından benzer olmayan testlere eşitleme yapılamaz ve elde edilen puanlar birbirinin yerine kullanılamaz (Kolen ve Brennan, 2014; Zhu, 1998).

## Test Eşitleme Desenleri

Test eşitleme sürecinde öncelikle eşitleme deseninin belirlenmesi gerekir. Eşitleme deseninin seçiminde uygulanabilirlik ve istatistiksel özellikler dikkate alınmalıdır (Kolen ve Brennan, 2014). Eşitleme desenleri, tek grup deseni (single group design), eşdeğer gruplar deseni (random groups design) ve eşdeğer olmayan gruplar ortak test deseni deseni (non-equivalent groups anchor test (NEAT) design) olarak sınıflandırılabilir (Crocker ve Algina, 1986; Kolen ve Brennan, 2014).

**Tek Grup Deseni.** Bu desende aynı bireyler iki ya da daha fazla test formunu alır. Bireyler her iki test formunu da aldığından, yeteneklerindeki değişim kontrol edilebilmektedir. Her iki test de aynı gruba uygulandığı için test alan bireylerin yetenek düzeyi değişimi kontrol altındadır. Tek grup desenin dezavantajı, formların uygulanma sırasının eşitlemeye olan etkisidir. İkinci formun uygulanmasında yorgunluk ve sorulara aşına olma durumuna “sıra etkisi” denilmektedir. Sıra etkisini kontrol altına almak için “dengelenmiş tek grup deseni” önerilmiştir (Kolen ve Brennan, 2014). Buna göre test uygulanacak grup ikiye ayrılır ve birinci grup sırasıyla X ve Y, ikinci grup sırasıyla Y ve X formlarını alır ve bu şekilde sıra etkisi kontrol altına alınabilir.

**Eşdeğer Gruplar Deseni.** Bu desende test formlarını alacak olan grup rastgele olacak şekilde ikiye bölünür. Kitapçıklar sırasıyla X formu ve Y formu olacak şekilde hazırlanır ve birinci kişi X formunu, ikinci kişi Y formunu, üçüncü kişi ise tekrar X formunu alır ve bu sarmal süreç sayesinde X ve Y formunu alan iki eşdeğer grup ortaya çıkar. Testin iki formuna ilişkin grupların performansları arasındaki fark, formların güçlük düzeyleri arasındaki farkı ortaya koyma konusunda bilgi verir (Kolen ve Brennan, 2014). Bu desende bireyler sadece bir form alırlar, tek grup desenindeki gibi aynı grubun iki formu alması gerekmez ve böylelikle sıra etkisi ortadan kalkar. Ancak bu desende tüm test formları aynı oturumda uygulanmalıdır. Bu desende her bir formu katılımcıların belirli bir kısmı aldığı için büyük örneklemelere ihtiyaç duyulur (Kolen ve Brennan, 2014).

**Eşdeğer Olmayan Gruplar Ortak Test Deseni.** Bu desen genellikle ulusal ve uluslararası düzeyde uygulanan geniş ölçekli sınavlarda kullanılmaktadır. Bu desende test formlarını alan gruplar eşdeğer yetenek dağılımına sahip değildir. Eşdeğer olmayan gruplar, içinde ortak madde setini içeren test formlarından birini alırlar. Test formlarında yer alan ortak maddeler üzerinden eşitleme ilişkisi kurulur ve gruplar arasındaki yetenek farklılıkları kontrol altına alınır (Angoff, 1984). Ortak test, testin bütünü ile aynı kapsama ve istatistiksel özelliklere sahip olmalı, testin bütünü temsil eden mini bir versiyonu olmalıdır (Kolen ve Brennan, 2014).

### **Kalibrasyon, Ölçek Dönüştürme ve Eşitleme**

Test eşitlemenin ilk adımı madde ve yetenek parametrelerinin kalibrasyonudur. Eş zamanlı (concurrent) kalibrasyon ve ayrı (separate) kalibrasyon yöntemleri ile madde ve yetenek parametrelerinin kalibrasyonu yapılabilir. Eş zamanlı kalibrasyonda test formlarına ait madde ve yetenek parametreleri eş zamanlı olarak tek seferde kestirilirken, ayrı kalibrasyonda test formlarında yer alan madde ve yetenek parametreleri her bir form için ayrı kestirilir. Eş zamanlı kalibrasyon sonucunda iki test formuna ait bütün parametreler aynı ölçekte elde edilirken, ayrı kalibrasyon yöntemiyle kestirilen madde parametreleri aynı ölçek üzerinde yer almaz. Ayrı kalibrasyon yapıldığında, örtük özelliğe ait ölçeğin ortalama ve standart sapması örtük özelliğin sahip olduğu dağılıma sabitlenmektedir (Hanson ve Beguin, 2002).

Eş zamanlı kalibrasyon uygulaması uygulanış açısından daha pratik gözükmesine rağmen, yapılan çalışmalar veri yapısının tek boyutlu olmadığı durumlarda ayrı kalibrasyon uygulamasının daha az hataya sahip ve güvenilir sonuçlar verdiğini ortaya koymuştur (Kim ve Kolen, 2006; Kolen ve Brennan, 2014). Çok boyutlu veri yapıları üzerinden analiz süreçleri yürütülen bu çalışmada madde ve yetenek parametrelerinin kestiriminde ayrı kalibrasyon yöntemi kullanılmıştır. Aşağıda ayrı kalibrasyon yöntemi açıklanmıştır.



**Ayrı Kalibrasyon (Separate Calibration).** Eşdeğer olmayan gruplar ortak test deseninde yapılan uygulama sonucu elde edilen madde parametreleri, gruplar farklı olduğundan aynı ölçekte yer almazlar. Farklı ölçeklerde yer alan madde parametrelerin karşılaştırılması doğru olmayacağından, kestirilen madde parametrelerinin aynı ölçeğe yerleştirilmesi gerekir. Ölçek dönüştürmek için bağlama (linking) katsayıları olan A (eğim) ve B (kesişim) sabitlerinin hesaplanması gerekir (Kolen ve Brennan, 2014). Bir teste ait olan yeni form I ve eski form J olarak düşünüldüğünde, ayrı kalibrasyon yöntemi uygulanarak Test I'daki i kişisine ait  $\theta$  yetenek parametresinin test J'deki karşılığı eşitlik 9'da verilen matematiksel ifade yardımıyla hesaplanır.

$$\theta_{Ji} = A\theta_{Ii} + B \quad (9)$$

A: Eğim sabiti

B: Kesişim sabiti

Eşitlik 10'da Test I'daki madde parametrelerinin Test J'ye dönüşümü verilmiştir. Eşitlikte görüleceği üzere, şans parametresi (c parametresi) için ölçek dönüştürme işlemine gerek yoktur (Kolen ve Brennan, 2014).

$$\begin{aligned} a_{Jj} &= \frac{a_{Ij}}{A} \\ b_{Jj} &= Ab_{Ij} + B \\ c_{Jj} &= c_{Ij} \end{aligned} \quad (10)$$

$a_{Ij}, b_{Ij}, c_{Ij}$  : j maddesinin I testine ait madde parametreleri

$a_{Jj}, b_{Jj}, c_{Jj}$  : j maddesinin J testine ait yeniden ölçeklenmiş madde parametreleri

ÇBMTK'da 3PLM için yetenek parametresinin dönüşümü tek boyutlu MTK ile aynı, madde parametrelerine uygulanacak dönüşüm ise tek boyutlu MTK'dan farklıdır. Yao ve Boughton (2009), madde parametrelerinin 3PLM'de dönüşümünü eşitlik 11'de yer aldığı şekilde açıklamıştır.

$$\begin{aligned}
\vec{\beta}_{2j}^* &= A^{-1} \vec{\beta}_{2j} \\
\beta_{1j}^* &= \beta_{1j} + \beta_{2j}^T A^{-1} \vec{B} \\
\beta_{3j}^* &= \beta_{3j}
\end{aligned} \tag{11}$$

$\beta_{2j}$  : j maddesine ait ayırt edicilik parametresi için D boyuta ilişkin vektör (MDISC)

$\beta_{1j}$  : j maddesine ait güçlük parametresi

$\beta_{3j}$  : j maddesine ait şans parametresi

$A$  : DxD'lik dönüştürme matrisi (transformation matrix)

$\vec{B}$  : Konum vektörü (location vector)

İki kategorili MTK modellerinde ayrı kalibrasyon için kullanılan ölçek dönüştürme yöntemleri moment yöntemleri ve karakteristik eğri yöntemleri olarak ikiye ayrılır. Moment yöntemleri ortalama-ortalama ve ortalama-sigma olarak açıklanabilir. Karakteristik eğri yöntemleri ise Stocking-Lord ve Haebara yöntemleridir. Bu çalışmada ölçek dönüştürme yöntemleri olarak ortalama-ortalama, ortalama-sigma ve Stocking-Lord yöntemleri kullanıldığından, bu yöntemlerin açıklamalarına ve hesaplama yöntemlerine yer verilmiştir.

**Ortalama-Ortalama (Mean-Mean) Yöntemi.** Bu yöntemde test formlarında yer alan ortak maddelerden kestirilen ayırıcılık ve güçlük parametrelerinin ortalamaları alınarak A ve B sabitleri elde edilir (Loyd ve Hoover, 1980). Eşitlik 12'de A ve B sabitlerinin elde edilmesine ilişkin ifade verilmiştir.

$$\begin{aligned}
A &= \frac{\mu(a_i)}{\mu(a_j)} \\
B &= \mu(b_j) - A\mu(b_i)
\end{aligned} \tag{12}$$

$\mu(a_i), \mu(a_j)$ : Sırasıyla i ve j ölçeklerindeki ortak maddelerden kestirilen ayırıcılık parametrelerinin ortalaması

$\mu(b_i), \mu(b_j)$ : Sırasıyla i ve j ölçeklerindeki ortak maddelerden kestirilen güçlük parametrelerinin ortalaması

**Ortalama-Sigma (Mean-Sigma) Yöntemi.** Bu yöntemde A ve B sabitleri, test formlarında yer alan ortak maddelerden kestirilen güçlük parametrelerinin ortalama ve standart sapmaları hesaplanarak elde edilir (Marco, 1977). Eşitlik 13'te ortalama-sigma yöntemine göre A ve B sabitlerinin hesaplanmasına ilişkin eşitlik verilmiştir.

$$A = \frac{\sigma(b_j)}{\sigma(b_i)} \quad (13)$$

$$B = \mu(b_j) - A\mu(b_i)$$

$\mu(b_i), \mu(b_j)$ : Sırasıyla i ve j ölçeklerindeki ortak maddelerden kestirilen güçlük parametrelerinin ortalaması

$\sigma(b_j), \sigma(b_i)$ : Sırasıyla j ve i ölçeklerindeki ortak maddelerden kestirilen güçlük parametrelerinin standart sapması

**Stocking-Lord Yöntemi.** Bu yöntemde madde parametrelerinin tümü aynı anda dikkate alınarak hesaplama yapılmaktadır. A ve B sabitleri elde edilirken, test formlarında yer alan ortak maddelerin madde karakteristik eğrileri arasındaki farka göre tanımlanan ölçüt fonksiyon minimize edilmektedir. Buna göre madde parametreleri üzerinden toplamın farkının karesi alınır. Stocking-Lord yöntemine ilişkin matematiksel ifade eşitlik 14'te verilmiştir (Stocking ve Lord, 1983).

$$SLdiff(\theta_i) = \left[ \sum_{j:V} p_{ij}(\theta_{ji}; a_{Jj}, b_{Jj}, c_{Jj}) - \sum_{j:V} p_{ij} \left( \theta_{ji}; \frac{a_{Ij}}{A}, Ab_{Ij} + B, c_{Ij} \right) \right]^2 \quad (14)$$

- $p_{ij}$  : i yanıtlayıcısı ve j maddesi için madde karakteristik fonksiyonu  
 $A$  : Eğim sabiti  
 $B$  : Kesişim sabiti  
 $p_{ij}$  : i yanıtlayıcısı ve j maddesi için madde karakteristik fonksiyonu  
 $a_{Jj}, b_{Jj}, c_{Jj}$  : J ölçeğindeki j. ortak madde için madde parametreleri  
 $a_{Ij}, b_{Ij}, c_{Ij}$  : I ölçeğindeki j. ortak madde için madde parametreleri  
 $j:V$  : Toplamın ortak maddeler üzerinden hesaplandığını gösterir.

SLdiff, sınava girenler üzerinden toplanır ve tahmin, eşitlik 15'teki kriteri minimize eden A ve B kombinasyonunu bulmak için ileri (Stocking ve Lord, 1983).

$$SLcrit = \sum_i SLdiff(\theta_i) \quad (15)$$

Yapılan çalışmalar incelendiğinde, karakteristik eğri yöntemlerinin moment yöntemlere oranla daha kararlı sonuçlar verdiği görülmektedir (Atar ve Yeşiltaş, 2017; Baker ve Al-Karni, 1991; Hanson ve Beguin, 2002; Kim ve Kolen, 2006). Buna ek olarak ortalama-ortalama yönteminin ortalama-sigma yöntemine oranla daha kararlı sonuçlar ürettiğini ortaya koyan çalışmalar da mevcuttur (Ogasawara, 2000; Gök ve Kelecioğlu, 2014)

Bir test eşitleme sürecinde desen seçiminin ardından eşitleme yöntemine karar verilir ve buna göre test formlarına ait madde ve yetenek parametreleri kestirilir. Parametreler kestirildikten sonra ortak bir ölçeğe yerleştirme işlemi yapılır. Ölçek dönüştürme işlemi eşdeğer olmayan gruplar ortak test deseni kullanıldığında yapılmalıdır. Tek ve eşdeğer grup desenlerinde kestirilen parametreler aynı ölçekte olacağından, bu aşamaya gerek yoktur. Madde ve yetenek parametreleri ortak bir ölçeğe yerleştirildikten sonra, test puanlarının raporlanacağı ölçek belirlenir. Test puanları yetenek parametreleri üzerinden raporlanacaksa süreç tamamlanır, fakat eğer raporlaştırma gerçek puanlar üzerinden yapılacaksa, gerçek puanlar farklı yetenek düzeylerine göre kestirilerek her iki forma ait gerçek puanlar eşitlenmelidir (Kabasakal, 2014).

## İlgili Araştırmalar

Bu bölümde, alan yazında sırasıyla tek boyutlu ve çok boyutlu ölçek dönüştürme/eşitleme yapılan çalışmalara ayrı başlıklar halinde değinilecektir.

### Tek Boyutlu Eşitleme Yapılan Çalışmalar

Hanson ve Beguin (2002), yaptıkları araştırmada ortalama-ortalama, ortalama-sigma ve karakteristik eğri yöntemlerinin performansını belirlemeyi amaçlamışlardır. Simülatif verilerin kullanıldığı araştırmada ortak maddelerin yer aldığı ve her biri 60 maddeden oluşan iki test formu kullanılmıştır. Örneklem büyüklüğü (1000, 3000), ortak madde sayısı (10, 20) ve grupların eşdeğer olup olmaması araştırma koşulları olarak alınmıştır. Araştırma sonucunda örneklem büyüklüğü ve ortak madde sayısı arttıkça, daha düşük hata elde edildiği görülmüştür. Eşdeğer gruplar kullanıldığında, eşdeğer olmayan gruba kıyasla daha az hatalı sonuçlar elde edilmiştir.

Kim ve Lee (2006), karma testler için sırasıyla ortalama-ortalama, ortalama-sigma, Stocking-Lord ve Haebara ölçek dönüştürme yöntemlerinin uzantısını geliştirmiş ve çeşitli koşullar (grupların yetenek dağılımı, örneklem büyüklüğü, karma formatta oluşturulan test kompozisyonu ve ortak madde seti) altında bu yöntemlerin performanslarını incelemişlerdir. Simülasyon verilerinin kullanıldığı çalışmada 3PLM altında kısmi puanlama modeli kullanılarak karma testler oluşturulmuştur. Değerlendirme ölçütleri olarak SB (square bias) ve MSE (mean squared errors) indekslerinin kullanıldığı araştırmada karakteristik eğri yöntemlerinin moment yöntemlerine oranla daha düşük eşitleme hataları ürettiği ve kullanılan yöntemler içinde en düşük hata veren yöntemin Haebara yöntemi olduğu bulgusuna ulaşılmıştır. Bununla birlikte ortak madde setinde madde türünün aynı olduğu durumlara kıyasla ortak madde setinin karma şekilde oluşturulduğu koşulların daha iyi performans gösterdiği ve eşdeğer gruplarla yürütülen eşitlemenin daha az hataya sahip olduğu bulunmuştur.

Speron (2009) çalışmasında moment ve karakteristik eğri ölçek dönüştürme yöntemlerini karşılaştırmayı amaçlamıştır. Verilerin 2PLM'ye göre üretildiği araştırmada simülasyon koşulu olarak örneklem büyüklüğü, ortak maddelerin sayısı, ortak maddelerin toplam testi temsil etme yüzdesi ve grupların yetenek dağılımları alınmıştır. Sonuç olarak, karakteristik eğri yöntemleri ve ortalama-ortalama yöntemlerinin benzer sonuçlar verdiği ve daha az hatalı kestirime neden olduğu görülmüştür. Bununla birlikte araştırmada ortalama-sigma yönteminin en yüksek hataya sahip yöntem olduğu da belirlenmiştir. Örneklem büyüklüğü arttıkça eşitleme hatalarının azaldığı gözlenmiştir.

Kilmen ve Demirtaşlı (2012), çalışmalarında ortalama-ortalama, ortalama-sigma, Heabera ve Stocking-Lord ölçek dönüştürme yöntemlerini yetenek dağılımı ve örneklem büyüklüğü değişkenleri açısından karşılaştırmıştır. Eşdeğer olmayan gruplar ortak test deseninin kullanıldığı bu çalışmada 600 maddeden oluşan 1-0 veri matrisi kullanılmıştır. Çalışma sonucunda, farklı örneklem büyüklüğü (500 ve 1000) ile benzer ve farklı yetenek dağılımına sahip grupların bulunduğu durumlarda, Stocking-Lord yöntemi ile test eşitleme uygulamasının daha az eşitleme hatası verdiği bulunmuştur.

Gök ve Kelecioğlu (2014), farklı koşullar altında üretilen test formlarını MTK'ya dayalı kestirim yöntemleri ile eşitlemiş ve bu yöntemlerden elde edilen sonuçları karşılaştırmıştır. Araştırma 2PLM ve 3PLM'ye uyumlu iki kategorili simülasyon verileri kullanılmıştır. Eşitleme deseni olarak eşdeğer olmayan gruplarda ortak test deseni kullanılmış ve elde edilen sonuçlar RMSE ölçütüne göre değerlendirilmiştir. Örneklem büyüklüğü 3000, test uzunluğu 80 ve yetenek dağılımı benzer olan gruplarda ortalama-ortalama yöntemi kullanıldığında en düşük eşitleme hatalarının elde edildiği görülmüştür. Test uzunluğu ve örneklem büyüklüğü arttıkça benzer yetenek dağılımına sahip gruplarda eşitleme yöntemlerinin daha az eşitleme hatası verdiği sonucuna varılmıştır.

Bökeoğlu, Uçar ve Balta (2022), MTK'ya ölçek dönüştürme yöntemlerinin farklı koşullar altındaki eşitleme hatalarını karşılaştırmıştır. Çalışmanın koşulları örneklem büyüklüğü, test uzunluğu, ortak madde oranı, parametre kestirim modeli ve yetenek

dağılımı olarak belirlenmiştir. Değerlendirme ölçütü olarak RMSD belirlenmiştir. SL yöntemine ilişkin RMSD değerlerinin diğer yöntemlere göre daha yüksek olduğu, 2PLM ve 3PLM kullanılan durumlarda benzer sonuçlar elde edildiği, örneklem büyüklüğü ve test uzunluğu arttıkça Stocking-Lord yöntemi hariç diğer yöntemlerin eşitleme hatalarında azalma gözleendiği, ortak madde oranının %40 ve gruplar benzer yetenek dağılımına sahip olduğunda yöntemlerin düşük RMSD değerleri ürettiği sonuçları elde edilmiştir.

Yapılan tek boyutlu ölçek dönüştürme çalışmaları değerlendirildiğinde, daha çok simülasyon verileri kullanıldığı ve eşdeğer grup deseni kullanıldığı durumlarda eşdeğer olmayan gruplar ortak test desenine oranla daha düşük eşitleme hataları elde edildiği bulgusuna ulaşılmıştır. Bununla birlikte test uzunluğu, örneklem büyüklüğü ve ortak madde oranında artırıldıkça daha düşük eşitleme hataları elde edilmiştir. Ortalama-sigma yönteminden elde edilen hatalar daha yüksek bulunmakla birlikte moment yöntemleri ya da karakteristik eğri yöntemlerinin birbirine yakın sonuçlar verdiği konusunda bulgular bulunmaktadır.

### **Çok Boyutlu Eşitleme Yapılan Çalışmalar**

Brossman ve Lee (2013), ÇBMTK çerçevesinde gözlenen ve gerçek puan eşitleme yöntemleri geliştirmişlerdir. Çalışma kapsamında iki gözlenen puan ve bir gerçek puan eşitleme yöntemi geliştirilmiş olup, gözlenen puan yöntemi tek boyutlu madde tepki kuramındaki gözlenen puan eşitlemenin doğrudan uzantısı olarak sunulmuş ve "Tam ÇBMTK Gözlenen Puan Eşitleme Yöntemi" olarak isimlendirilmiştir. Çalışmada yer alan diğer gerçek ve gözlenen puan yöntemleri, tek boyutlu yaklaşım prosedürleri içermektedir. Bu yöntemler sırasıyla "ÇBMTK Gerçek Puan Eşitlemenin Tek Boyutlu Yaklaşımı" ve "ÇBMTK Gözlenen Puan Eşitlemenin Tek Boyutlu Yaklaşımı" olarak isimlendirilmiştir. Çalışmada buna ek olarak tek boyutlu gözlenen ve gerçek puan eşitleme yöntemleri ile eşit yüzdelli eşitleme yöntemi de kullanılmıştır. ÇBMTK eşitleme yöntemlerinin, TBMTK eşitleme yöntemlerine oranla eşit yüzdelli eşitleme yöntemiyle daha yakın sonuçlar verdiği

sonucuna ulařılmış, bunun sebebi ise eřit yzdelikli eřitleme ynteminin MTK'nın tek boyutluluk varsayımını ihmal etmemesi olarak aıklanmıřtır.

Yao ve Boughton (2009), iki boyutlu veri setiyle yaptıkları simlasyon alıřmasında, 3000 rneklem blyklęu kullanılarak basit ve karmařık yapılı maddelerin bir karıřımı ile 10 farklı rneklem daęılımında, ortak madde seti uzunluklarının (6, 8, 16, 32, 60) deęiřen kořulları altında, gerek verilerden turetilen parametreler kullanarak ok boyutlu lek dnřtrme yapmıřlardır. lek dnřtrme yntemi olarak Stocking-Lord yntemi kullanılmıřtır. Tm rneklemeler iin, iyi seilmiř bir ortak madde setinde parametrelerin eřitlendikten sonra iyi olacak řekilde dzeldięi sonucuna ulařılmıřtır.

Lee, Lee ve Brennan (2014), yaptıkları alıřmada ok boyutlu testler zerinde eřitli eřitleme yntemlerinin performansını karřılařtırmıřlardır. alıřmada BMTK erevesinde simlasyon verileri kullanılmıřtır. Eřdeęer gruplar deseninin kullanıldıęı alıřmada (1) Tek Boyutlu MTK gzlenen puan eřitleme, (2) Tek Boyutlu MTK gerek puan eřitleme, (3) Tam BMTK gzlenen puan eřitleme, (4) BMTK gerek puan eřitleme ynteminin tek boyutlu yaklařımı, (5) MIRT gzlenen puan eřitleme ynteminin ok boyutlu yaklařımı ve (6) eřit yzdelikli eřitleme yntemleri kullanılmıřtır. Drt faktrn (test uzunluęu, rneklem blyklęu, formların glk dzeyleri arasındaki farklılıklar ve boyutlar arasındaki korelasyon) etkisi arařtırılmıř ve her faktr iin iki kořul yaratılmıřtır. Kořullar dikkate alınarak elde edilen eřitleme iliřkileri ve genel eřitleme iliřkileri yanlılık, standart hata ve genel hata aısından deęerlendirilmiřtir. alıřma sonunda řu sonular elde edilmiřtir. (1) Tam BMTK yntemi, zellikle boyutlar arasındaki korelasyonun dřk olduęu durumlarda, dięer eřitleme yntemlerinden daha az hataya sahip eřitleme sonuları retmiřtir, (2) Test edilen tm kořullar iin eřit yzdelikli eřitleme ynteminin TBMTK yntemlerine oranla daha fazla miktarda standart hata ve genel hata vermesi olasıdır; (3) Formlar arasındaki farklılıkların dřk, rneklem blyklęnn ve test uzunluęunun fazla olduęu kořullarda ok boyutlu testleri eřitlemek daha doęru sonulara ulařtırmaktadır, (4) ok boyutlu testlerin kullanıldıęı durumlar dahil olmak zere (tek boyutluluk varsayımı ihmal edilse bile)



tek boyutlu MTK yöntemlerinin daha doğru eşitleme sonuçları ürettiği bulgusuna ulaşılmıştır, (5) gözlenen ya da gerçek puan eşitleme yöntemlerinden hangisinin kullanıldığı eşitleme sonuçları üzerinde önemli bir farklılık göstermemektedir.

Lee ve Lee (2016) çalışmalarında karma formatlı testler için bifaktör çok boyutlu madde tepki kuramı (BF-MIRT) gözlenen puan eşitleme prosedürleri geliştirmek ve önerilen yöntemlerin göreceli uygunluğunu araştırmayı amaçlamışlardır. Büyük ölçekli bir test programından elde edilen veriler kullanılarak üç tür sözde veri seti formüle edilmiştir: eşleştirilmiş örnekler, sözde formlar ve simüle edilmiş veri setleri. Birincil genel faktörün etkisi kontrol edildikten sonra karma format testlerinde çok küçük format içi artık bağımlılığı bulunmuştur. Tek boyutlu IRT ve BF-MIRT eşitleme yöntemleri, bu çalışmada kullanılan veriler için benzer eşitleme sonuçları üretmiştir.

Atar ve Yeşiltaş (2017), çok boyutlu veri için düzelenmiş olan ortalama-ortalama ve ortalama-sigma ve Stocking-Lord eşitleme yöntemlerinin performanslarını eşdeğer olmayan gruplar ortak test deseninde örneklem büyüklüğü, yetenek dağılımı, boyutlar arasındaki korelasyon ve testte yer alan ortak madde yüzdesi koşulları altında incelemiştir. Değerlendirme ölçütü olarak eşitleme hatası (RMSE) ve yanlılık değerleri kullanılmıştır. Hem ayırt edicilik hem de güçlük parametre kestirimlerinde çoğu koşul için Stocking-Lord yönteminin daha düşük RMSE ve yanlılık değerleri verdiği sonucuna ulaşılmıştır.

Kim (2018), çalışmasında basit yapıları bir ÇBMTK modeli kullanarak gerçek puan eşitlemeye yönelik teorik ve kavramsal bir çerçeve geliştirmeyi, bu sayede ÇBMTK ile ilgili alanyazına katkıda bulunmayı hedeflemiştir. Bu çalışmada, basit yapıları ÇBMTK gerçek puan eşitleme yönteminin performansı farklı veri türleri kullanılarak dört farklı çalışma ile incelenmiştir: (1) gerçek veri, (2) simüle veri, (3) pseudo form verisi, ve (4) kimlik eşitleme yöntemiyle elde edilmiş tek form verisi. Gerçek puan eşitlemenin yanı sıra, bu yöntemin diğer yöntemlere göre göreceli avantajlarını ortaya koymak amacıyla dört alternatif yöntem daha analizlere dahil edilmiştir: (a) ön düzleştirilmiş eşit yüzdelikli eşitleme, (b) Tek boyutlu MTK gerçek puan eşitleme, (c) Tek Boyutlu MTK gözlenen puan eşitleme, (d) Basit

yapılı ÇBMTK gözlenen puan eşitleme. Sonuç olarak, çalışmada önerilen gerçek puan eşitleme yönteminin diğer yöntemlerle yakın performans gösterdiği bulgusuna ulaşılmıştır. Bunun yanında gerçek puan eşitleme yönteminin, geleneksel tek boyutlu eşitleme yöntemine kıyasla daha doğru eşitleme sonuçları verdiği sonucuna ulaşılmıştır.

Kim ve Lee (2023), çalışmalarında basit yapılı çok boyutlu madde tepki kuramı eşitleme yöntemlerinin çeşitli varyantlarını araştırmışlardır. İki farklı eşitleme deseni için önerilen eşitleme yöntemlerinin uygulanabilirliğini göstermek amacıyla eşdeğer gruplar deseni ve eşdeğer olmayan gruplar ortak test deseni kullanılmıştır. Bulgular, verilerin çok boyutluluğa ilişkin kanıtlar sergilediği durumlarda çok boyutlu ve tek boyutlu yaklaşımlar arasında bazı önemli farklılıklar olduğunu göstermiştir. Bununla birlikte, önerilen yöntemlerden bazıları, mevcut yöntemlerin çoğu tarafından başarısız olan, hem bölüm hem de bileşim düzeyindeki puanlar için eşitleme sonuçları sağlamada başarılı olmuştur. Özellikle yüksek boyutlu verileri eşitlemede için bir dizi kareleme noktası ve ağırlık kullanan geleneksel yöntemin hesaplama açısından yoğun olduğu ortaya çıkmıştır. Bu çalışma ayrıca çok değişkenli bir gözlenen puan dağılımına dayanmayan basit yapılı bir gerçek puan eşitleme yöntemi önermiştir.

Yapılan çok boyutlu eşitleme çalışmaları değerlendirildiğinde, yeni eşitleme yöntemlerinin önerildiği ve bunların performanslarının incelendiği çalışmalar mevcuttur. Boyutlar arası korelasyonun düşük olduğu durumlarda çok boyutlu eşitleme yöntemlerinin daha düşük eşitleme hatası ürettiği ortaya konulmuştur. Bunun yanında örneklem büyüklüğü ve ortak madde oranının artırıldığı durumlarda daha düşük eşitleme hataları elde edilmiştir. Veri yapısının çok boyutlu olduğu durumda uygulanan çok boyutlu ve tek boyutlu yöntemler arasında elde edilen hatalar bakımından farklılıklar bulunmuştur.

## Bölüm 3

### Yöntem

Bu bölümde araştırmanın türü, verilerin üretilmesinde kullanılan simülasyon koşulları, verilerin üretilmesi, verilerin analizi ve ölçek dönüştürme süreci ile değerlendirme kriteri başlıkları altında açıklamalar yapılmıştır.

#### Araştırmanın Türü

Bu çalışma, çok boyutlu testlere çok boyutlu ve tek boyutlu yöntemler kullanılarak ölçek dönüştürme yapılmasını ve ele alınan koşullar altında elde edilen eşitleme hatalarının değişim düzeyini incelemeyi amaçlamaktadır. Araştırma kapsamında yer alan koşulların etkisinin incelenebilmesi amacıyla simülasyon verileri kullanılmıştır. Sosyal bilimlerde deneysel geliştirme, araştırma sayesinde elde edilen bilgiyi sınama ve değerlendirme amaçları için yürütülen işlemsel programlara tercüme etme işlemi olarak tanımlanabilir (TÜBİTAK, 2002). Bu bağlamda bu çalışmanın simülatif deneysel araştırma olduğu ifade edilebilir.

#### Simülasyon Koşulları

Araştırma kapsamında ele alınan koşulların (örneklem büyüklüğü, boyutlar arası korelasyon, ortak madde oranı, gruplar arası yetenek dağılımı farkı, parametre kestirim modeli) gerçek veri yapılarında tamamının sağlanmasının olanaklı olmamasından dolayı simülasyon verileri kullanılmıştır. Araştırma kapsamında yapılan ölçek dönüştürme çalışmalarından elde edilen eşitleme hataları örneklem büyüklüğü, boyutlar arası korelasyon, ortak madde oranı, gruplar arası yetenek dağılımı farkı ve parametre kestirim modeli açısından incelenmiştir. Araştırma kapsamında iki boyutlu veri setleri üzerinden analizler gerçekleştirilmiştir.

R yazılımı kullanılarak araştırmada yer alan koşullar çerçevesinde testin 40 maddeden oluşan her iki formuna (X ve Y) ilişkin çok boyutlu veri setleri ile madde ve

yetenek parametreleri üretilmiş, yine R yazılımı kullanılarak madde ve yetenek parametrelerinden iki kategorili madde yanıt verileri elde edilmiştir. Testte yer alan 40 maddenin 20 tanesi bir boyutta, 20 tanesi diğer boyutta olacak şekilde iki boyutlu basit yapıları veri seti elde edilmiştir. Araştırmada yer alan değişkenler ve koşullar tablo 1'de açıklanmıştır.

**Tablo 1**

*Çalışma Kapsamında Yer Alan Değişkenler ve Koşullar*

Değişkenler	Koşullar	Koşul Sayısı
Örnekleme büyüklüğü	1000-2000	2
Boyutlar arası korelasyon	0.1-0.5-0.9	3
Ortak madde oranı	%20-%40	2
Gruplar arası yetenek dağılımı farkı	0.05-0.5	2
Parametre kestirim modeli	2PLM-3PLM	2

**Örnekleme Büyüklüğü.** Ölçek dönüştürme yöntemlerinde, örneklem büyüklüğünün eşitleme hatasına olan etkisini inceleyen çalışmalar bulunmaktadır. Skaggs ve Lissitz (1986), 3PLM için örneklem büyüklüğünün en az 1000 olması gerektiğini belirtmiş, Gübeş (2019) çalışmasında örneklem büyüklüğünü 1000 ve 2000 olarak, bunun yanında Hanson ve Beguin (2002), Gök ve Kelecioğlu (2014) ve Kumlu (2019) çalışmalarında örneklem büyüklüğünü 1000 ve 3000 olarak almışlardır. Bu çalışmada ise örneklem büyüklüğü 1000 ve 2000 olarak alınmıştır.

**Boyutlar Arası Korelasyon.** Çalışma kapsamında boyutlar arası korelasyon düşük (0.1), orta (0.5) ve yüksek (0.9) düzeyde olacak şekilde koşullar oluşturulmuştur. Boyutlar arasındaki korelasyonun fazla olması testin tek boyutluluğuna kanıt olarak gösterilebilir (Zhang, 2009). Beguin ve Hanson (2001), boyutlar arasındaki korelasyon artışının çok

boyutlu model parametre kestirimi kullanıldığında toplam hatada artışa neden olduğunu gözlemlemişlerdir. Gübeş (2019), iki boyutlu bir test verisine tek boyutlu ölçek dönüştürme yöntemleri uygulamış, boyutlar arası korelasyon azaldıkça daha yüksek eşitleme hataları elde etmiştir.

**Ortak Madde Oranı.** Eşdeğer olmayan gruplar ortak test deseninde ölçek dönüştürme işlemi ortak maddeler (ankor) aracılığıyla yapılmaktadır. Bu desende ankor formu iç ankor ve dış ankor olmak üzere ikiye ayrılır. Ortak maddeler bireyin toplam puanına dahilse iç ankor, dahil değilse dış ankor olarak adlandırılır (Kolen ve Brennan, 2014). Bu çalışmada iç ankor test kullanılmıştır. Angoff (1984) ve Kolen ve Brennan (2014), test formlarında yer alan ortak madde sayısının 20 maddeden ya da toplam madde sayısının %20'sinden az olmaması gerektiğini belirtmişlerdir. Bu çalışmada ortak madde oranı %20 (8 madde) ve %40 (16 madde) olarak alınmıştır. Ortak madde oranının %20 olduğu durumda ortak madde setinde yer alan 8 maddenin 4 tanesi bir boyutta, 4 tanesi diğer boyutta yer alan maddelerden oluşmaktadır. Ortak madde oranının %40 olduğu durumda ise ortak madde setinde yer alan 16 maddenin 8 tanesi bir boyutta, 8 tanesi ise diğer boyutta yer alacak şekilde veri seti oluşturulmuştur.

**Parametre Kestirim Modeli.** 2PLM ve 3PLM kullanılarak yapılan ölçek dönüştürmenin eşitleme hatasına olan etkisini inceleyen araştırmalar mevcuttur (Gök ve Kelecioğlu, 2014; Kim ve Kolen, 2006; Kim ve Lee, 2006). Buna göre, kullanılan modelin ölçek dönüştürme ve test eşitleme süreci üzerinde etkisi olduğu belirtilmiştir. Bu araştırmanın verileri 2PLM ve 3PLM'ye göre üretilmiş ve bu şekilde 3PLM kullanıldığı durumda şans ile doğru yanıtlanma olasılığının eşitleme hatasına olan etkisini ortaya koyabilmek amaçlanmıştır. Madde ayırt edicilik düzeyini gösteren a parametresi, her iki form için değerleri 0.6 ile 2 arasında değişen tek biçimli dağılımdan üretilmiştir. Güçlük düzeyini gösteren b parametresi (-3,+3) aralığında ortalaması 0, standart sapması 1 olan normal dağılımdan ve şans başarısını ifade eden c parametresi 0.01 ile 0.25 arasında değişen tek biçimli (uniform) dağılımdan üretilmiştir.

**Gruplar Arası Yetenek Dağılımı.** Gruplar arası yetenek dağılımlarına ait ortalamalar arasında bulunan farka dayalı olarak yapılan sınıflamaya göre eğer fark 0.05 ile 0.10 aralığında ise “nispeten büyük”, 0.25 ve daha büyük olduğunda ise “çok büyük” olarak ifade edilebilir (Wang, Lee, Brennan ve Kolen, 2008). Bu çalışmada, yukarıda belirtilen değerler dikkate alınarak gruplar arası ortalama yetenek farkının düşük (0.05) ve yüksek (0.5) olduğu iki koşul belirlenmiştir. Yeni formu alan grubun yetenek dağılımı  $N(0.05,1)$  ve  $N(0.5,1)$  olarak, eski formu alan grubun yetenek dağılımı ise  $N(0,1)$  olarak standart normal dağılım gösterecek şekilde yetenek parametreleri elde edilmiştir.

### **Verilerin Üretilmesi**

Çalışma kapsamında yer alan değişkenler ve koşullar tablo 1’de açıklanmıştır. Tablo 1’de görüldüğü üzere, bu çalışmada her bir ölçek dönüştürme yöntemine (tek ve çok boyutlu, 2 koşul) ilişkin örneklem büyüklüğü (2 koşul), boyutlar arası korelasyon (3 koşul), ortak madde oranı (2 koşul), gruplar arası yetenek dağılımı (2 koşul) ve parametre kestirim modeli (2 koşul) olmak üzere toplam 96 koşul incelenmiştir.

Literatürde yapılan çalışmalar incelendiğinde, araştırma sonuçlarının tutarlı ve kararlı olabilmesi açısından her veri seti için en az 50 tekrar yapıldığı (Hanson ve Beguin, 2002; Kim ve Kolen, 2006), kararlı sonuçların elde edilebilmesi açısından 100 tekrarın yeterli olduğu görülmüştür. Bu araştırmada da her bir koşul için veri seti üretilirken tekrar sayısı 100 olarak alınmıştır.

### **Ölçek Dönüştürme Süreci ve Verilerin Analizi**

Çalışmada hem tek boyutlu hem de çok boyutlu ortalama-ortalama (OO), ortalama-sigma (OS) ve Stocking-Lord (SL) ölçek dönüştürme yöntemlerinden elde edilen eşitleme hatalarını karşılaştırabilmek için R yazılımında testin her iki formuna (X ve Y) ilişkin 9600 veri seti üretilmiştir. R yazılımı üzerinden mirt paketi (Chalmers, 2012) kullanılarak üretilen madde ve yetenek parametrelerinden iki kategorili (1-0) madde yanıt verileri üretilmiştir. Testin her iki formu için IRTPRO 4.2 yazılımında 2PLM ve 3PLM için ayrı ayrı tek boyutlu

ve çok boyutlu (basit yapılı) olacak şekilde parametre kestirimleri yapılmıştır. Tek boyutlu parametre kestiriminde Bock-Aitkin yöntemi, çok boyutlu parametre kestiriminde ise Markov Chain Monte Carlo (MC-MC) yöntemi kullanılmıştır. Her iki formdan elde edilen parametreleri aynı ölçek üzerine yerleştirmek için birinci formda yer alan maddelerin parametrelerine ikinci formda yer alan ortak madde parametreleri baz alınarak tek boyutlu ve çok boyutlu madde tepki kuramı üzerinden ayrı kalibrasyon yöntemleri (ortalama-ortalama, ortalama-sigma ve Stocking-Lord) kullanılarak ölçek dönüştürme yapılmıştır. Tek boyutlu ölçek dönüştürme için IRTEQ (Han, 2009), çok boyutlu ölçek dönüştürme için Linkmirt (Yao, 2009) yazılımı kullanılmıştır. Çalışmada kullanılan yazılımlar, veri setlerine ilişkin belirlenen 100 tekrarın tek seferde analiz edilebilmesi amacıyla R yazılımı ile grup komut dosyaları üzerinden kontrol edilmiştir. OO, OS ve SL ölçek dönüştürme yöntemleri kullanılarak eğim (A) ve kesişim (B) sabitleri elde edilmiş ve bu sabitler kullanılarak ölçek dönüştürme işlemi gerçekleştirilmiştir.

### Değerlendirme Kriteri

Her bir ölçek dönüştürme yöntemi için dönüştürülen madde ve yetenek parametrelerine ait hata miktarını veren RMSE (Root Mean Squared Error) değeri hesaplanmıştır. Madde ve yetenek parametrelerinin her biri için, her tekrardan sonra kestirilen RMSE değerlerinin ortalaması alınarak her bir koşul için her bir parametreye ilişkin ortalama RMSE değeri elde edilmiştir. Eşitlik 16'da RMSE değerinin hesaplanmasına ilişkin matematiksel ifade verilmiştir.

$$RMSE(\tau_j) = \sqrt{\frac{\sum_{r=1}^R (\tau_{jr} - \tau_j)^2}{R}} \quad (16)$$

$\tau_j$  : j parametresinin gerçek değeri

$\tau_{jr}$  : Her bir tekrar için (r=1, 2, 3, ..., R) için j parametresine ait kestirilen değer

### **Çalışmada Ele Alınan Koşullara Göre Elde Edilen RMSE Değerlerinin Karşılaştırılması**

Çalışma kapsamında ele alınan koşullar çerçevesinde elde edilen RMSE değerlerinin ortalamaları arasında istatistiksel olarak anlamlı farklılık olup olmadığının belirlenmesi amacıyla karşılaştırma testleri yapılmıştır. Tüm koşullar göz önüne alınarak çalışma kapsamında kullanılan çok boyutlu veri setine öncelikle çok boyutlu, sonrasında tek boyutlu parametre kestirimleri yapılarak ölçek dönüştürme sürecine alınmış ve elde edilen RMSE değerleri karşılaştırılarak anlamlı bir fark olup olmadığı sorusuna ilişkisiz örneklem t testi ile yanıt aranmıştır. Sonrasında çok boyutlu ve tek boyutlu analizler için ayrı ayrı olacak şekilde çalışmada yer alan diğer koşullara (örneklem büyüklüğü, boyutlar arası korelasyon, ortak madde oranı, gruplar arası yetenek dağılımı farkı ve parametre kestirim modeli) ve bunların etkileşimlerine göre elde edilen RMSE değerleri karşılaştırılarak anlamlı farklılık olup olmadığı sorusuna 5 yönlü varyans analizi (ANOVA) testi yapılarak yanıt aranmıştır. Karşılaştırma testleri yapılırken SPSS 23 yazılımı kullanılmıştır.



## Bölüm 4

### Bulgular, Yorumlar ve Tartışma

Çalışma kapsamında üretilmiş olan çok boyutlu test formlarına sırasıyla 3PLM ve 2PLM kullanılarak araştırma koşulları çerçevesinde sırasıyla çok boyutlu ve tek boyutlu ölçek dönüştürme yapılmış, elde edilen bulgular sırasıyla ayrı başlıklar halinde yorumlanmıştır.

#### Alt Problem 1'e İlişkin Bulgular ve Yorumlar

*“Çok boyutlu test verileri eşdeğer olmayan gruplar ortak test deseninde çok boyutlu ölçek dönüştürme yöntemleriyle eşitlendiğinde madde ve yetenek parametrelerinden kestirilen eşitleme hatası (RMSE) değerleri çalışmada ele alınan koşullara göre nasıl değişmektedir?”*

İlgili alt problemin çözümüne yönelik olarak, testin her iki formundan elde edilen verilere 3PLM ve 2PLM üzerinden yapılan çok boyutlu ölçek dönüştürme işlemi sonucunda elde edilen madde ve yetenek parametrelerine ilişkin RMSE değerleri elde edilmiştir. Elde edilen RMSE değerleri gruplar arasındaki yetenek dağılımı farkı düşük ve yüksek olan durumlar için ayrı ayrı olacak şekilde tablolarla ve şekillerle gösterilmiştir.

**3PLM Kullanılarak Yapılan Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri.** Bu kısımda 3PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri, çalışmada ele alınan koşullara göre gruplar arası yetenek dağılımı ortalamaları arasındaki fark düşük (0.05) ve yüksek (0.5) olduğu durumlar için ayrı tablolar ve şekiller halinde gösterilmiş ve yorumlanmıştır. İlk olarak, gruplar arası yetenek dağılımı farkı düşük olan durumda 3PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri tablo 2'de sunulmuştur.

**Tablo 2**

*Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

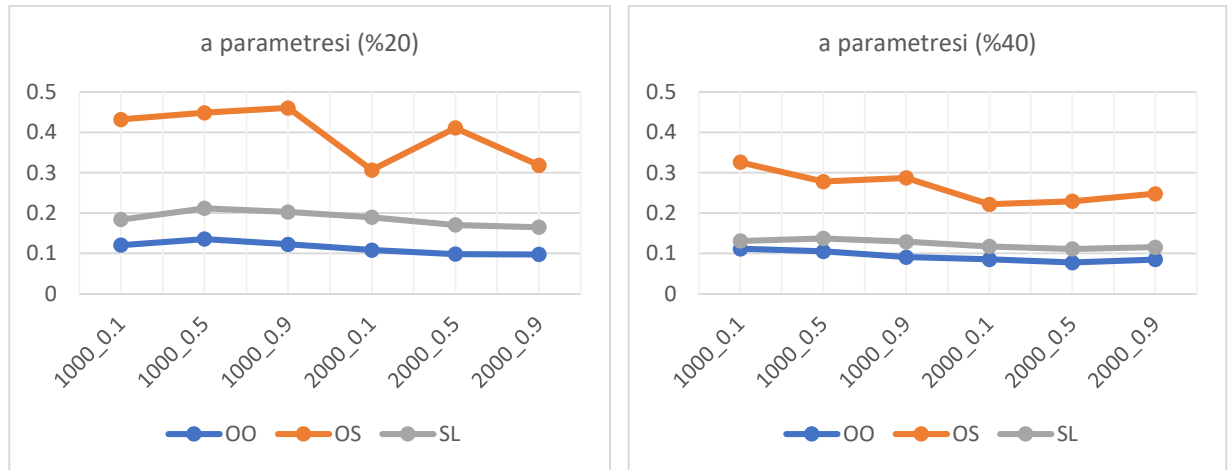
Boyutlar Arası Korelasyon	Örneklem Büyüklüğü	Ortak Madde Oranı	RMSE Değeri								
			a			b			theta (yetenek)		
			OO	OS	SL	OO	OS	SL	OO	OS	SL
0.1	1000	%20	0.12	0.43	0.18	0.21	0.38	0.26	0.15	0.25	0.21
		%40	0.11	0.33	0.13	0.14	0.23	0.21	0.10	0.17	0.16
	2000	%20	0.11	0.31	0.19	0.15	0.25	0.32	0.10	0.17	0.27
		%40	0.09	0.22	0.12	0.10	0.21	0.26	0.07	0.15	0.20
0.5	1000	%20	0.14	0.45	0.21	0.22	0.38	0.33	0.15	0.25	0.27
		%40	0.11	0.28	0.14	0.13	0.24	0.23	0.09	0.17	0.17
	2000	%20	0.10	0.41	0.17	0.17	0.30	0.25	0.12	0.21	0.20
		%40	0.08	0.23	0.11	0.10	0.17	0.20	0.07	0.13	0.15
0.9	1000	%20	0.12	0.46	0.20	0.21	0.39	0.30	0.14	0.25	0.24
		%40	0.09	0.29	0.13	0.13	0.27	0.23	0.09	0.19	0.18
	2000	%20	0.10	0.32	0.17	0.15	0.28	0.28	0.11	0.19	0.23
		%40	0.08	0.25	0.12	0.10	0.23	0.24	0.07	0.15	0.19

Tablo 2 incelendiğinde, gruplar arası yetenek dağılımı farkı düşük olan durumda 3PLM kullanılarak yapılan analizlerde tüm yöntemler için örneklem büyüklüğü arttıkça elde edilen RMSE değerinin azaldığı görülmektedir. Benzer şekilde, ortak madde oranının artmasının da genel olarak daha düşük RMSE değeri elde edilmesinde etkili olduğu yorumu yapılabilir. Yöntemlere ilişkin elde edilen RMSE değerleri incelendiğinde, bütün parametreler için ortalama-sigma yöntemine ilişkin RMSE değerlerinin diğer yöntemlere oranla genel olarak daha yüksek olduğu görülmektedir. Bu durum Atar ve Yeşiltaş (2017) tarafından elde edilen bulguları desteklemektedir. Bunun yanında en düşük RMSE değerlerinin ortalama-ortalama yönteminden elde edildiği görülmektedir. Bu durum ortalama-ortalama yönteminin ortalama-sigma yöntemine göre daha iyi sonuçlar verdiğini gösteren Ogasawara (2000) ile Gök ve Kelecioğlu'nun (2014) bulgularıyla tutarlılık

gösterirken, Hanson ve Beguin (2002) ile Baker ve Al-Karni (1991) tarafından yapılan çalışmalarda raporlanan karakteristik eğri yöntemlerinin, moment yöntemlerinden daha düşük hata ürettiği bulgusuyla çelişmektedir. Boyutlar arası korelasyon 0.1 (düşük), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %40 (yüksek) iken ortalama-ortalama ve ortalama-sigma yöntemlerine ait RMSE değerlerinin bütün parametreler için en düşük değeri aldığı görülmektedir. Stocking-Lord yöntemi için bütün parametrelere ilişkin en düşük RMSE değerleri boyutlar arası korelasyon 0.5 (orta), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %40 (yüksek) olan koşullarda elde edilmiştir. Şekil 4'te gruplar arası yetenek dağılımı farkı düşük olduğu durumda a, b ve yetenek parametrelerine ilişkin 3PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri gösterilmiştir.

#### Şekil 4

*Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*





Şekil 4 incelendiğinde, örneklem büyüklüğündeki artışın genel olarak tüm yöntemlerden kestirilen RMSE değerlerinde düşüşe neden olduğu görülmektedir. Bununla birlikte boyutlar arası korelasyon 0.1 olduğu koşullarda ortalama-ortalama ve ortalama-sigma yöntemlerinde en düşük değerler elde edilmiştir. Ortak madde yüzdesindeki artış, RMSE değerlerini düşürücü yönde etki etmiştir. Tüm parametreler için en düşük RMSE değerlerinin, genel olarak ortak madde oranının %40 ve örneklem büyüklüğünün 2000 olduğu koşullarda ortalama-ortalama yönteminden elde edildiği görülmektedir. Gruplar arası yetenek dağılımı farkı yüksek olduğunda ölçek dönüştürme yöntemlerinden elde edilen ortalama RMSE değerleri tablo 3'te gösterilmiştir.

**Tablo 3**

*Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

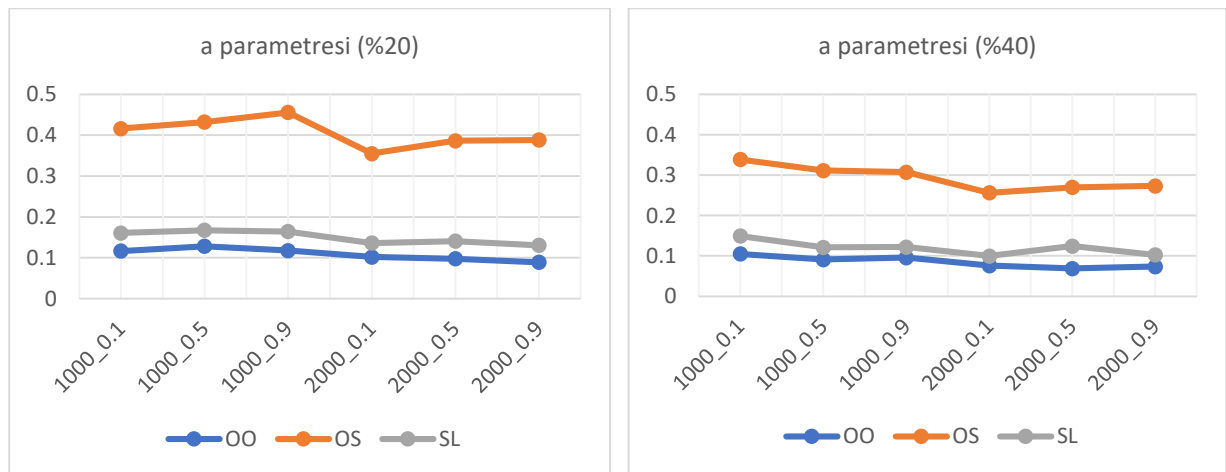
Boyutlar Arası Korelasyon	Örneklem Büyüklüğü	Ortak Madde Oranı	RMSE Değeri								
			a			b			theta (yetenek)		
			OO	OS	SL	OO	OS	SL	OO	OS	SL
0.1	1000	%20	0.12	0.42	0.16	0.35	0.61	0.40	0.29	0.40	0.34
		%40	0.10	0.34	0.15	0.34	0.46	0.43	0.30	0.34	0.38
	2000	%20	0.10	0.36	0.14	0.29	0.51	0.43	0.27	0.35	0.40
		%40	0.08	0.26	0.10	0.27	0.38	0.44	0.27	0.29	0.42
0.5	1000	%20	0.13	0.43	0.17	0.33	0.62	0.43	0.29	0.40	0.38
		%40	0.09	0.31	0.12	0.29	0.46	0.43	0.28	0.33	0.39
	2000	%20	0.10	0.39	0.14	0.33	0.50	0.37	0.28	0.35	0.33
		%40	0.07	0.27	0.12	0.29	0.37	0.39	0.27	0.29	0.35
0.9	1000	%20	0.12	0.46	0.16	0.34	0.63	0.40	0.28	0.40	0.35
		%40	0.10	0.31	0.12	0.30	0.41	0.43	0.27	0.31	0.39
	2000	%20	0.09	0.39	0.13	0.31	0.54	0.42	0.28	0.36	0.38
		%40	0.07	0.27	0.10	0.28	0.38	0.43	0.27	0.29	0.40

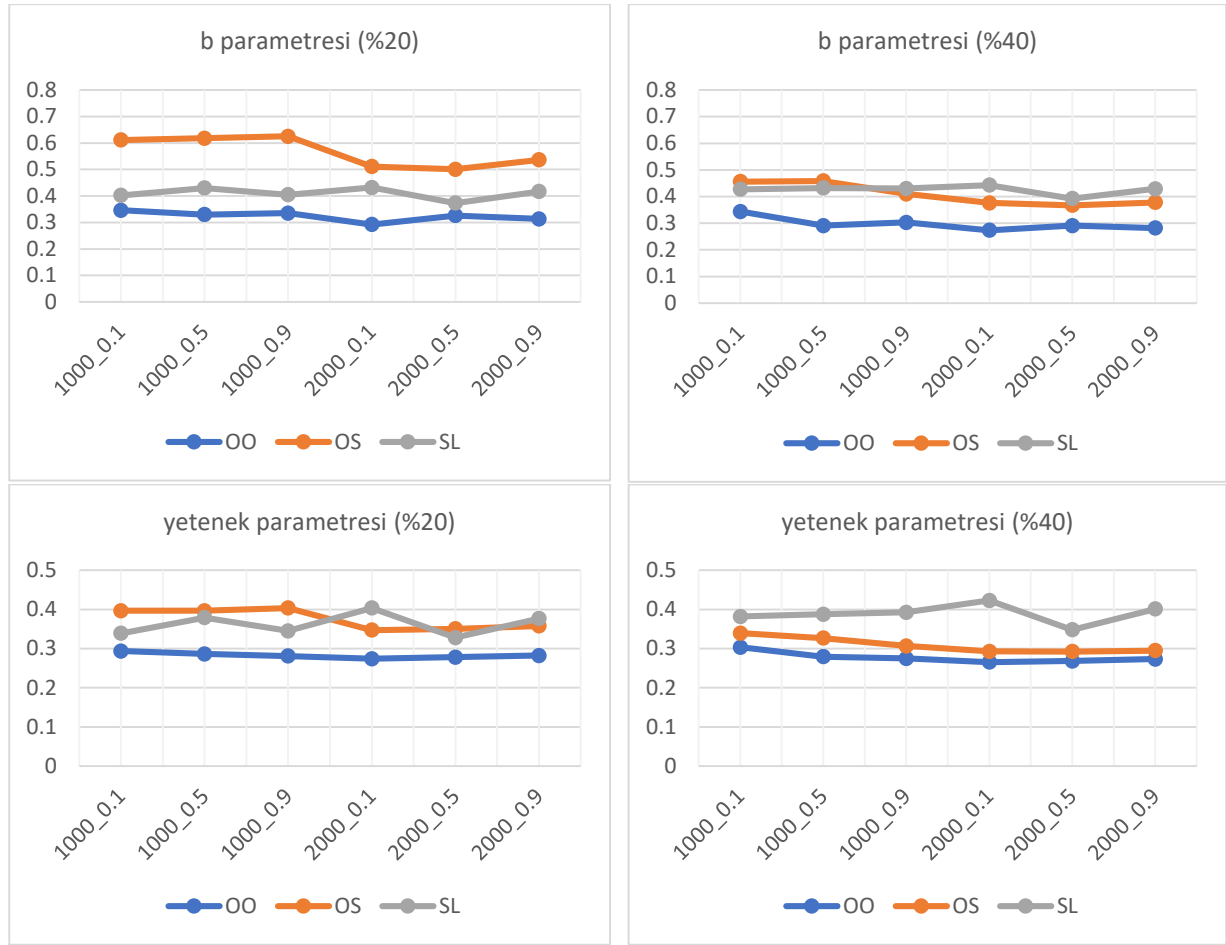
Tablo 3 incelendiğinde, gruplar arası yetenek dağılımı farkı yüksek olduğu durumda 3PLM kullanılarak yapılan analizlerde tüm yöntemler için örneklem büyüklüğü arttıkça genel olarak RMSE değerinin azaldığı görülmektedir. Benzer şekilde, ortak madde oranının artmasının da genel olarak daha düşük RMSE değeri edilmesinde etkili olduğu yorumu yapılabilir, yalnızca b ve yetenek parametrelerine ilişkin Stocking-Lord yönteminden kestirilen RMSE değerleri ortak madde oranı artışından ters yönde etkilenmiştir. Yöntemlere ilişkin elde edilen RMSE değerleri incelendiğinde, ortak madde oranının %40 olduğu ve boyutlar arası korelasyonun yüksek olduğu durumlarda b ve yetenek parametrelerine ait Stocking-Lord yöntemi kullanılarak kestirilen RMSE değerleri en yüksek bulunurken, diğer durumlarda ortalama-sigma yöntemine ilişkin RMSE değerleri diğer yöntemlere oranla daha yüksek olarak kestirilmiştir. Bununla birlikte en düşük RMSE değerlerinin ortalama-ortalama

yönteminden elde edildiği görülmektedir. Bu durum Ogasawara (2000) ile Gök ve Kelecioğlu'nun (2014)'nin bulgularıyla tutarlılık gösterirken, Hanson ve Beguin (2002) ile Baker ve Al-Karni (1991)'nin bulgularıyla örtüşmemektedir. Ortalama-ortalama ve ortalama-sigma yöntemleri için en düşük RMSE değerleri örneklem büyüklüğü 2000 (yüksek), boyutlar arası korelasyon 0.1 (düşük) ve ortak madde oranı %40 (yüksek) olan koşulda elde edilmiştir. Stocking-Lord yöntemi için en düşük RMSE değerleri örneklem büyüklüğü 2000 (yüksek), boyutlar arası korelasyon 0.5 (orta) ve ortak madde oranı %40 (yüksek) olan koşulda elde edilmiştir. Ortak madde oranındaki artış bütün parametrelere ilişkin ortalama-ortalama ve ortalama-sigma yöntemleri kullanılarak kestirilen RMSE değerlerini düşürücü yönde etki ederken, b ve yetenek parametrelerine ilişkin Stocking-Lord yöntemi kullanılarak kestirilen RMSE değerlerinde artışa neden olmuştur. Tablo 2 ile kıyaslandığında, gruplar arası yetenek dağılımı farkı yüksek olduğunda özellikle b ve yetenek parametrelerinden kestirilen RMSE değerlerinde artış görülmektedir, a parametrelerine ait RMSE değerleri birbirine yakın olarak bulunmuştur. Şekil 5'te, gruplar arası yetenek dağılımı yüksek olan durumda a, b ve yetenek parametrelerine ilişkin 3PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri gösterilmiştir.

### Şekil 5

*Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*





Şekil 5 incelendiğinde, örneklem büyüklüğündeki artış a ve b parametrelerinden kestirilen RMSE değerlerini düşürücü yönde etki etmiştir. Boyutlar arası korelasyonun 0.1 olduğu durumlarda ortalama-ortalama ve ortalama-sigma yöntemlerine ait RMSE değerleri düşüş göstermiştir. Stocking-Lord yöntemi ile kestirilen yetenek parametrelerinden kestirilen RMSE değerleri incelendiğinde, örneklem büyüklüğü 2000 ve boyutlar arası korelasyon 0.1 olduğu durumda en büyük değerleri aldığı görülmektedir. Bunun dışında tüm parametrelere ilişkin en düşük RMSE değerlerinin, genel olarak örneklem büyüklüğünün 2000 (yüksek), boyutlar arası korelasyonun 0.1 (düşük) ve ortak madde oranının %40 (yüksek) olduğu koşullarda ortalama-ortalama yönteminden elde edildiği görülmektedir.

**2PLM Kullanılarak Yapılan Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerlerine İlişkin Bulgular.** Bu bölümde 2PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri, gruplar arası yetenek dağılımı arasındaki fark düşük (0.05) ve yüksek (0.5) olduğu durumlar için ayrı tablolar ve

şekiller halinde gösterilmiş ve yorumlanmıştır. İlk olarak, gruplar arası yetenek dağılımı farkı düşük olan durumda 2PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri tablo 4'te sunulmuştur.

**Tablo 4**

*Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

Boyutlar Arası Korelasyon	Örneklem Büyüklüğü	Ortak Madde Oranı	RMSE Değeri								
			a			b			theta (yetenek)		
			OO	OS	SL	OO	OS	SL	OO	OS	SL
0.1	1000	%20	0.07	0.32	0.09	0.25	0.37	0.17	0.18	0.27	0.12
		%40	0.05	0.29	0.06	0.17	0.30	0.14	0.12	0.22	0.11
	2000	%20	0.06	0.23	0.07	0.14	0.22	0.13	0.10	0.16	0.10
		%40	0.07	0.20	0.17	0.15	0.25	0.34	0.11	0.18	0.19
0.5	1000	%20	0.08	0.39	0.10	0.28	0.40	0.18	0.19	0.28	0.13
		%40	0.09	0.33	0.19	0.22	0.33	0.35	0.16	0.23	0.20
	2000	%20	0.05	0.25	0.07	0.19	0.27	0.15	0.13	0.20	0.11
		%40	0.08	0.22	0.17	0.17	0.27	0.34	0.12	0.19	0.20
0.9	1000	%20	0.08	0.45	0.10	0.32	0.38	0.19	0.22	0.28	0.15
		%40	0.09	0.29	0.19	0.22	0.40	0.35	0.16	0.26	0.21
	2000	%20	0.05	0.23	0.06	0.16	0.25	0.14	0.11	0.18	0.10
		%40	0.08	0.23	0.17	0.16	0.22	0.34	0.10	0.16	0.20

Tablo 4 incelendiğinde, 3PLM'de olduğu gibi, 2PLM kullanılarak yapılan analizlerde de tüm ölçek dönüştürme yöntemleri için örneklem büyüklüğü arttıkça RMSE değerinin azaldığı görülmektedir. Bunun yanında ortak madde oranının artmasının ortalama-ortalama ve ortalama-sigma yöntemlerinden elde edilen RMSE değerlerini azaltıcı yönde, Stocking-Lord yönteminden elde edilen RMSE değerlerini ise artırıcı yönde etki ettiği yorumu yapılabilir. Yöntemlere ilişkin elde edilen RMSE değerleri incelendiğinde, örneklem büyüklüğü 2000 ve ortak madde oranı %40 olan durumlarda b ve yetenek parametrelerine



ilişkin Stocking-Lord yönteminden kestirilen RMSE değerleri diğer yöntemlere oranla daha yüksek olarak kestirilirken, diğer durumlarda ise ortalama-sigma yöntemine ilişkin değerler daha yüksektir. Bütün yöntemler için en düşük RMSE değerleri genel olarak boyutlar arası korelasyon 0.1 (düşük), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %20 (düşük) olan durumda elde edilmiştir.

Gruplar arası yetenek dağılımı farkı düşük olan durumlar için parametre kestirim modellerine göre RMSE değerleri karşılaştırıldığında, a parametresine ilişkin RMSE değerleri tüm yöntemler için 2PLM'de daha düşük kestirilmiştir. b parametresi ile yetenek parametresine ilişkin RMSE değerleri incelendiğinde ortalama-sigma ve Stocking-Lord yöntemine ait değerler genel olarak ortak madde oranı %20 olduğunda 2PLM'de, %40 olduğunda 3PLM'de daha düşük bulunmuş, ortalama-ortalama yöntemine ait değerler ise genel olarak 3PLM'de daha düşük olarak kestirilmiştir. Bu bağlamda 2PLM kullanıldığında genel olarak daha düşük RMSE değerleri elde edildiği yorumu yapılabilir. Bu bulgu, Kaskowitz ve De Ayala (2001) tarafından 3PLM'nin daha yüksek hatayla a ve b parametrelerine ilişkin kestirim yaptığı bulgusuyla kısmen örtüşmektedir. Parametre kestirim modeli olarak 2PLM'nin kullanıldığı durumlarda, büyük örneklerde daha kararlı madde parametre kestirimleri elde edilmesi bu durumu açıklayabilir (Bökeoğlu ve diğ., 2022). Şekil 6'da gruplar arası yetenek dağılımı farkı düşük olduğu durumda a, b ve yetenek parametrelerine ilişkin 2PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri gösterilmiştir.

## Şekil 6

Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri



Şekil 6 incelendiğinde, örneklem büyüklüğündeki artışın ve boyutlar arasındaki korelasyonun düşüşünün tüm parametreler için yöntemlerden kestirilen RMSE değerlerinde

keskin bir düşüşe neden olduğu görülmektedir. Çok boyutlu veri yapısına çok boyutlu ölçek dönüştürme yöntemleri uygulandığından, bu durumun olmasının beklendiği yorumu yapılabilir. Tüm parametrelere ilişkin RMSE değerleri incelendiğinde, ortalama-sigma yönteminden elde edilen değerler diğer yöntemlere oranla daha yüksek olup, ortalama-ortalama yönteminden elde edilen değerler daha düşük bulunmuştur. Örneklem büyüklüğü arttıkça tüm parametreler için tüm yöntemlere ait RMSE değerleri azalma eğilimindedir. Gruplar arası yetenek dağılımı farkı yüksek olduğunda 2PLM kullanılarak yapılan ölçek dönüştürme yöntemlerinden elde edilen ortalama RMSE değerleri tablo 5'te gösterilmiştir.

**Tablo 5**

*Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

Boyutlar Arası Korelasyon	Örneklem Büyüklüğü	Ortak Madde Oranı	RMSE Değeri									
			a			b			theta (yetenek)			
			OO	OS	SL	OO	OS	SL	OO	OS	SL	
0.1	1000	%20	0.10	0.38	0.10	0.47	0.50	0.40	0.35	0.37	0.42	
		%40	0.10	0.45	0.09	0.49	0.37	0.42	0.38	0.41	0.47	
	2000	%20	0.10	0.32	0.08	0.43	0.40	0.40	0.32	0.33	0.43	
		%40	0.15	0.38	0.22	0.50	0.30	0.36	0.35	0.38	0.41	
	0.5	1000	%20	0.11	0.49	0.11	0.50	0.52	0.40	0.36	0.38	0.41
			%40	0.16	0.47	0.22	0.55	0.36	0.37	0.39	0.39	0.41
2000		%20	0.10	0.34	0.08	0.46	0.42	0.40	0.34	0.33	0.43	
		%40	0.14	0.39	0.22	0.50	0.31	0.36	0.36	0.38	0.41	
0.9		1000	%20	0.12	0.54	0.11	0.56	0.50	0.39	0.40	0.36	0.42
			%40	0.16	0.45	0.23	0.51	0.34	0.37	0.36	0.38	0.41
	2000	%20	0.10	0.35	0.08	0.45	0.42	0.40	0.34	0.33	0.43	
		%40	0.14	0.44	0.22	0.52	0.32	0.36	0.38	0.40	0.42	

Tablo 5 incelendiğinde, tüm ölçek dönüştürme yöntemleri için örneklem büyüklüğü arttıkça RMSE değerinin genel olarak azaldığı görülmektedir. Bunun yanında ortak madde

oranının artmasının ortalama-ortalama yönteminden elde edilen RMSE değerlerini artırıcı yönde, b ve yetenek parametrelerine ilişkin Stocking-Lord yönteminden elde edilen RMSE değerlerini ise azaltıcı yönde etki ettiği yorumu yapılabilir. Yöntemlere ilişkin elde edilen RMSE değerleri incelendiğinde, a parametrelerinin kestiriminde ortalama-sigma yönteminden, b parametrelerinin kestiriminde ortalama-ortalama yönteminden, yetenek parametrelerinden kestiriminde ise Stocking-Lord yönteminden elde edilen hataların daha yüksek olduğu görülmektedir. Boyutlar arası korelasyon 0.1 (düşük), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %20 (düşük) iken genel olarak bütün parametreler için tüm yöntemlerden kestirilen ortalama RMSE değerlerinin için en düşük değerleri aldığı yorumu yapılabilir.

Gruplar arası yetenek dağılımı farkı yüksek olan durumlar için parametre kestirim modellerine göre RMSE değerleri karşılaştırıldığında, ortalama-ortalama yönteminin bütün parametreler için genel olarak 3PLM'de daha düşük değerler ürettiği görülmektedir. Ortalama-sigma ve Stocking-Lord yöntemlerinin, b parametresi için 2PLM'de, yetenek parametresi için genel olarak 3PLM'de daha düşük değerler ürettiği yorumu yapılabilir. Ortalama-sigma yöntemi kullanıldığında yetenek parametresinden kestirilen RMSE değerleri boyutlar arası korelasyonun %20 olduğu durumlarda 2PLM'de daha düşük kestirilmiştir. Bu bağlamda 3PLM ya da 2PLM'den hangisinin kullanılmasının avantaj sağlayacağı hakkında net bir kanıya varmak güç olmakla birlikte ortalama-ortalama yöntemi için 3PLM üzerinden yapılan analizlerde daha düşük hatalar elde edileceği yorumu yapılabilir. Şekil 7'de gruplar arası yetenek dağılımı farkı yüksek olduğu durumda a, b ve yetenek parametrelerine ilişkin 2PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri gösterilmiştir.

## Şekil 7

Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Çok Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri



Şekil 7 incelendiğinde, boyutlar arası korelasyonun artmasının genel olarak RMSE değerlerini artırıcı yönde etkide bulunmuştur. Boyutlar arası korelasyonun 0.9 olduğu koşullarda tüm parametreler ve tüm yöntemler için en yüksek RMSE değerleri elde edilmiştir. Çok boyutlu veride boyutlar arası korelasyonun artmasının testin tek boyutluluğunu güçlendirdiği düşünüldüğünde, bunun beklenen bir durum olduğu yorumu yapılabilir. Örneklem büyüklüğünün artması RMSE değerlerini düşürücü etkiye sahiptir.

### **Alt Problem 2'ye İlişkin Bulgular ve Yorumlar**

*“Çok boyutlu test verileri eşdeğer olmayan gruplar ortak test deseninde tek boyutlu ölçek dönüştürme yöntemleriyle eşitlendiğinde madde ve yetenek parametrelerinden kestirilen eşitleme hatası (RMSE) değerleri çalışmada ele alınan koşullara göre nasıl değişmektedir?”*

İlgili alt problemin çözümüne yönelik olarak, testin her iki formundan elde edilen verilere 3PLM ve 2PLM üzerinden yapılan tek boyutlu ölçek dönüştürme işlemi sonucunda elde edilen madde ve yetenek parametrelerine ilişkin RMSE değerleri elde edilmiştir. Elde edilen RMSE değerleri gruplar arasındaki yetenek dağılımı farkı düşük ve yüksek olan durumlar için ayrı ayrı olacak şekilde tablolarla ve şekillerle gösterilmiştir.

**3PLM Kullanılarak Yapılan Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri.** Bu kısımda 3PLM kullanılarak yapılan tek boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri, gruplar arası yetenek dağılımı arasındaki fark düşük ve yüksek olduğu durumlar için ayrı tablolar ve şekiller halinde sunulmuş ve yorumlanmıştır. İlk olarak, gruplar arası yetenek dağılımı farkı düşük olan durumda 3PLM kullanılarak yapılan tek boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri tablo 6'da sunulmuştur.

**Tablo 6**

*Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

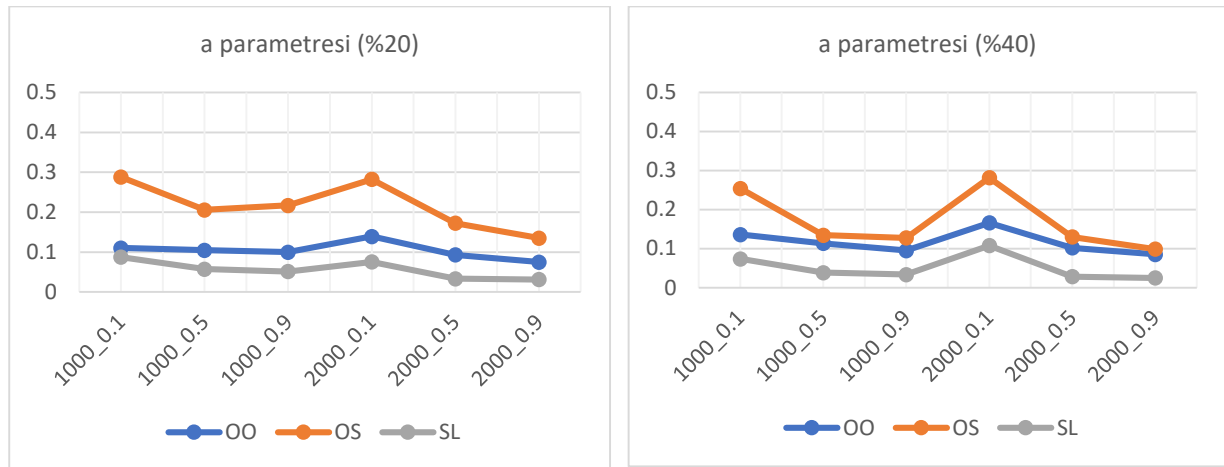
Boyutlar Arası Korelasyon	Örnekleme Büyüklüğü	Ortak Madde Oranı	RMSE Değeri								
			a			b			theta (yetenek)		
			OO	OS	SL	OO	OS	SL	OO	OS	SL
0.1	1000	%20	0.11	0.29	0.09	0.53	0.76	0.26	0.38	0.40	0.15
		%40	0.14	0.25	0.07	0.54	0.75	0.15	0.32	0.43	0.10
	2000	%20	0.14	0.28	0.08	0.55	0.82	0.23	0.39	0.41	0.13
		%40	0.17	0.28	0.11	0.51	0.75	0.15	0.30	0.45	0.12
0.5	1000	%20	0.10	0.21	0.06	0.18	0.25	0.09	0.15	0.21	0.08
		%40	0.11	0.13	0.04	0.14	0.16	0.06	0.09	0.12	0.05
	2000	%20	0.09	0.17	0.03	0.16	0.20	0.05	0.14	0.18	0.05
		%40	0.10	0.13	0.03	0.13	0.16	0.04	0.09	0.12	0.04
0.9	1000	%20	0.10	0.22	0.05	0.15	0.19	0.07	0.13	0.18	0.07
		%40	0.09	0.13	0.03	0.11	0.12	0.04	0.07	0.10	0.05
	2000	%20	0.08	0.14	0.03	0.10	0.12	0.04	0.09	0.12	0.04
		%40	0.09	0.10	0.02	0.09	0.10	0.03	0.06	0.08	0.03

Tablo 6 incelendiğinde, boyutlar arası korelasyon arttıkça tüm parametreler için tüm yöntemlerden elde edilen RMSE değerleri düşme eğilimindedir. Bu durum, veri yapısının çok boyutlu olması ve boyutlar arası korelasyonun artmasının veriyi tek boyutlu analiz için daha elverişli hale getirmesi olarak açıklanabilir. Boyutlar arası korelasyon 0.5 ve 0.9 olan durumlarda tüm yöntemler için örneklem büyüklüğü arttıkça RMSE değerlerinin azaldığı, boyutlar arası korelasyonun 0.1 olduğu durumda ise sadece Stocking-Lord yöntemi için örneklem büyüklüğündeki artışın RMSE değerlerini düşürücü yönde etki ettiği yorumu yapılabilir. Ortak madde oranının artması yalnızca ortalama-ortalama yönteminden elde edilen a parametrelerine ait RMSE değerlerini artırıcı yönde, diğer bütün durumlarda RMSE değerlerini azaltıcı yönde etki etmiştir. Yöntemlere ilişkin elde edilen RMSE değerleri incelendiğinde, bütün parametreler için ortalama-sigma yöntemine ilişkin RMSE

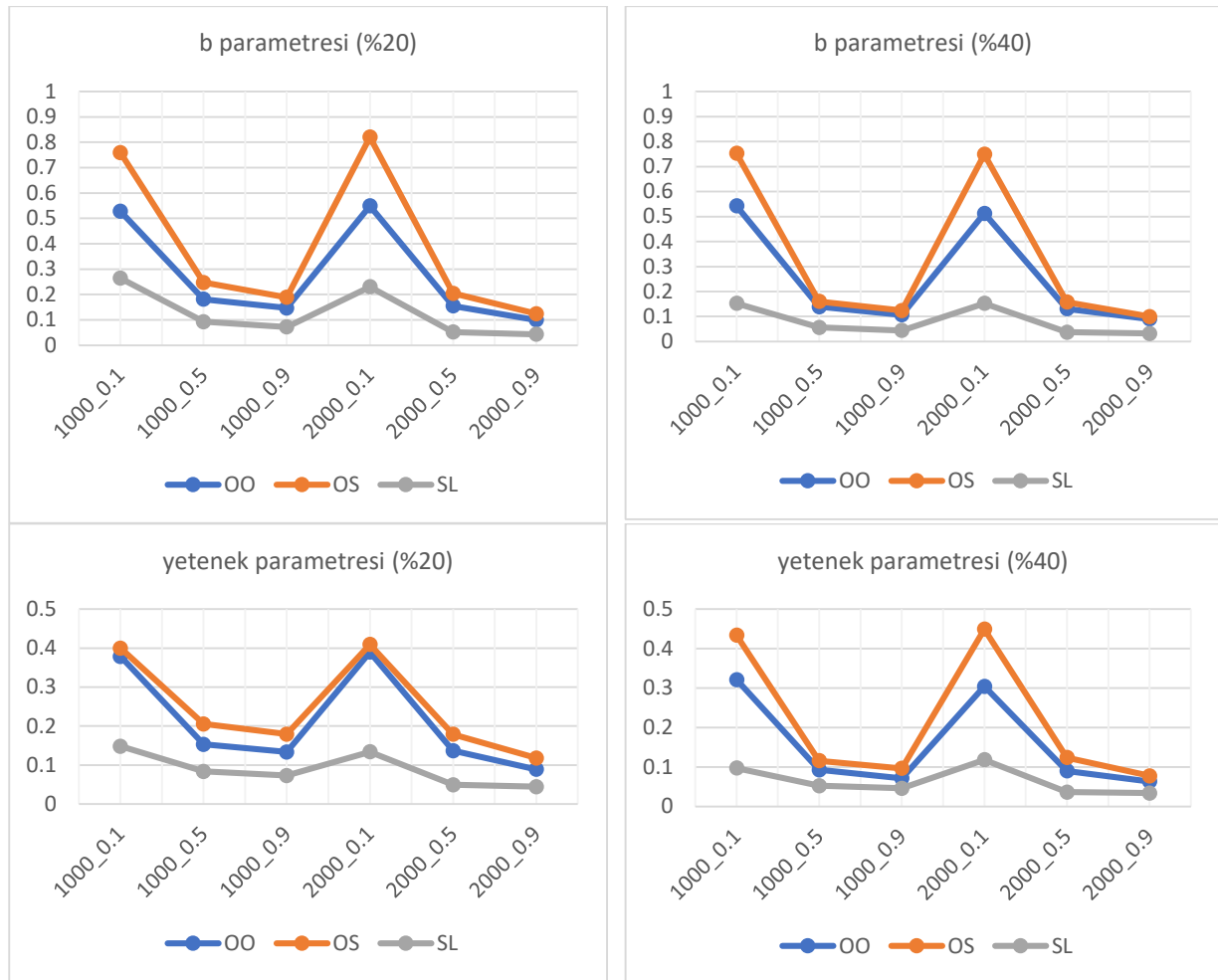
değerlerinin diğer yöntemlere oranla daha yüksek olduğu görülmektedir. Bu durum Atar ve Yeşiltaş (2017) tarafından elde edilen bulguları desteklemektedir. Bunun yanında en düşük RMSE değerlerinin Stocking-Lord yönteminden elde edildiği görülmektedir. Bu durum literatürde bulunan sonuçlarla uyumludur (Baker ve Al-Karni, 1991; Hanson ve Beguin, 2002; Gök ve Kelecioğlu, 2014). Boyutlar arası korelasyon 0.9 (yüksek), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %40 (yüksek) iken tüm yöntemlerden elde edilen RMSE değerlerinin bütün parametreler için en düşük değeri aldığı görülmektedir. Şekil 8'de gruplar arası yetenek dağılımı farkı düşük olduğu durumda a, b ve yetenek parametrelerine ilişkin 3PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri gösterilmiştir.

### Şekil 8

*Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*







Şekil 8 incelendiğinde, boyutlar arası korelasyon arttıkça bütün parametreler için RMSE değerlerinin azaldığı, boyutlar arası korelasyonun 0.1 olduğu koşullarda tüm parametreler ve tüm yöntemler için keskin yükselişler görülmektedir. Çok boyutlu veri yapısında boyutlar arası korelasyon düşükken testin tek boyutluluk derecesinin azalması ve bunun sonucunda uygulanan tek boyutlu ölçek dönüştürme sonucu elde edilen hataların artması beklenen bir durumdur. Örneklem büyüklüğünün artması ise RMSE değerlerini düşürücü yönde etkide bulunmuştur. Bütün parametreler için yüksek örneklem büyüklüklerinde daha düşük RMSE değerleri elde edilmiştir. Bununla birlikte ortak madde oranındaki artışın genel olarak tüm parametrelerden kestirilen RMSE değerlerini düşürücü yönde etki ettiği yorumu yapılabilir. Bütün parametreler için en yüksek RMSE değerleri ortalama-sigma yönteminden, en düşük RMSE değerleri Stocking-Lord yönteminden kestirilmiştir. Tablo 7’de gruplar arası yetenek dağılımı farkı yüksek olduğu durumda a, b ve

yetenek parametrelerine ilişkin 3PLM kullanılarak yapılan çok boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri gösterilmiştir.

**Tablo 7**

*Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

Boyutlar Arası Korelasyon	Örneklem Büyüklüğü	Ortak Madde Oranı	RMSE Değeri								
			a			b			theta (yetenek)		
			OO	OS	SL	OO	OS	SL	OO	OS	SL
0.1	1000	%20	0.11	0.27	0.11	1.10	1.28	0.81	1.02	1.11	0.78
		%40	0.09	0.27	0.10	1.26	1.46	0.84	1.16	1.19	0.82
	2000	%20	0.13	0.23	0.08	1.09	1.18	0.79	1.02	1.08	0.79
		%40	0.08	0.18	0.05	1.11	1.19	0.75	1.00	1.13	0.73
0.5	1000	%20	0.09	0.18	0.05	0.58	0.62	0.60	0.57	0.60	0.59
		%40	0.11	0.13	0.05	0.63	0.63	0.61	0.60	0.63	0.59
	2000	%20	0.08	0.16	0.04	0.57	0.59	0.58	0.56	0.59	0.57
		%40	0.10	0.12	0.05	0.62	0.62	0.59	0.60	0.63	0.57
0.9	1000	%20	0.10	0.18	0.05	0.51	0.54	0.52	0.52	0.54	0.53
		%40	0.10	0.12	0.04	0.53	0.54	0.52	0.53	0.55	0.53
	2000	%20	0.07	0.13	0.03	0.53	0.54	0.53	0.53	0.54	0.52
		%40	0.08	0.09	0.03	0.54	0.54	0.53	0.53	0.54	0.52

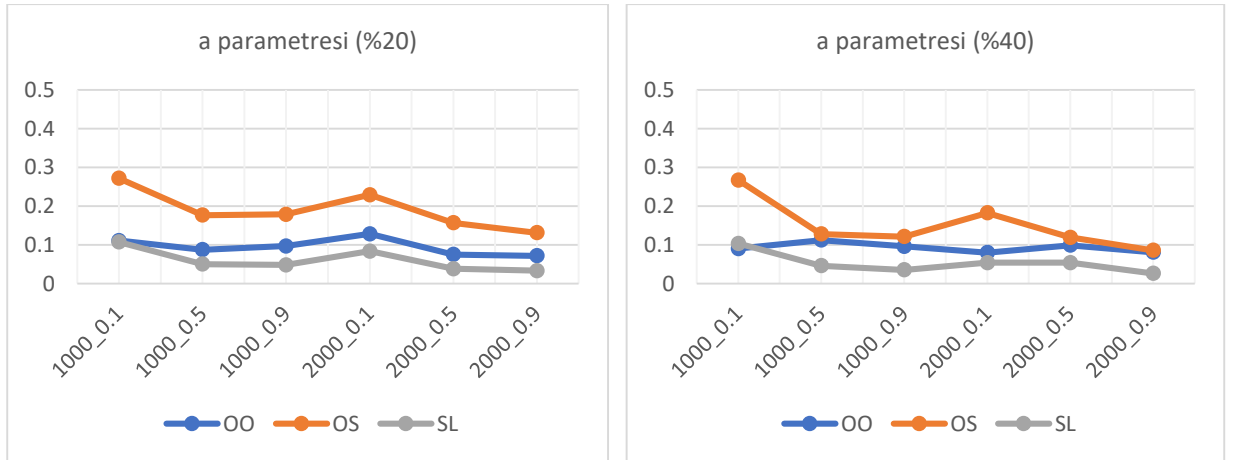
Tablo 7 incelendiğinde, tablo 6'ya kıyasla gruplar arası yetenek dağılımı ortalamaları arasındaki fark yüksek olduğunda tüm ölçek dönüştürme yöntemleri için b ve yetenek parametrelerinden elde edilen RMSE değerlerinin arttığı görülmektedir. Bunun sebebi olarak, gruplar arasındaki yetenek dağılımının farklılaşmasının yetenek parametrelerinin kestiriminde yüksek hataya neden olması gösterilebilir. Örneklem büyüklüğündeki artış, genel olarak RMSE değerlerini düşürücü yönde etkide bulunmuştur. Ortak madde oranının artmasının a parametrelerinden elde edilen RMSE değerlerini azaltıcı yönde, b ve yetenek parametrelerinden elde edilen RMSE değerlerini artırıcı yönde etkide bulunduğu yorumu yapılabilir. Yöntemlere ilişkin elde edilen RMSE değerleri incelendiğinde, bütün parametreler için ortalama-sigma yöntemine ilişkin RMSE değerlerinin diğer yöntemlere

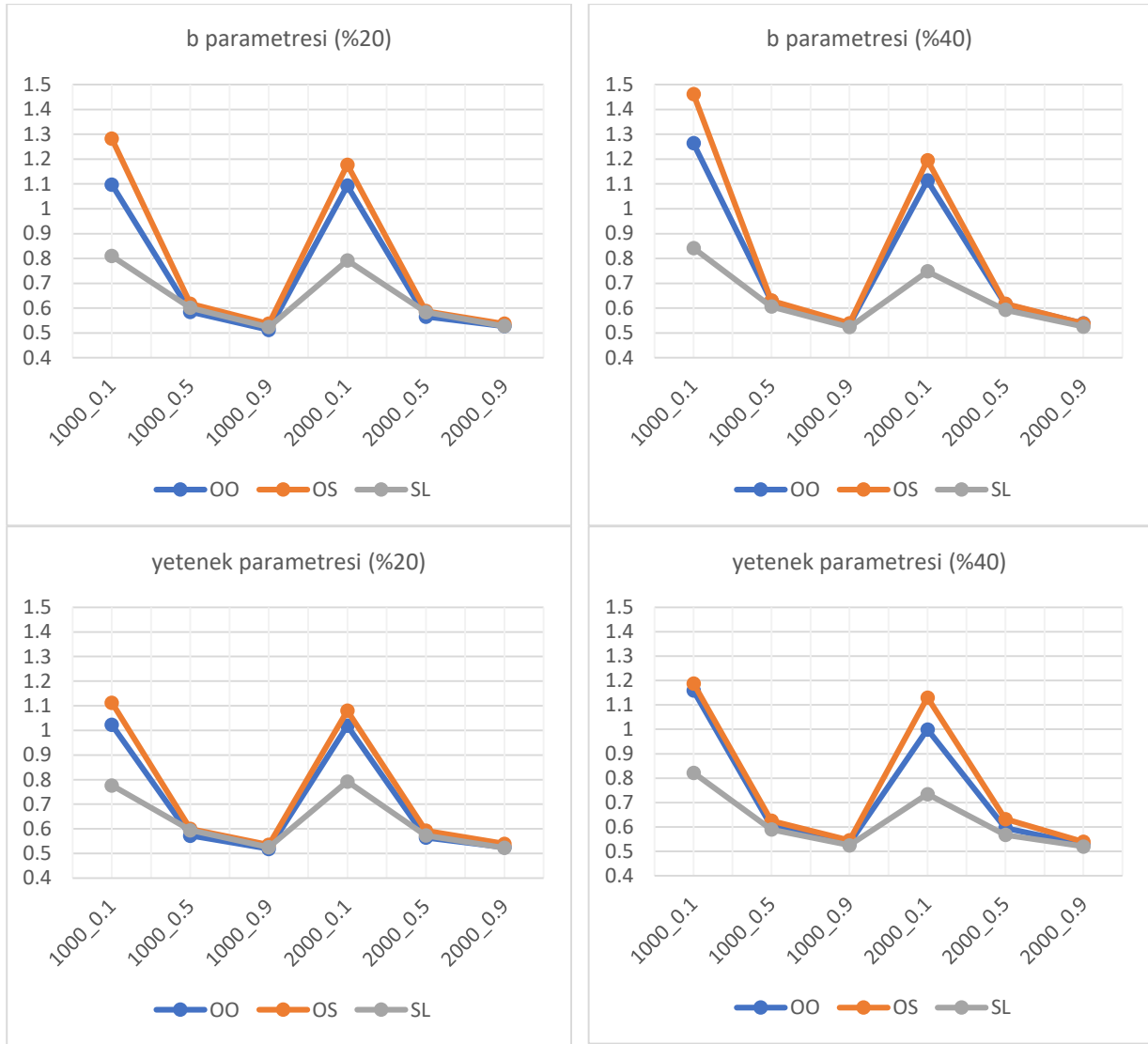
oranla daha yüksek olduğu, en düşük RMSE değerlerinin Stocking-Lord yönteminden elde edildiği görülmektedir. Bu durum Hanson ve Beguin (2002) ile Baker ve Al-Karni (1991)'nin bulgularıyla uyumludur. Boyutlar arası korelasyon 0.9 (yüksek), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %40 (yüksek) iken tüm yöntemlere ilişkin RMSE değerlerinin bütün parametreler için en düşük değeri aldığı görülmektedir.

Şekil 9'da gruplar arası yetenek dağılımı farkı yüksek olduğunda a, b ve yetenek parametrelerine ilişkin 3PLM kullanılarak yapılan tek boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri her bir parametre için ayrı ayrı grafikler halinde gösterilmiştir.

### Şekil 9

*Gruplar Arası Yetenek Dağılımı Farkı Yüksek Olduğunda (0.5) 3PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*





Şekil 9 incelendiğinde, boyutlar arası korelasyon arttıkça genel olarak bütün parametreler için RMSE değerlerinin azaldığı görülmektedir. Bununla birlikte boyutlar arası korelasyonun 0.1 olduğu koşullarda tüm parametreler ve tüm yöntemler için en yüksek değerler elde edilmiştir. Daha önce açıklandığı üzere bu beklenen bir durumdur. Örneklem büyüklüğünün artması ise genel olarak tüm parametrelerden kestirilen RMSE değerlerini düşürücü yönde etkide bulunmuş ve yüksek örneklem büyüklüklerinde daha düşük RMSE değerleri elde edilmiştir. Ortak madde oranının artmasının da bütün parametreler için RMSE değerlerini düşürücü yönde etki ettiği yorumu yapılabilir.

## 2PLM Kullanılarak Yapılan Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen

**RMSE Değerlerine İlişkin Bulgular.** Bu kısımda 2PLM kullanılarak yapılan tek boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri, gruplar arası yetenek dağılımı arasındaki fark düşük ve yüksek olduğu durumlar için ayrı tablolar ve şekiller halinde sunulmuş ve yorumlanmıştır. İlk olarak gruplar arası yetenek farkı düşük iken 2PLM üzerinden yapılan tek boyutlu ölçek dönüştürme işlemi sonucunda elde edilen madde ve yetenek parametrelerinden kestirilen RMSE değerleri tablo 8'de verilmiştir.

**Tablo 8**

*Gruplar Arası Yetenek Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

Boyutlar Arası Korelasyon	Örneklem Büyüklüğü	Ortak Madde Oranı	RMSE Değeri									
			a			b			theta (yetenek)			
			OO	OS	SL	OO	OS	SL	OO	OS	SL	
0.1	1000	%20	0.05	0.25	0.06	0.58	1.05	0.31	0.44	0.60	0.18	
		%40	0.06	0.17	0.06	0.70	1.04	0.29	0.39	0.59	0.15	
	2000	%20	0.04	0.31	0.07	0.79	1.14	0.52	0.64	0.64	0.36	
		%40	0.04	0.16	0.03	0.59	0.92	0.16	0.35	0.50	0.10	
	0.5	1000	%20	0.06	0.16	0.04	0.15	0.21	0.08	0.12	0.20	0.07
			%40	0.06	0.10	0.02	0.13	0.18	0.05	0.08	0.16	0.05
2000		%20	0.04	0.11	0.02	0.10	0.15	0.04	0.08	0.16	0.04	
		%40	0.04	0.08	0.01	0.10	0.15	0.03	0.06	0.14	0.03	
0.9		1000	%20	0.06	0.15	0.04	0.11	0.16	0.06	0.10	0.18	0.07
			%40	0.06	0.10	0.02	0.10	0.13	0.04	0.07	0.12	0.04
	2000	%20	0.04	0.09	0.02	0.07	0.10	0.04	0.06	0.11	0.04	
		%40	0.04	0.07	0.01	0.06	0.08	0.03	0.04	0.08	0.03	

Tablo 8 incelendiğinde, boyutlar arası korelasyon 0.1 ve ortak madde oranı %20 olan durumda örneklem büyüklüğünün artması tüm yöntemlerde RMSE değerlerini artırıcı yönde etki ederken, bunun dışındaki durumlarda örneklem büyüklüğünün artmasıyla daha düşük RMSE değerleri elde edilmiştir. Bunun yanında ortak madde oranının artması RMSE

değerlerini düşürücü yönde etkide bulunmuştur. Yöntemlere ilişkin elde edilen RMSE değerleri incelendiğinde, ortalama-sigma yöntemine ilişkin değerlerin diğer yöntemlere oranla daha yüksek olduğu, tüm parametreler için en düşük değerlerin Stocking-Lord yöntemine ait olduğu görülmektedir. Bu durum alan yazında bulunan karakteristik eğri yöntemlerinin, moment yöntemlerinden daha düşük hata ürettiği bulgusuyla örtüşmektedir (Baker ve Al-Karni, 1991; Hanson ve Beguin, 2002). Boyutlar arası korelasyon 0.9 (düşük), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %40 (düşük) iken bütün yöntemlere ait RMSE değerlerinin bütün parametreler için en düşük değeri aldığı görülmektedir.

Gruplar arası yetenek dağılımı farkı düşük olan durumlar için parametre kestirim modellerine göre RMSE değerleri karşılaştırıldığında, a parametresine ilişkin RMSE değerlerinin tüm yöntemler için 2PLM'de daha düşük kestirildiği görülmektedir. b ve yetenek parametrelerine ilişkin RMSE değerleri incelendiğinde, boyutlar arası korelasyonun 0.1 olduğu durumlarda 3PLM'de, boyutlar arası korelasyonun orta (0.5) ve yüksek (0.5) olduğu durumlarda 2PLM'de daha düşük değerler elde edildiği görülmektedir. Bu bağlamda 2PLM kullanıldığında genel olarak daha düşük RMSE değerleri elde edildiği yorumu yapılabilir. Bu bulgu, 3PLM'nin daha yüksek hatayla a ve b parametrelerine ilişkin kestirim yaptığı bulgusuyla örtüşmektedir (Kaskowitz ve De Ayala, 2001). Şekil 10'da gruplar arası yetenek dağılımı farkı düşük olduğunda a, b ve yetenek parametrelerine ilişkin 2PLM kullanılarak yapılan tek boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri her bir parametre için ayrı grafikler halinde gösterilmiştir.

## Şekil 10

Gruplar Arası Yetenek Dağılımı Farkı Düşük Olduğunda (0.05) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri



Şekil 10 incelendiğinde, boyutlar arası korelasyonun 0.1 olduğu koşullarda tüm parametreler ve tüm yöntemlerden kestirilen RMSE değerlerinde keskin yükselişler görülmektedir. Boyutlar arası korelasyondaki artış genel olarak bütün parametreler için RMSE değerlerini düşürücü yönde etki etmiştir. Örneklem büyüklüğünün artması ise genel olarak tüm parametrelerden kestirilen RMSE değerlerini düşürücü yönde etkide bulunmuş ve yüksek örneklem büyüklüklerinde daha düşük RMSE değerleri elde edilmiştir. Ortak madde oranının artmasının da bütün parametreler için RMSE değerlerini düşürücü yönde etki ettiği yorumu yapılabilir. En yüksek RMSE değerleri bütün parametreler için ortalama-sigma yönteminden, en düşük RMSE değerleri ise Stocking-Lord yönteminden kestirilmiştir. Tablo 9'da, gruplar arası yetenek dağılımı farkı yüksek olan durumda 2PLM kullanılarak yapılan tek boyutlu ölçek dönüştürme sonucu elde edilen RMSE değerleri sunulmuştur.

**Tablo 9**

*Gruplar Arası Yetenek Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*

Boyutlar Arası Korelasyon	Örneklem Büyüklüğü	Ortak Madde Oranı	RMSE Değeri									
			a			b			theta (yetenek)			
			OO	OS	SL	OO	OS	SL	OO	OS	SL	
0.1	1000	%20	0.08	0.25	0.11	0.92	1.26	0.69	0.93	1.12	0.79	
		%40	0.08	0.19	0.15	1.36	1.64	0.92	1.27	1.47	0.97	
	2000	%20	0.08	0.25	0.14	1.11	1.43	0.81	1.16	1.25	0.93	
		%40	0.09	0.18	0.12	1.09	1.36	0.72	1.07	1.38	0.86	
	0.5	1000	%20	0.08	0.14	0.07	0.56	0.59	0.54	0.64	0.65	0.62
			%40	0.09	0.12	0.05	0.60	0.62	0.53	0.68	0.73	0.61
2000		%20	0.08	0.11	0.06	0.56	0.57	0.53	0.64	0.65	0.61	
		%40	0.08	0.12	0.05	0.60	0.61	0.52	0.67	0.76	0.60	
0.9		1000	%20	0.11	0.14	0.07	0.49	0.51	0.47	0.58	0.58	0.55
			%40	0.11	0.14	0.07	0.51	0.51	0.46	0.59	0.63	0.55
	2000	%20	0.09	0.12	0.06	0.50	0.51	0.48	0.57	0.58	0.55	
		%40	0.09	0.12	0.06	0.52	0.53	0.47	0.60	0.64	0.55	



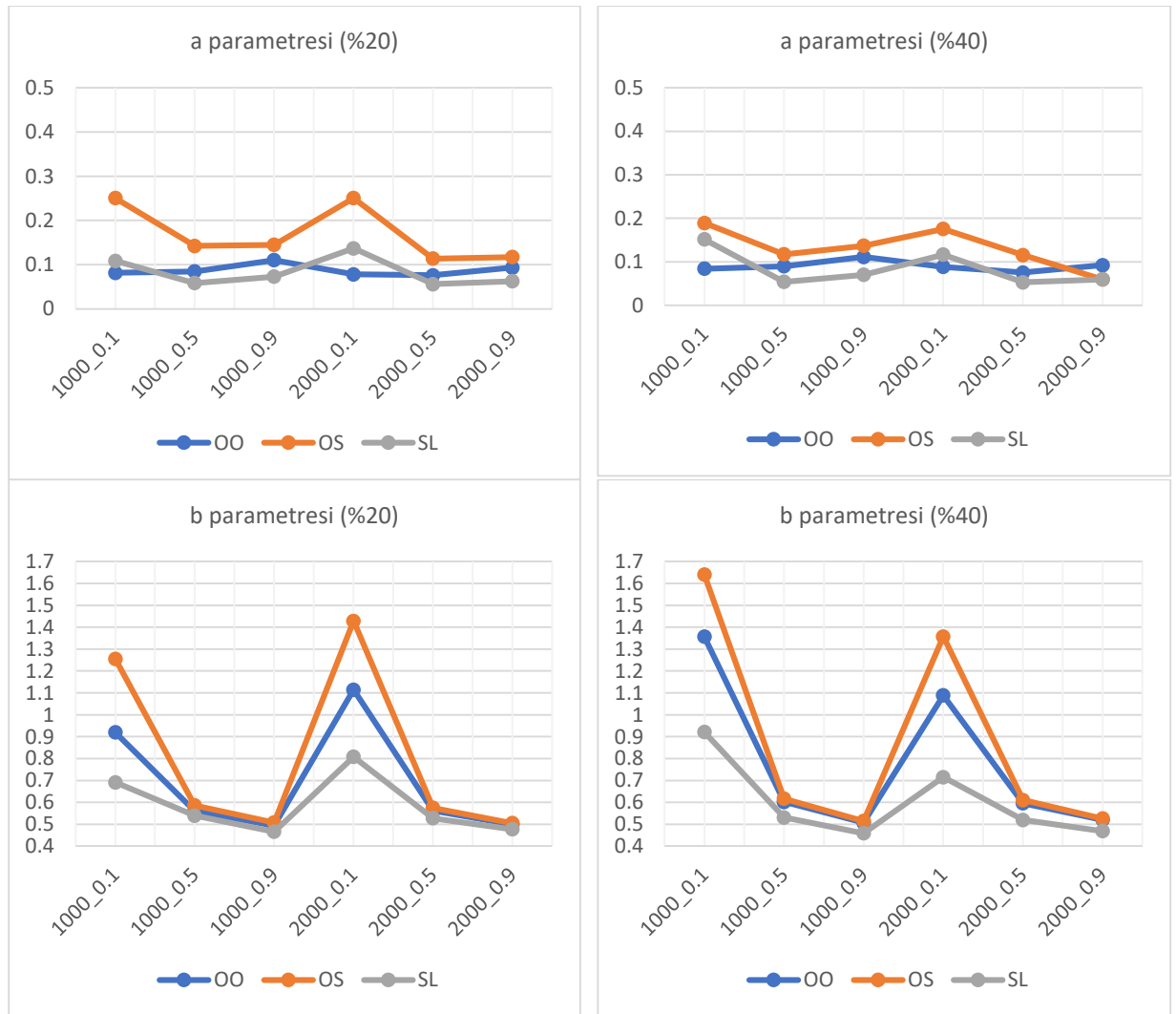
Tablo 9 incelendiğinde, boyutlar arası korelasyon 0.1 ve ortak madde yüzdesi %20 olan durumda örneklem büyüklüğünün artması RMSE değerlerini artırıcı yönde, diğer durumlarda ise örneklem büyüklüğünün artması RMSE değerlerini düşürücü yönde etkide bulunmuştur. Bunun yanında ortak madde oranının artması, ortalama-ortalama ve ortalama-sigma yöntemi kullanılarak kestirilen parametrelere ilişkin RMSE değerlerini genel olarak artırıcı yönde, Stocking-Lord yöntemi kullanılarak elde edilen RMSE değerlerini ise düşürücü yönde etkide bulunmuştur. Tablo 8'e kıyasla, gruplar arası yetenek dağılımları arasındaki fark yüksek olduğunda elde edilen RMSE değerlerinde artış görülmektedir. Bu artış b ve yetenek parametrelerinden kestirilen RMSE değerlerinde daha belirgindir. Yöntemlere ilişkin elde edilen ortalama RMSE değerleri incelendiğinde, ortalama-sigma yöntemine ilişkin değerlerin diğer yöntemlere oranla daha yüksek olduğu görülmektedir. Tüm parametreler için en düşük RMSE değerleri Stocking-Lord yönteminden elde edilmiştir. Bu durum alan yazında bulunan karakteristik eğri yöntemlerinin, moment yöntemlerinden daha düşük hata ürettiği bulgusuyla örtüşmektedir (Baker ve Al-Karni, 1991; Hanson ve Beguin, 2002). Boyutlar arası korelasyon 0.9 (yüksek), örneklem büyüklüğü 2000 (yüksek) ve ortak madde oranı %20 (düşük) olduğu durumda bütün yöntemlere ait RMSE değerlerinin tüm parametreler için en düşük değeri aldığı görülmektedir.

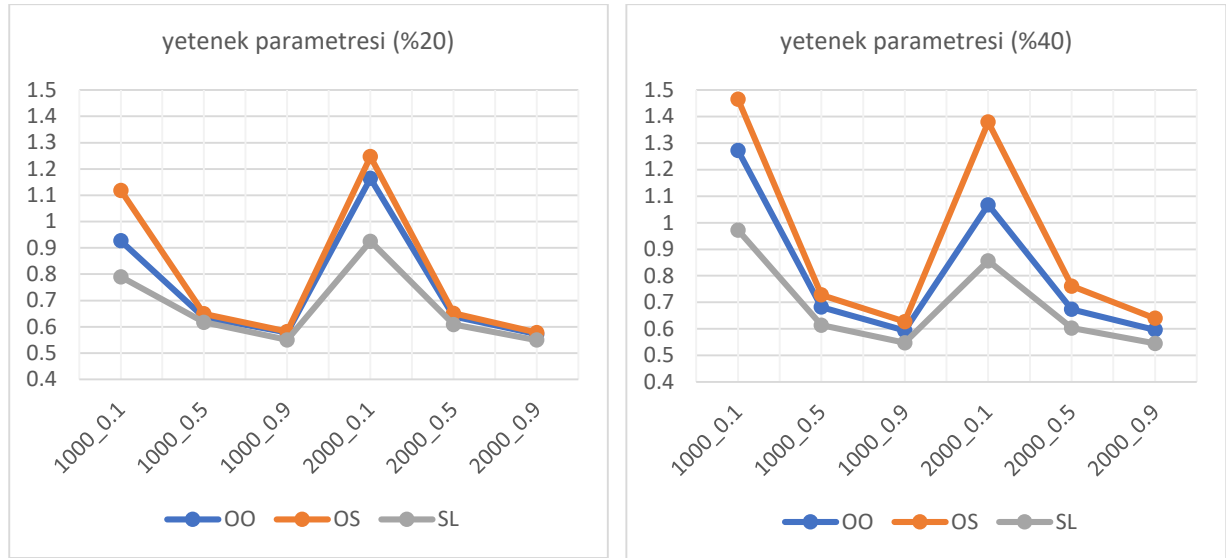
Gruplar arası yetenek dağılımı farkı yüksek olan durumlar için parametre kestirim modellerine göre RMSE değerleri karşılaştırıldığında, b parametrelerinden kestirilen bütün yöntemlere ilişkin RMSE değerleri genel olarak 2PLM'de, yetenek parametrelerinden kestirilen bütün yöntemlere ilişkin RMSE değerleri ise genel olarak 3PLM'de daha düşük olarak hesaplanmıştır. a parametresinden kestirilen RMSE değerleri incelendiğinde, ortalama-ortalama yöntemi kullanılarak kestirilen RMSE değerleri boyutlar arası korelasyon 0.9 iken 3PLM'de, diğer durumlarda 2PLM'de daha düşük kestirilmiştir. Bununla birlikte ortalama-sigma yöntemi kullanılarak kestirilen a parametresine ilişkin RMSE değerleri genel olarak 2PLM'de daha düşük kestirilirken, Stocking-Lord yöntemi kullanılarak kestirilen değerler 3PLM'de daha düşüktür. Bu bağlamda 3PLM ya da 2PLM'den hangisinin

kullanılmasının avantaj sağlayacağı hakkında net bir yorum yapmak güç olmakla birlikte, b parametresinin kestiriminde 2PLM kullanılmasının, a ve yetenek parametrelerinin kestiriminde ise 3PLM kullanılmasının daha doğru kestirimler yapılmasını sağlayacağı yorumu yapılabilir. Şekil 11'de gruplar arası yetenek dağılımı farkı yüksek olduğunda a, b ve yetenek parametrelerine ilişkin 2PLM kullanılarak yapılan tek boyutlu ölçek dönüştürme sonucu tüm koşullar altında elde edilen RMSE değerleri her bir parametre için ayrı grafikler halinde gösterilmiştir.

### Şekil 11

*Gruplar Arası Yetenek Farkı Yüksek Olduğunda (0.5) 2PLM Kullanılarak Tek Boyutlu Ölçek Dönüştürme Sonucu Elde Edilen RMSE Değerleri*





Şekil 11 incelendiğinde, boyutlar arası korelasyon 0.1 olduğunda tüm parametreler ve tüm yöntemler için en yüksek RMSE değerleri elde edilmiş, boyutlar arası korelasyon arttıkça RMSE değerlerinde düşüşler gözlemlenmiştir. Örneklem büyüklüğünün artması ise genel olarak tüm parametrelerden kestirilen RMSE değerlerini düşürücü yönde etkide bulunmuş ve yüksek örneklem büyüklüklerinde daha düşük RMSE değerleri elde edilmiştir. En yüksek RMSE değerleri bütün parametreler için ortalama-sigma yönteminden, en düşük RMSE değerleri ise Stocking-Lord yönteminden elde edilmiştir.

### Alt Problem 3'e İlişkin Bulgular ve Yorumlar

*“Ele alınan koşullara göre yapılan çok boyutlu ve tek boyutlu ölçek dönüştürme işlemleri sonucu elde edilen RMSE değerlerinin koşullar arası etkileşimi ve anlamlılık düzeyi nasıldır?”*

Alt problemin çözümüne yönelik olarak, çalışma kapsamında türetilmiş olan çok boyutlu test verilerine yapılan çok boyutlu ve tek boyutlu ölçek dönüştürme işlemleri sonucunda elde edilen RMSE değerleri çalışma kapsamında ele alınan koşullara göre sırasıyla karşılaştırılmıştır. İlk olarak çok boyutlu veri yapısına uygulanan çok boyutlu ve tek boyutlu ölçek dönüştürme sonucu kestirilen a, b ve yetenek parametrelerine ilişkin kestirilen RMSE değerleri arasında manidar farklılık olup olmadığı ilişkisiz örneklem t testi ile analiz edilmiştir. Böylelikle çok boyutlu veri yapısına çok boyutlu ve tek boyutlu ölçek dönüştürme

süreçleri uygulandığında elde edilen eşitleme hataları arasında anlamlı bir farklılık olup olmadığı sorusuna yanıt aranmıştır. Daha sonra ise çok boyutlu ve tek boyutlu ölçek dönüştürme yöntemlerinden elde edilen RMSE değerleri için ayrı ayrı, ele alınan koşullara göre a, b ve yetenek parametrelerinden her biri için yapılan ölçek dönüştürme işlemi sonucu elde edilen RMSE değerlerinin değişiminin anlamlılık düzeyini belirlemek için 5 yönlü ANOVA testi yapılmıştır. Elde edilen tablolar ayrı ayrı sunulmuş ve yorumlanmıştır.

**Testin Bütününe Uygulanan Çok Boyutlu ve Tek Boyutlu Ölçek Dönüştürme Yöntemlerinden Kestirilen RMSE Değerlerinin Karşılaştırılması.** Bu kısımda çok boyutlu test formlarına sırasıyla uygulanan çok boyutlu ve tek boyutlu ölçek dönüştürme işlemleri sonucu elde edilen RMSE değerleri karşılaştırılmış ve bu değerler arasında bulunan farkın anlamlılığı ilişkisiz örneklem t testi ile incelenmiştir. Bununla birlikte etki büyüklüğünün hesaplanmasında Cohen standardize edilmiş etki büyüklüğü indeksi olan d istatistiği kullanılmıştır (Cohen, 1988). Buna göre d değeri .2 ise küçük, .5 ise orta ve .8 ise büyük etki büyüklüğü olarak yorumlanabilir. Tablo 10'da çok boyutlu test verisine uygulanan çok boyutlu ve tek boyutlu ölçek dönüştürme sonucu kestirilen ortalama RMSE değerleri arasındaki farkın anlamlılığını araştıran karşılaştırma testi sonuçları sunulmuştur.

**Tablo 10**

*Çok Boyutlu ve Tek Boyutlu Ölçek Dönüştürme Yöntemlerine Göre Kestirilen Ortalama RMSE Değerlerinin Karşılaştırılması*

	a			b			yetenek		
	OO	OS	SL	OO	OS	SL	OO	OS	SL
Çok Boyutlu									
Ort. RMSE	0.10	0.35	0.14	0.30	0.37	0.32	0.23	0.28	0.29
Tek Boyutlu									
Ort. RMSE	0.09	0.17	0.06	0.52	0.62	0.37	0.48	0.54	0.37
t	6.018	20.946	35.442	-20.418	-19.885	-6.613	-25.557	-24.345	-11.186
p değeri	.013	.000	.000	.000	.000	.000	.000	.000	.000
d	0.19	0.49	4.42	0.77	0.54	0.16	0.69	0.57	0.28

Tablo 10 incelendiğinde, a parametresi için tek boyutlu yöntemlerde, b parametresi ve yetenek parametresi için çok boyutlu yöntemlerde daha düşük ortalama RMSE değerleri elde edildiği görülmektedir. Elde edilen RMSE değerleri a parametresi için tek boyutlu yöntemler lehine, b parametresi ve yetenek parametresi için çok boyutlu yöntemler lehine anlamlı şekilde farklılık göstermiştir ( $p < .05$ ). a parametresinde ortalama-ortalama yöntemi için düşük düzeyde etki bulunurken, ortalama-sigma yöntemi için orta ve Stocking-Lord yöntemi için yüksek düzeyde etki bulunmuştur. b parametresinde ve yetenek parametresinde ortalama-ortalama ve ortalama-sigma yöntemi için etki büyüklüğü orta düzeyde bulunurken, Stocking-Lord yöntemi için düşük düzeyde etki bulunmuştur. Bu durumda, testin boyutluluk yapısına uygun model seçilerek parametre kestirimi yapılmasının model-veri uyumunu ve dolayısıyla testin yapı geçerliğini artıracaklarını, böylelikle daha düşük hatalı ölçek dönüştürme işlemleri gerçekleştirilebileceği yorumu yapılabilir.

**Çok Boyutlu Ölçek Dönüştürme İşleminde Elde Edilen RMSE Değerlerinin Çalışmada Yer Alan Koşullara Göre İncelenmesi ve Karşılaştırılması.** Bu bölümde, çok boyutlu test verisine yapılan çok boyutlu ölçek dönüştürme işleminden elde edilen RMSE değerleri çalışmada ele alınan koşullara ve bunların birbirleriyle olan etkileşimlerine göre incelenip karşılaştırılmış ve elde edilen RMSE değerlerinin ortalamaları arasındaki farkın anlamlılığı 5 yönlü ANOVA testi ile incelenmiştir. Koşulların birbirleriyle etkileşimleri incelenirken, ikili etkileşimler alınmış, ikiden fazla olan etkileşimler anlamlı etkiye sahip olmaması ve yorumlanmasındaki güçlük nedeniyle tabloya eklenmemiştir. Tablo 11’de çok boyutlu ölçek dönüştürme sonucu koşullara ve alt koşullara ilişkin a parametresinden kestirilen ortalama RMSE değerleri verilmiştir.

**Tablo 11**

*Çok boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre a parametresinden kestirilen ortalama RMSE değerleri*

Koşul	Alt Koşul	OO	OS	SL
Örnekleme	1000	0.11	0.40	0.15
Büyüklüğü	2000	0.09	0.31	0.13
Boyutlar Arası Korelasyon	0.1	0.09	0.32	0.12
	0.5	0.10	0.36	0.15
	0.9	0.10	0.38	0.15
Ortak Madde Oranı	%20	0.10	0.40	0.13
	%40	0.10	0.31	0.15
Gruplar Arası Yetenek Dağılımı Farkı	Düşük	0.09	0.31	0.14
	Yüksek	0.11	0.39	0.14
Parametre	3PLM	0.10	0.35	0.15
Kestirim Modeli	2PLM	0.10	0.36	0.13

Tablo 11 incelendiğinde a parametresi için örneklem büyüklüğü 2000, boyutlar arası korelasyon 0.1 ve gruplar arası yetenek dağılımı farkı düşük olduğu durumlarda tüm yöntemlerden elde edilen ortalama RMSE değerlerinin daha düşük olduğu görülmektedir. Tablo 12'de a parametresine ilişkin elde edilen RMSE değerlerinin ele alınan koşullar ve bunların etkileşimlerine göre değişimini incelemeye yönelik yapılan ANOVA testinin sonuçları sunulmuştur.

**Tablo 12**

*Çok boyutlu ölçek dönüştürmede a parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması*

Koşullar	OO (a)		OS (a)		SL (a)	
	Test	p değeri	Test	p değeri	Test	p değeri
	İstatistiği		İstatistiği		İstatistiği	
	(F)		(F)		(F)	
Örn	69.081	.000	33.599	.000	17.425	.000
Kor	1.629	.196	5.307	.005	25.738	.000
Ort	1.529	.216	28.496	.000	104.310	.000
Yet	127.755	.000	23.725	.000	1.507	.220
Par	3.541	.060	1.315	.252	20.044	.000
Örn*Kor	15.971	.000	1.823	.162	28.738	.000
Örn*Ort	1.381	.240	11.590	.001	48.794	.000
Örn*Yet	0.117	.732	0.388	.534	0.003	.954
Örn*Par	9.261	.002	1.351	.245	25.373	.000
Kor*Ort	2.123	.120	3.025	.049	12.437	.000
Kor*Yet	1.261	.284	0.356	.701	1.096	.334
Kor*Par	12.532	.000	1.371	.254	20.612	.000
Ort*Yet	5.622	.018	1.031	.310	18.878	.000
Ort*Par	134.899	.000	5.042	.025	561.981	.000
Yet*Par	204.196	.000	12.579	.000	76.707	.000

\*Örn: örneklem büyüklüğü, Kor: boyutlar arası korelasyon, Ort: ortak madde oranı, Yet: gruplar arası yetenek dağılımı farkı, Par: parametre kestirim modeli

Tablo 12 incelendiğinde, tüm yöntemler için örneklem büyüklüğü 2000 olan koşul lehine, ortalama-sigma ve Stocking-Lord yöntemleri için boyutlar arası korelasyonun 0.1 olduğu koşul lehine, ortalama-ortalama ve ortalama-sigma yöntemleri için gruplar arası yetenek dağılımı farkı düşük olan koşul lehine a parametresinden kestirilen ortalama RMSE değerleri arasında anlamlı farklılık bulunmuştur ( $p < .05$ ). Bununla birlikte gruplar arası

yetenek dağılımı farkı ve parametre kestirim modelinin ortak etkisinin tüm yöntemler için a parametresinden kestirilen RMSE değerinin değişiminde anlamlı etkiye sahip olduğu görülmektedir. Boyutlar arası korelasyonda 3 düzey bulunduğundan, anlamlılığın hangi düzeyler arasında olduğunun tespiti için ikili karşılaştırma (post-hoc) testi yapılmıştır. İkili karşılaştırma için Tukey testi kullanılmıştır. Ortalama-sigma ve Stocking-Lord yöntemlerinden kestirilen RMSE değerleri üzerinden gerçekleştirilen Tukey testi sonuçları tablo 13'de verilmiştir.

**Tablo 13**

*Çok boyutlu ölçek dönüştürmede a parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları*

Boyutlar Arası Korelasyon	OS (a)		SL (a)	
	Ortalama Fark	p değeri	Ortalama Fark	p değeri
0.1-0.5	-0.411	.100	-0.218	.000
0.5-0.9	-0.234	.473	0.002	.649
0.1-0.9	-0.646	.004	-0.189	.000

Tablo 12 incelendiğinde, çok boyutlu ölçek dönüştürmede a parametresinden kestirilen RMSE değerlerinin ortalamaları arasındaki fark ortalama-sigma yöntemi için boyutlar arası korelasyonun 0.1 ile 0.5 olduğu koşullar arasında 0.1 olan koşul lehine, Stocking-Lord yöntemi için boyutlar arası korelasyonun 0.1 ile 0.5 olduğu ve 0.1 ile 0.9 olduğu koşullar arasında 0.1 olan koşul lehine anlamlı bulunmuştur ( $p < .05$ ). Diğer koşullar arasında istatistiksel olarak anlamlı bir fark bulunmamaktadır. Tablo 14'te çok boyutlu ölçek dönüştürme sonucu koşullara ve alt koşullara göre b parametresinden kestirilen ortalama RMSE değerleri verilmiştir.



**Tablo 14**

*Çok boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre b parametresinden kestirilen ortalama RMSE değerleri*

Koşul	Alt Koşul	OO	OS	SL
Örnekleme	1000	0.33	0.41	0.33
Büyükklüğü	2000	0.26	0.33	0.32
Boyutlar Arası Korelasyon	0.1	0.27	0.36	0.31
	0.5	0.30	0.37	0.32
	0.9	0.31	0.39	0.33
Ortak Madde Oranı	%20	0.31	0.42	0.31
	%40	0.27	0.32	0.33
Gruplar Arası Yetenek Dağılımı Farkı	Düşük	0.18	0.29	0.25
	Yüksek	0.41	0.45	0.40
Parametre	3PLM	0.23	0.39	0.34
Kestirim Modeli	2PLM	0.36	0.36	0.31

Tablo 14'e göre, b parametresi için örneklem büyüklüğü 2000, boyutlar arası korelasyon 0.1, ortak madde oranı %40 ve gruplar arası yetenek dağılımı farkı düşük olan koşullarda daha düşük ortalama RMSE değerleri elde edilmiştir. Tablo 15'te çok boyutlu ölçek dönüştürme sonucu b parametresine ilişkin ele alınan koşullar çerçevesinde elde edilen ortalama RMSE değerlerinin değişimine ilişkin yapılan ANOVA testi sonuçları sunulmuştur.

**Tablo 15**

*Çok boyutlu ölçek dönüştürmede b parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması*

Koşullar	OO (b)		OS (b)		SL (b)	
	Test İstatistiği (F)	p değeri	Test İstatistiği (F)	p değeri	Test İstatistiği (F)	p değeri
Örn	31.916	.000	66.396	.000	1.922	.166
Kor	2.430	.088	2.134	.119	1.926	.146
Ort	9.720	.002	94.746	.000	39.091	.000
Yet	394.071	.000	210.249	.000	1517.201	.000
Par	118.438	.000	8.267	.004	42.958	.000
Örn*Kor	1.255	.285	0.233	.793	35.657	.000
Örn*Ort	9.011	.003	2.245	.134	19.525	.000
Örn*Yet	.666	.415	0.703	.402	1.692	.193
Örn*Par	1.593	.207	1.124	.289	0.113	.737
Kor*Ort	1.316	.268	0.353	.703	2.748	.064
Kor*Yet	0.171	.842	1.574	.207	16.628	.000
Kor*Par	2.253	.105	0.219	.804	8.734	.000
Ort*Yet	6.012	.014	20.133	.000	41.936	.000
Ort*Par	1.133	.287	9.816	.002	82.774	.000
Yet*Par	37.436	.000	29.923	.000	1.049	.306

\*Örn: örneklem büyüklüğü, Kor: boyutlar arası korelasyon, Ort: ortak madde oranı, Yet: gruplar arası yetenek dağılımı farkı, Par: parametre kestirim modeli

Tablo 15 incelendiğinde ortak madde oranı, gruplar arası yetenek dağılımı farkı ve parametre kestirim modeli tüm yöntemlerden kestirilen b parametrelerine ilişkin RMSE değerlerinin değişiminde anlamlı bir etkiye sahiptir ( $p < .05$ ). Örneklem büyüklüğü ise ortalama-ortalama ve ortalama-sigma yöntemleri kullanılarak kestirilen RMSE değerlerinde anlamlı farklılığa neden olmuştur. Boyutlar arası korelasyon, b parametrelerinden kestirilen

RMSE değerlerinin değişiminde anlamlı etkiye sahip değildir. Gruplar arası yetenek dağılımı farkı - ortak madde oranı ve yine gruplar arası yetenek dağılımı farkı - parametre kestirim modeli ortak etkisi, b parametresine ait RMSE değerinin kestiriminde anlamlı etkiye sahiptir. Çok boyutlu ölçek dönüştürme sonucu koşullara ve alt koşullara göre yetenek parametresinden kestirilen ortalama RMSE değerleri tablo 16'da verilmiştir.

**Tablo 16**

*Çok boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre yetenek parametresinden kestirilen ortalama RMSE değerleri*

Koşul	Alt Koşul	OO	OS	SL
Örneklem	1000	0.25	0.31	0.29
Büyüklüğü	2000	0.20	0.26	0.29
Boyutlar Arası	0.1	0.22	0.28	0.29
	0.5	0.23	0.28	0.28
	0.9	0.24	0.29	0.29
Ortak Madde	%20	0.24	0.30	0.28
	Oranı	%40	0.21	0.27
Gruplar Arası	Düşük	0.13	0.21	0.18
	Yetenek Dağılımı	Yüksek	0.33	0.36
Farkı				
Parametre	3PLM	0.19	0.27	0.29
Kestirim Modeli	2PLM	0.26	0.30	0.29

Tablo 16 incelendiğinde, örneklem büyüklüğü 2000, ortak madde oranı %40 ve gruplar arası yetenek dağılımı farkı düşük olan koşullarda, 3PLM kullanıldığında daha düşük RMSE değerleri elde edildiği görülmektedir. Çok boyutlu ölçek dönüştürme sonucu yetenek parametresine ilişkin incelenen koşullar çerçevesinde elde edilen RMSE değerlerinin değişimine ilişkin yapılan ANOVA testi sonuçları tablo 17'de sunulmuştur.

**Tablo 17**

*Çok boyutlu ölçek dönüştürmede yetenek parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması*

Koşullar	OO (yetenek)		OS (yetenek)		SL (yetenek)	
	Test	p değeri	Test	p değeri	Test	p değeri
	İstatistiği		İstatistiği		İstatistiği	
	(F)		(F)		(F)	
Örn	28.901	.000	53.198	.000	0.369	.543
Kor	1.645	.193	2.319	.099	2.573	.077
Ort	8.619	.003	21.496	.000	7.918	.005
Yet	604.097	.000	569.782	.000	2752.025	.000
Par	63.156	.000	16.671	.000	0.176	.675
Örn*Kor	0.666	.514	0.186	.000	26.800	.000
Örn*Ort	6.627	.010	2.880	.090	8.914	.003
Örn*Yet	1.510	.219	5.401	.020	1.286	.257
Örn*Par	2.575	.109	0.889	.346	0.306	.580
Kor*Ort	1.088	.337	0.559	.572	0.535	.585
Kor*Yet	0.401	.670	2.179	.113	7.882	.000
Kor*Par	2.252	.105	0.188	.829	4.295	.014
Ort*Yet	5.682	.017	3.217	.073	0.806	.369
Ort*Par	0.239	.625	35.513	.000	24.100	.000
Yet*Par	9.661	.002	0.015	.902	157.858	.000

\*Örn: örneklem büyüklüğü, Kor: boyutlar arası korelasyon, Ort: ortak madde oranı, Yet: gruplar arası yetenek dağılımı farkı, Par: parametre kestirim modeli

Çok boyutlu ölçek dönüştürmede yetenek parametrelerinin kestiriminde ortak madde oranı ve gruplar arası yetenek dağılımları farkı tüm yöntemlere ilişkin kestirilen RMSE değerlerinde anlamlı şekilde farklılığa neden olmuştur ( $p < .05$ ). Benzer şekilde ortalama-ortalama ve ortalama-sigma yöntemlerinde örneklem büyüklüğü ve parametre kestirim modeli, yetenek parametresinden kestirilen RMSE değerlerinin değişiminde

anlamli etkiye sahiptir. Boyutlar arasi korelasyon ise yontemlerden kestirilen RMSE degerlerinin deęişiminde anlamli etkiye sahip deęildir. Örnekleme büyüklüğü – boyutlar arasi korelasyon ortak etkisi ile ortak madde oranı – parametre kestirim modeli ortak etkisi ortalama-sigma ve Stocking-Lord yöntemlerinden kestirilen RMSE degerlerinde anlamli etkiye sahiptir. Yine gruplar arasi yetenek daęılımı farkı – parametre kestirim modeli ortak etkisi ortalama-ortalama ve Stocking-Lord yöntemlerinden kestirilen RMSE degerlerinde anlamli etkiye sahiptir. Tablo 18’de tek boyutlu ölçek dönüştürme sonucu koşullara ve alt koşullara göre a parametresinden elde edilen ortalama RMSE degerleri verilmiştir.

**Tablo 18**

*Tek boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre a parametresinden kestirilen ortalama RMSE degerleri*

Koşul	Alt Koşul	OO	OS	SL
Örnekleme Büyüküğü	1000	0.09	0.18	0.06
	2000	0.08	0.15	0.05
Boyutlar Arasi Korelasyon	0.1	0.09	0.23	0.09
	0.5	0.08	0.13	0.04
	0.9	0.08	0.13	0.04
Ortak Madde Oranı	%20	0.08	0.19	0.06
	%40	0.09	0.14	0.05
Gruplar Arasi Yetenek Daęılımı Farkı	Düşük	0.07	0.17	0.05
	Yüksek	0.09	0.16	0.07
Parametre Kestirim Modeli	3PLM	0.10	0.18	0.06
	2PLM	0.07	0.15	0.06

Tablo 18’e göre örnekleme büyüklüğü 2000, boyutlar arasi korelasyon 0.9 ve gruplar arasi yetenek daęılımı farkı düşük olan koşullarda daha düşük ortalama RMSE degerleri elde edildięi görülmektedir. Tablo 19’da tek boyutlu ölçek dönüştürme sonucu a

parametresine ilişkin incelenen koşullar çerçevesinde elde edilen RMSE değerlerinin değişimine ilişkin yapılan ANOVA testi sonuçları sunulmuştur.

**Tablo 19**

*Tek boyutlu ölçek dönüştürmede a parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması*

Koşullar	OO (a)		OS (a)		SL (a)	
	Test İstatistiği (F)	p değeri	Test İstatistiği (F)	p değeri	Test İstatistiği (F)	p değeri
Örn	7.218	.007	20.470	.000	7.292	.007
Kor	4.474	.011	149.369	.000	127.020	.000
Ort	0.580	.446	84.199	.000	4.487	.034
Yet	11.031	.001	1.598	.206	75.218	.000
Par	101.745	.000	25.094	.000	0.516	.472
Örn*Kor	7.368	.001	2.716	.066	0.450	.637
Örn*Ort	0.129	.720	2.286	.131	0.034	.854
Örn*Yet	0.342	.558	0.194	.660	0.029	.865
Örn*Par	2.468	.116	0.772	.380	0.051	.821
Kor*Ort	0.276	.759	2.334	.097	0.849	.428
Kor*Yet	5.526	.004	1.776	.170	3.866	.021
Kor*Par	10.795	.000	0.391	.677	2.070	.126
Ort*Yet	1.566	.211	3.255	.071	2.205	.138
Ort*Par	0.044	.835	0.403	.526	0.714	.398
Yet*Par	63.924	.000	6.839	.009	53.026	.000

\*Örn: örneklem büyüklüğü, Kor: boyutlar arası korelasyon, Ort: ortak madde oranı, Yet: gruplar arası yetenek dağılımı farkı, Par: parametre kestirim modeli

Tablo 19 incelendiğinde, tek boyutlu ölçek dönüştürmede örneklem büyüklüğü ve boyutlar arası korelasyonun tüm yöntemlere ilişkin a parametresinden kestirilen RMSE değerlerinde anlamlı bir etkiye sahip olduğu görülmektedir ( $p < .05$ ). Bununla birlikte gruplar

arası yetenek dağılımı farkı ortalama-ortalama ve Stocking-Lord yöntemlerinde, parametre kestirim modeli ortalama-ortalama ve ortalama-sigma yöntemlerinde, ortak madde oranı ise ortalama-sigma ve Stocking-Lord yöntemlerinde  $a$  parametresinden kestirilen RMSE değerlerinde anlamlı farklılığa neden olmuştur. Bununla birlikte gruplar arası yetenek dağılımı farkı ve parametre kestirim modelinin ortak etkisi tüm yöntemlere ilişkin kestirilen RMSE değerlerinin değişiminde anlamlı etkiye sahiptir. Boyutlar arası korelasyonda anlamlılığın hangi düzeyler arasında olduğunun tespiti için Tukey ikili karşılaştırma (post-hoc) testi yapılmıştır ve sonuçlar tablo 20'de verilmiştir.

**Tablo 20**

*Tek boyutlu ölçek dönüştürmede  $a$  parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları*

Boyutlar Arası Korelasyon	OO (a)		OS (a)		SL (a)	
	Ortalama Fark	p değeri	Ortalama Fark	p değeri	Ortalama Fark	p değeri
0.1-0.5	0.011	.013	0.105	.000	0.048	.000
0.5-0.9	-0.002	.066	0.002	.967	0.001	.921
0.1-0.9	0.009	.815	0.106	.000	0.503	.000

Tablo 20 incelendiğinde, tek boyutlu ölçek dönüştürmede  $a$  parametrelerinden kestirilen RMSE değerlerinin ortalamaları arasındaki fark tüm ölçek dönüştürme yöntemleri için boyutlar arası korelasyonun 0.1 ile 0.5 arasında olduğu koşullarda 0.5 olan koşul lehine, ortalama-sigma ve Stocking-Lord yöntemleri için boyutlar arası korelasyonun 0.1 ile 0.9 arasında olduğu koşullarda 0.9 olan koşul lehine anlamlı bulunmuştur. Tek boyutlu ölçek dönüştürme sonucu koşullara ve alt koşullara göre  $b$  parametresinden kestirilen ortalama RMSE değerleri tablo 21'de verilmiştir.

**Tablo 21**

*Tek boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre b parametresinden kestirilen ortalama RMSE değerleri*

Koşul	Alt Koşul	OO	OS	SL
Örneklem	1000	0.53	0.63	0.38
Büyükklüğü	2000	0.50	0.60	0.36
Boyutlar Arası	0.1	0.87	1.13	0.53
	0.5	0.36	0.39	0.31
	0.9	0.31	0.33	0.27
Ortak Madde	%20	0.50	0.62	0.39
Oranı	%40	0.52	0.61	0.36
Gruplar Arası	Düşük	0.28	0.40	0.12
Yetenek Dağılımı	Yüksek	0.75	0.83	0.62
Parametre	3PLM	0.51	0.58	0.37
Kestirim Modeli	2PLM	0.52	0.65	0.37

Tablo 21 incelendiğinde, örneklem büyüklüğü 2000, boyutlar arası korelasyon 0.9 ve gruplar arası yetenek dağılımı farkı düşük olan koşullarda daha düşük ortalama RMSE değerleri elde edildiği görülmektedir. Tek boyutlu ölçek dönüştürme sonucu b parametresine ilişkin incelenen koşullar çerçevesinde elde edilen RMSE değerlerinin değişimine ilişkin yapılan ANOVA testi sonuçları tablo 22'de sunulmuştur.



**Tablo 22**

*Tek boyutlu ölçek dönüştürmede b parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması*

Koşullar	OO (b)		OS (b)		SL (b)	
	Test İstatistiği (F)	p değeri	Test İstatistiği (F)	p değeri	Test İstatistiği (F)	p değeri
Örn	3.963	.047	7.088	.008	3.759	.053
Kor	1436.862	.000	1526.829	.000	366.812	.000
Ort	6.224	.013	1.110	.292	8.921	.003
Yet	2478.842	.000	1048.402	.000	3462.171	.000
Par	0.166	.684	23.841	.000	0.221	.638
Örn*Kor	0.634	.531	0.241	.786	0.342	.710
Örn*Ort	15.668	.000	5.453	.020	7.176	.007
Örn*Yet	0.012	.913	0.036	.850	0.159	.690
Örn*Par	0.523	.470	0.558	.455	0.386	.534
Kor*Ort	4.584	.010	0.027	.973	2.328	.098
Kor*Yet	17.058	.000	1.767	.171	8.233	.000
Kor*Par	7.880	.000	37.735	.000	12.206	.000
Ort*Yet	25.996	.000	21.571	.000	20.463	.000
Ort*Par	0.114	.735	0.292	.589	0.776	.379
Yet*Par	6.405	.011	4.588	.032	24.496	.000

\*Örn: örneklem büyüklüğü, Kor: boyutlar arası korelasyon, Ort: ortak madde oranı, Yet: gruplar arası yetenek dağılımı farkı, Par: parametre kestirim modeli

Tablo 22 incelendiğinde, tek boyutlu ölçek dönüştürmede örneklem büyüklüğü, boyutlar arası korelasyon ve gruplar arası yetenek dağılımı farkı, tüm yöntemlere ilişkin b parametresinden kestirilen RMSE değerlerinde anlamlı şekilde farklılığa neden olmuştur ( $p < .05$ ). Bununla birlikte ortak madde oranı ortalama-ortalama ve Stocking- Lord yöntemleri, parametre kestirim modeli ise ortalama-sigma yöntemi kullanılarak b parametresinden elde

edilen RMSE değerlerinin değişiminde anlamlı etkiye sahiptir. Örneklem büyüklüğü – ortak madde oranı, boyutlar arası korelasyon – parametre kestirim modeli, ortak madde oranı – gruplar arası yetenek dağılımı farkı ve gruplar arası yetenek dağılımı farkı – parametre kestirim modeli değişkenlerinin ortak etkileri de tüm yöntemlerden kestirilen RMSE değerlerinin değişiminde anlamlı etkiye sahiptir. Boyutlar arası korelasyonda anlamlılığın hangi düzeyler arasında olduğunun tespiti için yapılan Tukey ikili karşılaştırma (post-hoc) testi sonuçları tablo 23'te verilmiştir.

**Tablo 23**

*Tek boyutlu ölçek dönüştürmede b parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları*

Boyutlar Arası Korelasyon	OO (b)		OS (b)		SL (b)	
	Ortalama Fark	p değeri	Ortalama Fark	p değeri	Ortalama Fark	p değeri
0.1-0.5	.512	.000	.738	.000	.219	.000
0.5-0.9	.053	.000	.062	.000	.041	.000
0.1-0.9	.565	.000	.800	.000	.260	.000

Tablo 23 incelendiğinde, tek boyutlu ölçek dönüştürmede b parametrelerinden kestirilen RMSE değerlerinin ortalamaları arasındaki fark tüm ölçek dönüştürme yöntemleri için boyutlar arası korelasyonun tüm alt koşulları arasında boyutlar arası korelasyonun yüksek olduğu koşullar lehine istatistiksel anlamlı olarak bulunmuştur ( $p < .05$ ). Tablo 24'te tek boyutlu ölçek dönüştürme sonucu koşullara ve alt koşullara göre yetenek parametresinden kestirilen ortalama RMSE değerleri verilmiştir.

**Tablo 24**

*Tek boyutlu ölçek dönüştürmede koşullara ve alt koşullara göre yetenek parametresinden kestirilen ortalama RMSE değerleri*

Koşul	Alt Koşul	OO	OS	SL
Örneklem	1000	0.49	0.55	0.38
Büyükklüğü	2000	0.47	0.54	0.37
Boyutlar Arası	0.1	0.76	0.87	0.50
	0.5	0.36	0.41	0.33
	0.9	0.32	0.35	0.29
Ortak Madde	%20	0.48	0.54	0.38
Oranı	%40	0.48	0.54	0.36
Gruplar Arası	Düşük	0.20	0.26	0.09
Yetenek Dağılımı	Yüksek	0.76	0.82	0.66
Parametre	3PLM	0.46	0.50	0.36
Kestirim Modeli	2PLM	0.50	0.59	0.39

Tablo 24'e göre, boyutlar arası korelasyon 0.9, gruplar arası yetenek dağılımı farkı düşük olan koşullarda ve 3PLM kullanılan koşullarda daha düşük ortalama RMSE değerleri elde edildiği görülmektedir. Tek boyutlu ölçek dönüştürme sonucu yetenek parametresine ilişkin incelenen koşullar çerçevesinde elde edilen RMSE değerlerinin değişimine ilişkin yapılan ANOVA testi sonuçları tablo 25'te sunulmuştur.

**Tablo 25**

*Tek boyutlu ölçek dönüştürmede yetenek parametresinden kestirilen RMSE değerlerinin koşullara göre karşılaştırılması*

Koşullar	OO (yetenek)		OS (yetenek)		SL (yetenek)	
	Test	p değeri	Test	p değeri	Test	p değeri
	İstatistiği		İstatistiği		İstatistiği	
	(F)		(F)		(F)	
Örn	1.397	.237	1.017	.313	1.750	.186
Kor	755.062	.000	636.509	.000	368.570	.000
Ort	0.109	.742	0.002	.967	5.821	.016
Yet	3095.451	.000	1869.236	.000	7048.279	.000
Par	17.903	.000	51.487	.000	31.114	.000
Örn*Kor	0.717	.488	0.847	.429	1.366	.255
Örn*Ort	8.499	.004	0.543	.461	7.539	.006
Örn*Yet	0.304	.582	1.546	.214	0.256	.613
Örn*Par	1.456	.228	0.085	.771	1.457	.228
Kor*Ort	0.167	.846	1.876	.153	0.334	.716
Kor*Yet	43.242	.000	43.035	.000	72.464	.000
Kor*Par	4.939	.007	13.387	.000	15.621	.000
Ort*Yet	38.137	.000	21.402	.000	14.167	.000
Ort*Par	0.100	.751	0.005	.945	1.461	.227
Yet*Par	6.554	.011	3.506	.061	7.781	.005

\*Örn: örneklem büyüklüğü, Kor: boyutlar arası korelasyon, Ort: ortak madde oranı, Yet: gruplar arası yetenek dağılımı farkı, Par: parametre kestirim modeli

Tek boyutlu ölçek dönüştürmede yetenek parametrelerinin kestiriminde elde edilen RMSE değerlerinde boyutlar arası korelasyon, gruplar arası yetenek dağılımı farkı ve parametre kestirim modeli anlamlı etkiye sahiptir ( $p < .05$ ). Ortak madde oranı ise yalnızca Stocking-Lord yönteminden kestirilen RMSE değerlerinde anlamlı etkiye sahiptir. Örneklem büyüklüğü tüm yöntemlerden kestirilen RMSE değerlerinin değişiminde anlamlı etkiye sahip

değildir. Buna ek olarak boyutlar arası korelasyon – gruplar arası yetenek dağılımı farkı, boyutlar arası korelasyon – parametre kestirim modeli ve ortak madde oranı – gruplar arası yetenek dağılımı farkı değişkenlerinin ortak etkileri tüm yöntemlerden kestirilen RMSE değerlerinin değişiminde anlamlı etkiye sahiptir. Gruplar arası yetenek dağılımı farkı – parametre kestirim modeli değişkenlerinin ortak etkisi ortalama-ortalama ve Stocking-Lord yöntemlerinden kestirilen RMSE değerlerinin değişiminde anlamlı etkiye sahiptir. Boyutlar arası korelasyonda anlamlılığın hangi düzeyler arasında olduğunun tespiti için Tukey ikili karşılaştırma (post-hoc) testi yapılmıştır ve sonuçlar tablo 26'da verilmiştir.

**Tablo 26**

*Tek boyutlu ölçek dönüştürmede yetenek parametresinden kestirilen hataların boyutlar arası korelasyona ilişkin ikili karşılaştırma testi sonuçları*

Boyutlar Arası Korelasyon	OO (yetenek)		OS (yetenek)		SL (yetenek)	
	Ortalama Fark	p değeri	Ortalama Fark	p değeri	Ortalama Fark	p değeri
0.1-0.5	.393	.000	.460	.000	.176	.000
0.5-0.9	.043	.002	.054	.002	.034	.000
0.1-0.9	.436	.000	.514	.000	.210	.000

Tablo 26 incelendiğinde, tek boyutlu ölçek dönüştürmede yetenek parametrelerinden kestirilen RMSE değerlerinin ortalamaları arasındaki fark tüm ölçek dönüştürme yöntemleri için boyutlar arası korelasyonun tüm alt koşulları arasında boyutlar arası korelasyonun yüksek olduğu koşul lehine istatistiksel anlamlı olarak bulunmuştur ( $p < .05$ ).

## Bölüm 5

### Sonuç ve Öneriler

Bu çalışmada çok boyutlu veri yapısına uygulanan çok boyutlu ve tek boyutlu ölçek dönüştürme yöntemlerinin performansının çeşitli koşullar altında incelenmesi ve karşılaştırılması amaçlanmıştır. Çalışma kapsamında ele alınan koşullar sırasıyla çok boyutlu ve tek boyutlu ölçek dönüştürme yöntemleri (ortalama-ortalama, ortalama-sigma ve Stocking-Lord), örneklem büyüklüğü (1000 ve 2000), boyutlar arası korelasyon (0.1-0.5-0.9), ortak madde oranı (%20 ve %40), gruplar arası yetenek dağılımı farkı (0.05-0.5) ve parametre kestirim modelidir (3PLM-2PLM). Yapılan analizler sonucu elde edilen bulgular, tablolarla ve şekillerle temsil edilerek dördüncü bölümde detaylı olarak açıklanmıştır. Bu bölümde, çalışma sonucu elde edilen bulgular özetlenmiş ve alan yazında bulunan benzer çalışmalarda elde edilen bulgularla karşılaştırılmıştır. Devamında ise çalışmanın sonucuna yönelik öneriler ve ileride yapılacak çalışmalara yönelik öneriler ayrı başlıklar halinde sunulmuştur.

### Sonuçlar

Çalışma kapsamında elde edilen RMSE değerlerinin çalışmada ele alınan koşullara göre değişimi (ölçek dönüştürme yöntemi, örneklem büyüklüğü, ortak madde oranı, boyutlar arası korelasyon, gruplar arası yetenek dağılımı farkı, parametre kestirim modeli) ayrı ayrı başlıklar halinde açıklanmıştır.

#### **Kullanılan Ölçek Dönüştürme Yöntemine Göre RMSE Değerlerinin Değişimi.**

Çok boyutlu veri yapısına hem çok boyutlu hem de tek boyutlu ölçek dönüştürme işlemleri uygulanmış ve yapılan ölçek dönüştürme sonucunda elde edilen ortalama RMSE değerleri karşılaştırılmıştır. Bunun sonucunda a parametresi için tek boyutlu yöntemler kullanıldığında, b parametresi ve yetenek parametresi için ise çok boyutlu yöntemler kullanıldığında RMSE değerleri düşüş göstermiştir. Bu bağlamda, testin sahip olduğu boyutluluk yapısının bilinerek, en az bilgi kaybıyla ve en az hatayla testin yapısının ortaya

konulabilmesi açısından seçilen ölçek dönüştürme yöntemin önemli olduğu sonucuna varılabilir.

**Örneklem büyüklüğüne göre RMSE değerlerinin değişimi.** Hem çok boyutlu hem de tek boyutlu ölçek dönüştürmede ele alınan tüm koşullar için örneklem büyüklüğündeki artış ile daha düşük RMSE değerleri elde edilmiştir. Bu bulgu alan yazında elde edilen sonuçlarla uyumludur (Atar ve Yeşiltaş, 2017; Bökeoğlu ve diğ, 2022; Gök ve Kelecioğlu, 2014; Hanson ve Beguin, 2002; Speroni 2009). Örneklem büyüklüğündeki artış, a ve b parametreleri için hem çok boyutlu hem de tek boyutlu ölçek dönüştürmede, yetenek parametreleri için ise çok boyutlu ölçek dönüştürmede istatistiksel olarak anlamlı bulunmuştur.

**Ortak Madde Oranına Göre RMSE Değerlerinin Değişimi.** Çok boyutlu ölçek dönüştürmede 3PLM üzerinden gerçekleştirilen analizlerde gruplar arası yetenek dağılımı farkı yüksek olduğunda Stocking-Lord yöntemi kullanılarak b parametresinden ve yetenek parametresinden elde edilen RMSE değerleri ortak madde oranı artışından olumsuz yönde etkilenmiştir. Yine 2PLM üzerinden gerçekleştirilen analizlerde gruplar arası yetenek dağılımı farkı düşük olan durumda Stocking-Lord yöntemi kullanıldığında, gruplar arası yetenek dağılımı farkı yüksek olan durumda ise ortalama-ortalama yöntemi kullanıldığında, ortak madde oranı artışından madde ve yetenek parametrelerine ait RMSE değerleri olumsuz yönde etkilenmiştir.

Tek boyutlu ölçek dönüştürmede 3PLM üzerinden gerçekleştirilen analizlerde gruplar arası yetenek dağılımı farkı düşük olduğunda ortalama-ortalama yöntemi kullanılarak a parametresinden kestirilen RMSE değerleri, yine gruplar arası yetenek dağılımı farkı yüksek olduğunda bütün yöntemlere ilişkin b parametresinden ve yetenek parametresinden kestirilen RMSE değerleri ortak madde oranı artışından olumsuz yönde etkilenmiştir. 2PLM üzerinden gerçekleştirilen analizlerde gruplar arası yetenek dağılımı farkı yüksek olduğunda ortalama-ortalama ve ortalama-sigma yöntemi kullanılarak b

parametresinden ve yetenek parametresinden kestirilen RMSE değerleri ortak madde oranı artışıyla birlikte artış göstermiştir.

Çok boyutlu ve tek boyutlu ölçek dönüştürme yöntemleri için yukarıda açıklanan koşullar dışında ele alınan tüm koşullarda ortak madde oranının artması genel olarak daha düşük RMSE değerlerinin elde edilmesini sağlamıştır. RMSE değerlerindeki değişim çok boyutlu ölçek dönüştürmede a, b ve yetenek parametreleri için istatistiksel olarak anlamlı bulunmuştur. Tek boyutlu ölçek dönüştürmede RMSE değerleri arasındaki fark a ve b parametreleri için anlamlı bulunurken, yetenek parametreleri arasındaki fark yalnızca Stocking-Lord yöntemi için istatistiksel olarak anlamlı bulunmuştur.

**Boyutlar Arası Korelasyona Göre RMSE Değerlerinin Değişimi.** Çok boyutlu ölçek dönüştürmede boyutlar arası korelasyonun düşük olduğu durumlarda, tek boyutlu ölçek dönüştürmede ise boyutlar arası korelasyonun yüksek olduğu durumlarda daha düşük RMSE değerlerinin elde edildiği görülmüştür. Boyutlar arası korelasyonun yüksek olması testin tek boyutluluğuna kanıt olarak gösterilebileceğinden (Zhang, 2012), bu sonucun beklendik olduğu önceki bölümlerde açıklanmıştır. Bu durum alan yazında yer alan boyutlar arasında korelasyonun artışının çok boyutlu ölçek dönüştürmede yüksek hataya sebep olduğu bulgusuyla tutarlıdır (Beguin ve Hanson, 2001; Gübeş, 2019). Boyutlar arası korelasyon değişkeni için yapılan karşılaştırma analizleri çok boyutlu ve tek boyutlu yöntemler için ayrı ayrı olacak şekilde yapılmış ve yorumlanmıştır. Çok boyutlu ölçek dönüştürmede a parametresine ilişkin elde edilen RMSE değerleri ortalama-sigma ve Stocking-Lord yöntemleri için boyutlar arası korelasyonun 0.1 olduğu koşul lehine anlamlı farklılık göstermiştir. Tek boyutlu ölçek dönüştürmede ise a, b ve yetenek parametrelerinin hepsi için RMSE değerleri arasında bulunan farklılık, tüm yöntemler için boyutlar arası korelasyonun 0.9 olduğu koşullar lehine istatistiksel olarak anlamlı bulunmuştur.

**Gruplar Arası Yetenek Dağılımı Farkına Göre RMSE Değerlerinin Değişimi.** Hem çok boyutlu hem de tek boyutlu ölçek dönüştürme yöntemleri için 3PLM üzerinden analizler gerçekleştirildiğinde a parametresine ilişkin RMSE değerleri gruplar arası yetenek



dağılımı farkı yüksek olan koşulda daha düşük olarak elde edilmiştir. Bunun dışında, çalışma kapsamında yer alan bütün parametreler ve ölçek dönüştürme yöntemleri için gruplar arası yetenek dağılımı farkının düşük olması, daha düşük RMSE değerleri elde edilmesini sağlamıştır. Bu bulgu Gök ve Kelecioğlu'nun (2014) test uzunluğu ve örneklem büyüklüğü arttıkça benzer yetenek dağılımına sahip gruplarda daha az eşitleme hatası elde edildiği bulgusuyla örtüşmektedir. Hem çok boyutlu hem de tek boyutlu ölçek dönüştürme yöntemleri için tüm yöntemlerden elde edilen RMSE değerleri arasındaki fark, bütün parametreler için gruplar arası yetenek dağılımı farkı düşük olan koşul lehine istatistiksel olarak anlamlı bulunmuştur.

**Parametre Kestirim Modeline Göre RMSE Değerlerinin Değişimi.** Bu çalışmada sırasıyla 3PLM ve 2PLM kullanılarak çok boyutlu ve tek boyutlu analizler yürütülmüştür. Çok boyutlu ölçek dönüştürme yöntemi kullanıldığı durumlarda elde edilen RMSE değerleri incelendiğinde, gruplar arası yetenek dağılımı farkı düşük olduğunda a parametresinden kestirilen RMSE değerleri bütün yöntemler için 2PLM'de daha düşük kestirilmiştir. b parametresinden kestirilen RMSE değerleri ortalama-ortalama yöntemi için 3PLM'de, ortalama-sigma ve Stocking-Lord yöntemleri için ortak madde oranının %20 olduğu koşullarda 2PLM'de, %40 olduğu koşullarda 3PLM'de daha düşük kestirilmiştir. Yetenek parametresi için ortalama-ortalama ve ortalama-sigma yöntemleri için 3PLM'de, Stocking-Lord yöntemi için 2PLM'de daha düşük RMSE değerleri elde edilmiştir. Gruplar arası yetenek dağılımı farkı yüksek olan durumlarda ise a parametresi için ortak madde oranı %20 olduğunda 2PLM'de, %40 olduğunda 3PLM'de daha düşük RMSE değerleri elde edilmiştir. b parametresi için ortalama-ortalama yöntemi kullanıldığında elde edilen RMSE değerleri 3PLM'de, ortalama-sigma ve Stocking-Lord yöntemleri kullanıldığında 2PLM'de daha düşük RMSE değerleri elde edilmiştir. Yetenek parametresinden kestirilen RMSE değerleri genel olarak 3PLM kullanıldığında daha düşük olarak hesaplanmıştır.

Tek boyutlu ölçek dönüştürme yöntemlerinden elde edilen değerler incelendiğinde, gruplar arası yetenek dağılımı farkı düşük olduğunda a parametresinden tüm yöntemler için

2PLM'de daha düşük RMSE değerleri elde edilmiştir. Boyutlar arası korelasyon 0.1 olan koşulda b ve yetenek parametreleri için tüm yöntemlere ait RMSE değerleri 3PLM'de, boyutlar arası korelasyonun daha yüksek değerleri için 2PLM'de daha düşük RMSE değerleri elde edilmiştir. Gruplar arası yetenek dağılımı farkı yüksek olduğunda a parametresinden kestirilen RMSE değerleri ortalama-ortalama yöntemi için boyutlar arası korelasyon yüksek olduğunda 3PLM'de, diğer koşullarda 2PLM'de daha düşük kestirilmiştir. Ortalama-sigma yöntemi kullanıldığına a parametresinden kestirilen RMSE değerleri 2PLM'de daha düşükken, Stocking-Lord yöntemi kullanıldığında 3PLM'de daha düşüktür. b parametrelerinden kestirilen RMSE değerleri, boyutlar arası korelasyon düşük olduğunda 3PLM'de, boyutlar arası korelasyonun yüksek değerlerinde ise 2PLM'de daha düşük kestirilmiştir. Yetenek parametrelerine ait RMSE değerleri tüm yöntemler için 3PLM'de daha düşük bulunmuştur.

Parametre kestirim modeline göre elde edilen madde ve yetenek parametrelerinden kestirilen RMSE değerlerinin değişimine yönelik yapılan varyans analizi sonuçlarına göre çok boyutlu ölçek dönüştürmede b parametrelerinden kestirilen RMSE değerleri arasında anlamlı farklılık vardır. a parametresinde Stocking-Lord yöntemiyle kestirilen, yetenek parametresinde ise ortalama-ortalama ve ortalama-sigma yöntemiyle kestirilen RMSE değerleri arasındaki fark istatistiksel olarak anlamlı bulunmuştur. Tek boyutlu ölçek dönüştürmede ise yetenek parametrelerinden kestirilen bütün yöntemlere ilişkin RMSE değerleri arasında istatistiksel olarak anlamlı farklılık vardır. a ve b parametrelerine ilişkin ortalama-sigma yöntemi kullanılarak elde edilen RMSE değerleri arasında anlamlı farklılık bulunmuştur. Bökeoğlu ve diğ. (2022) tarafından yapılan çalışmada tek boyutlu eşitlemede 2PLM ve 3PLM kullanılan durumlarda benzer sonuçlar elde edilmiştir. Bu çalışmada yapılan karşılaştırma testi sonuçlarına kullanılan modele göre parametreler için elde edilen RMSE değerleri arasında bulunan istatistiksel olarak anlamlı fark bu durumla çelişmektedir.

## Öneriler

Bu araştırmanın bulgularına yönelik öneriler ile gelecekte yapılacak araştırmalara yönelik öneriler ayrı başlıklar halinde verilmiştir.

**Araştırmanın Sonuçlarına Yönelik Öneriler.** Araştırma sonucu elde edilen bulgular doğrultusunda yapılacak öneriler aşağıda maddeler halinde sıralanmıştır.

- Bu araştırma kapsamında çok boyutlu iki test formuna ele alınan koşullar altında sırasıyla çok boyutlu ve tek boyutlu ölçek dönüştürme yöntemleri uygulanmış ve bunun sonucunda elde edilen RMSE değerleri karşılaştırılmıştır. Devamında ise uygulanan çok boyutlu ve tek boyutlu yöntemlerden kestirilen RMSE değerleri arasında çok boyutlu yöntemler lehine istatistiksel olarak anlamlı fark bulunmuştur. Bu bağlamda eşitlenecek olan testin boyutluluk yapısının bilinerek uygun koşulların, yöntemlerin ve modelin seçilmesi önerilmektedir.
- Örneklem büyüklüğü arttıkça hem tek boyutlu hem de çok boyutlu ölçek dönüştürme yöntemlerinde eşitleme hatalarının anlamlı şekilde azaldığı bulgusuna ulaşılmıştır. Bu bağlamda alan yazında yapılan çalışmalar da dikkate alınarak büyük örneklemeler üzerinden ölçek dönüştürme çalışmaları gerçekleştirilebilir.
- Ortak madde oranı %40 olan koşullarda genel olarak daha düşük RMSE hataları elde edilmiştir. Alan yazında ortak madde oranının test formunda yer alan madde sayısının %20'sinden az olmaması gerektiği belirtilmiştir (Angoff 1984; Kolen ve Brennan, 2014). Bu bağlamda yapılacak ölçek dönüştürme çalışmalarında ortak madde oranı %40 ve daha fazla olacak şekilde analizler yürütülebilir.
- Bu çalışma kapsamında çok boyutlu ölçek dönüştürmede en düşük RMSE değerlerinin genel olarak boyutlar arası korelasyonun 0.1 (düşük) olduğu koşullarda, tek boyutlu ölçek dönüştürmede ise en düşük RMSE değerlerinin genel olarak boyutlar arası korelasyonun 0.9 (yüksek) olduğu koşullarda elde

edildiği bulgusuna ulaşılmıştır. Bu bağlamda, eşitlenecek olan test formlarının boyutluluk yapısının ve boyutlar arasındaki korelasyonun derecesi dikkate alınarak buna uygun analiz süreçlerinin yürütülmesi önerilmektedir. Eğitimde ve sosyal bilimlerde genellikle veri yapısının tek boyutlu olarak kabul edilip tek boyutlu analiz süreçleri yürütülmesinin yüksek hataya neden olacağı açıktır. Boyutlar arası korelasyonun değerine göre uygulanan ölçek dönüştürme yöntemlerinden kestirilen RMSE değerleri arasındaki istatistiksel olarak anlamlı bulunan farklar durumun önemini ortaya koymaktadır.

- Bu çalışmada eşitleme deseni olarak eşdeğer olmayan gruplar ortak test deseni seçildiğinden, gruplar arası yetenek dağılımı ortalamalarının farkı araştırma koşulu olarak ele alınmıştır. Hem tek boyutlu hem de çok boyutlu ölçek dönüştürme yöntemleri için çalışma kapsamında yer alan bütün koşullarda gruplar arası yetenek dağılımı yüksek olan koşullarda genel olarak daha yüksek RMSE değerleri elde edilmiştir. Bu bağlamda daha düşük eşitleme hataları kestirebilmek için gruplar arası yetenek dağılımı farkı düşük olan koşullar tercih edilebilir ya da eşitleme deseni olarak eşdeğer gruplar deseni seçilebilir.
- Bu araştırma kapsamında bazı koşullarda ve bazı yöntemlerde 3PLM kullanılarak kestirilen RMSE değerleri daha düşük değer üretirken, bazı koşullarda ise 2PLM kullanılarak kestirilen RMSE değerleri daha düşük bulunmuştur. Bununla ilgili net bir yargıya varmak güçtür. Şans başarısının parametre kestirimine ve eşitleme hatasına olan etkisini ortaya koymak isteyen bir araştırmacının 3PLM üzerinden analiz süreçlerini yürütmesi önerilebilir.

**Gelecek Araştırmalara Yönelik Öneriler.** Araştırma sonucu elde edilen bulgular doğrultusunda gelecekte yapılacak araştırmalara yönelik öneriler aşağıda maddeler halinde sıralanmıştır.

- Bu araştırma kapsamında simülasyon verileri kullanılmıştır. İleride yapılacak çalışmalarda gerçek veri seti üzerinden ölçek dönüştürme süreçleri yürütülebilir.

- Bu arařtırmada 6rneklem b6y6kl6đ6 1000 ve 2000 olarak alınmıř ve 6rneklem b6y6kl6đ6 2000 olan durumlarda genel olarak daha d6ř6k eřitleme hataları elde edilmiřtir. Bundan sonra yapılacak 7alıřmalarda daha b6y6k 6rneklem b6y6kl6kleri ele alınıp elde edilen eřitleme hataları karřılařtırılabilir.
- Bu 7alıřmada 6l7ek d6n6řt6rme y6ntemleri olarak ortalama-ortalama, ortalama-sigma ve Stocking-Lord kullanılarak analizler ger7ekleřtirilmiřtir. 7ok boyutlu veri yapısına farklı 6l7ek d6n6řt6rme ya da eřitleme y6ntemleri kullanılarak elde edilen sonu7lar karřılařtırılabilir.
- Bu 7alıřma, boyut sayısı 2 olacak řekilde basit yapılı veri seti ile y6r6t6lm6řt6r. Benzer 7alıřma 2'den fazla boyuttan oluřan veri setleri oluřturulabilir. Bununla birlikte karmařık yapılı ya da karıřık formatlı testler 6zerinden de 6l7ek d6n6řt6rme 7alıřmaları y6r6t6lebilir.
- Bu 7alıřmada deđerlendirme 6l76t6 olarak RMSE deđerini kullanılmıřtır. Bařka 7alıřmalarda RMSE dıřında bias, yanlılık, eřitlemenin standart hatası gibi deđerler de karřılařtırılarak oluřan farkların istatistiksel olarak anlamlı olup olmadıđı incelenebilir.

### Kaynaklar

- Ackerman, T. (1989). Unidimensional IRT calibration of compensatory and noncompensatory items. *Applied Psychological Measurement, 13*, 113-127.
- Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice, 22*(3), 37-51.
- Angoff, W. H. (1984). *Scales, norms and equivalent scores*. New Jersey: Educational Testing Service.
- Atar, B. ve Yeşiltaş, G. (2017). Çok boyutlu eşitleme yöntemlerinin eşdeğer olmayan gruplarda ortak madde deseni için performanslarının incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 8*(4), 421-434.
- Baker, F. B. & Al-Karni, A. (1991). A comparison of two procedures for computing IRT equating coefficients. *Journal of Educational Measurement, 28*(2), 147-162.
- Braun, H. I and Holland, P. W. (1982). Observed-score test equating: A mathematical analysis of some ETS equating procedures. In P. W. Holland and D.B. Rubin (Ed.), *Test equating* (s. 9-49). New York: Academic Press.
- Beguín, A. A., & Hanson, B. A. (2001). *Effect of noncompensatory multidimensionality on separate and concurrent estimation in IRT observed score equating*. Paper presented at the The Annual Meeting of the National Council on Measurement in Education, Seattle, WA.
- Bökeoğlu, Ö., Uçar, A. ve Balta, E. (2022). Madde tepki kuramına dayalı gerçek puan eşitlemede ölçek dönüştürme yöntemlerinin incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi. 55*(1), 1-36.
- Brossman, B. G. (2010). *Observed score and true score equating procedures for multidimensional item response theory* (Doctoral dissertation). University of Iowa, Iowa.

- Brossman, B. G., & Lee, W. (2013). Observed score and true score equating procedures for multidimensional item response theory. *Applied Psychological Measurement, 37*, 460-481.
- Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*, 1-29.
- Chu, K. L. & Kamata, A. (2003). *Test equating with the presence of DIF*. Paper presented at the annual meeting of American Educational Research Association, Chicago.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum
- Cook L. L. & Eignor R. E. (1991). An NCME instructional module on IRT equating methods. Instructional topics in educational measurement. *Educational Measurement: Issues and Practice, 10*(1), 37-45.
- Crocker, L., and Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: Guilford Press.
- DeMars, C. (2016). *Madde Tepki Kuramı*. Ankara: Nobel Akademi Yayıncılık.
- Dorans, N. J., & Holland, P. W. (2000). Population invariance and the equatability of tests: basic theory and the linear case. *Journal of Educational Measurement, 37*(4), 281-306.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. London: Lawrence Erlbaum Associates, Publishers.
- Gibbons, R. D., Immekus, J., ve Bock, R. D. (2007). *Didactic workbook: The added value of multidimensional IRT models*. National Cancer Institute Technical Report.

- Gök, B. ve Kelecioğlu, H. (2014). Denk olmayan gruplarda ortak madde deseni kullanılarak madde tepki kuramına dayalı eşitleme yöntemlerinin karşılaştırılması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 10(1), 120-136
- Gübeş, N. Ö. (2019). Test Eşitlemede Çok Boyutluluğun Eş Zamanlı ve Ayrı Kalibrasyona Etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34(4), 1061-1074.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hanson, B. A., & Beguin, A.A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common-item equating design. *Applied Psychological Measurement*, 26(1), 3-24.
- Huggins, A. C. (2012). *The effect of differential item functioning on population invariance of item response theory true score equating* (Doctoral dissertation). University of Miami, Coral Gables.
- Kabasakal, K. A. (2014). *Değişen Madde Fonksiyonunun Test Eşitlemeye Etkisi* (Doktora tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Kaskowitz, G. S., & De Ayala, R. J. (2001). The effect of error in item parameter estimates on the test response function method of linking. *Applied Psychological Measurement*, 25, 39-52.
- Kilmen, S. and Demirtaşlı, N. (2012). Comparison of test equating methods based on item response theory. *Procedia-Social and Behavioral Sciences* 46(2012), 130-134 according to the sample size and ability distribution
- Kim, S., & Kolen, M.J. (2006). Robustness to format effects of IRT linking methods for mixed-format tests. *Applied Measurement in Education*, 19(4), 357-381.



- Kim, S., & Lee, W. (2006). An extension of four IRT linking methods for mixed- format tests. *Journal of Educational Measurement, 43*(1), 53-76.
- Kim, S. Y. (2018). *Simple structure MIRT equating for multidimensional tests* (Doctoral Dissertation). University of Iowa, Iowa.
- Kim, S., Lee, W. C. ve Kolen, M. J. (2020). Simple-Structure Multidimensional Item Response Theory Equating for Multidimensional Tests. *Educational and Psychological Measurement. 80*(1), 91-125.
- Kim, S. & Lee, W. (2023). Several Variations of Simple-Structure MIRT Equating. *Journal of Educational Measurement*. <https://doi.org/10.1111/jedm.12341>
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices (3rd edition)*. New York: Springer.
- Kreiter, C.D. (1993). *An empirical investigation of compensatory and noncompensatory test items in simulated and real data* (Doctoral Dissertation). The University of Iowa, Iowa.
- Kumlu, G. (2019). *Test ve alt testlerde eşitlemenin farklı koşullar açısından incelenmesi* (Doktora tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Lee, E., Lee, W., & Brennan, R. L. (2014). *Equating multidimensional tests under a random groups design: A comparison of various equating procedures*. (CASMA Research Report No. 40). Iowa City, IA: Center for Advanced Studies in Measurement and Assessment, The University of Iowa.
- Livingston, S. A. (2004). *Equating test scores (Without IRT)* (2nd edition). Educational Testing Service.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Loyd, B. H., & Hoover, H. D. (1980). Vertical equating using the rasch model. *Journal of Educational Measurement*, 17(3), 179-193.
- Marco, G. L. (1977). Item characteristic curve solutions to three intractable testing problems. *Journal of Educational Measurement*, 14(2), 139-160.
- Ogasawara, H. (2000). Asymptotic standard errors of IRT equating coefficients using moments. *Economic Review (Otaru University of Commerce)*, 51(1), 1-23.
- Öztürk Gübeş, N. (2014). *Testlerin boyutluluğunun, ortak madde formatının, yetenek dağılımının ve ölçek dönüştürme yöntemlerinin karma testlerin eşitlenmesine etkisi* (Doktora tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement*, 9, 401-412.
- Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer.
- Skaggs, G., and Lissitz, R. W. (1986). IRT test equating: Relevant issues and a review of recent research. *Review of Educational Research*, 56(4), 495-529.
- Smith, J. (2009). *Some issues in item response theory: Dimensionality assessment and models for Guessing* (Unpublished Doctoral Dissertation). University of South California.
- Speron, E. (2009). *A comparison of metric linking procedures in Item Response Theory* (doctoral dissertation). University of Illinois, Chicago, Illinois.
- Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7, 201-210.
- Tian, F. (2011). *A comparison of equating/linking using the Stocking-Lord method and concurrent calibration with mixed-format tests in the non-equivalent groups common-item design under IRT* (Doctoral dissertation). Boston College University, Boston.

TÜBİTAK (2002). *Araştırma ve Deneysel Geliştirme Taramaları İçin Önerilen Standart Uygulama.*

[https://www.tubitak.gov.tr/tubitak\\_content\\_files/BTYPD/kilavuzlar/frascati\\_tr.pdf](https://www.tubitak.gov.tr/tubitak_content_files/BTYPD/kilavuzlar/frascati_tr.pdf)

adresinden erişilmiştir.

Wang, T., Lee, W. C., Brennan, R. J., & Kolen, M. J. (2008). A comparison of the frequency estimation and chained equipercentile methods under the common-item non-equivalent groups design. *Applied Psychological Measurement, 32*, 632-651.

Yao, L. (2009). *LinkMIRT: Linking of multivariate item response model*. Monterey, CA: Defense Manpower Data Center.

Yao, L., & Boughton, K. A. (2009). Multidimensional Linking for Tests with Mixed Item Types. *Journal of Educational Measurement, 46*(2), 177-197.

Xu, Y. (2009). *Measuring change in jurisdiction achievement over time: Equating issues in current international assessment programs* (Doctoral dissertation). University of Toronto, Toronto.

Zhang, B. (2009). Application of unidimensional item response models to tests with item sensitive to secondary dimensions. *The Journal of Experimental Education, 77*(2), 147-166.

Zhang, J. (2012). Calibration of response data using MIRT models with simple and mixed structures. *Applied Psychological Measurement, 36*(5), 375-398.

Zhu, W. (1998). Test equating: What, why, who?. *Research Quarterly for Exercise and Sport, 69*(1), 11-23.

## EK-A: Araştırmada Yer Alan Koşullara İlişkin Elde Edilen RMSE Değerleri

ÖLÇ	ÖRN	KOR	ORT	YET	PAR	a			b			yetenek		
						OO	OS	SL	OO	OS	SL	OO	OS	SL
ÇOK B.	1000	0.1	%20	0.05	3PLM	0.12	0.43	0.18	0.21	0.38	0.26	0.15	0.25	0.21
ÇOK B.	1000	0.5	%20	0.05	3PLM	0.14	0.45	0.21	0.22	0.38	0.33	0.15	0.25	0.27
ÇOK B.	1000	0.9	%20	0.05	3PLM	0.12	0.46	0.20	0.21	0.39	0.30	0.14	0.25	0.24
ÇOK B.	2000	0.1	%20	0.05	3PLM	0.11	0.31	0.19	0.15	0.25	0.32	0.10	0.17	0.27
ÇOK B.	2000	0.5	%20	0.05	3PLM	0.10	0.41	0.17	0.17	0.30	0.25	0.12	0.21	0.20
ÇOK B.	2000	0.9	%20	0.05	3PLM	0.10	0.32	0.17	0.15	0.28	0.28	0.11	0.19	0.23
ÇOK B.	1000	0.1	%40	0.05	3PLM	0.11	0.33	0.13	0.14	0.23	0.21	0.10	0.17	0.16
ÇOK B.	1000	0.5	%40	0.05	3PLM	0.11	0.28	0.14	0.13	0.24	0.23	0.09	0.17	0.17
ÇOK B.	1000	0.9	%40	0.05	3PLM	0.09	0.29	0.13	0.13	0.27	0.23	0.09	0.19	0.18
ÇOK B.	2000	0.1	%40	0.05	3PLM	0.09	0.22	0.12	0.10	0.21	0.26	0.07	0.15	0.20
ÇOK B.	2000	0.5	%40	0.05	3PLM	0.08	0.23	0.11	0.10	0.17	0.20	0.07	0.13	0.15
ÇOK B.	2000	0.9	%40	0.05	3PLM	0.08	0.25	0.12	0.10	0.23	0.24	0.07	0.15	0.19
ÇOK B.	1000	0.1	%20	0.05	2PLM	0.07	0.32	0.09	0.25	0.37	0.17	0.18	0.27	0.12
ÇOK B.	1000	0.5	%20	0.05	2PLM	0.08	0.39	0.10	0.28	0.40	0.18	0.19	0.28	0.13
ÇOK B.	1000	0.9	%20	0.05	2PLM	0.08	0.45	0.10	0.32	0.38	0.19	0.22	0.28	0.15
ÇOK B.	2000	0.1	%20	0.05	2PLM	0.06	0.23	0.07	0.14	0.22	0.13	0.10	0.16	0.10
ÇOK B.	2000	0.5	%20	0.05	2PLM	0.05	0.25	0.07	0.19	0.27	0.15	0.13	0.20	0.11
ÇOK B.	2000	0.9	%20	0.05	2PLM	0.05	0.23	0.06	0.16	0.25	0.14	0.11	0.18	0.10
ÇOK B.	1000	0.1	%40	0.05	2PLM	0.05	0.29	0.06	0.17	0.30	0.14	0.12	0.22	0.11
ÇOK B.	1000	0.5	%40	0.05	2PLM	0.09	0.33	0.19	0.22	0.33	0.35	0.16	0.23	0.20
ÇOK B.	1000	0.9	%40	0.05	2PLM	0.09	0.29	0.19	0.22	0.40	0.35	0.16	0.26	0.21
ÇOK B.	2000	0.1	%40	0.05	2PLM	0.07	0.20	0.17	0.15	0.25	0.34	0.11	0.18	0.19
ÇOK B.	2000	0.5	%40	0.05	2PLM	0.08	0.22	0.17	0.17	0.27	0.34	0.12	0.19	0.20
ÇOK B.	2000	0.9	%40	0.05	2PLM	0.08	0.23	0.17	0.16	0.22	0.34	0.10	0.16	0.20
ÇOK B.	1000	0.1	%20	0.5	3PLM	0.12	0.42	0.16	0.35	0.61	0.40	0.29	0.40	0.34
ÇOK B.	1000	0.5	%20	0.5	3PLM	0.13	0.43	0.17	0.33	0.62	0.43	0.29	0.40	0.38
ÇOK B.	1000	0.9	%20	0.5	3PLM	0.12	0.46	0.16	0.34	0.63	0.40	0.28	0.40	0.35
ÇOK B.	2000	0.1	%20	0.5	3PLM	0.10	0.36	0.14	0.29	0.51	0.43	0.27	0.35	0.40
ÇOK B.	2000	0.5	%20	0.5	3PLM	0.10	0.39	0.14	0.33	0.50	0.37	0.28	0.35	0.33
ÇOK B.	2000	0.9	%20	0.5	3PLM	0.09	0.39	0.13	0.31	0.54	0.42	0.28	0.36	0.38
ÇOK B.	1000	0.1	%40	0.5	3PLM	0.10	0.34	0.15	0.34	0.46	0.43	0.30	0.34	0.38
ÇOK B.	1000	0.5	%40	0.5	3PLM	0.09	0.31	0.12	0.29	0.46	0.43	0.28	0.33	0.39
ÇOK B.	1000	0.9	%40	0.5	3PLM	0.10	0.31	0.12	0.30	0.41	0.43	0.27	0.31	0.39
ÇOK B.	2000	0.1	%40	0.5	3PLM	0.08	0.26	0.10	0.27	0.38	0.44	0.27	0.29	0.42
ÇOK B.	2000	0.5	%40	0.5	3PLM	0.07	0.27	0.12	0.29	0.37	0.39	0.27	0.29	0.35
ÇOK B.	2000	0.9	%40	0.5	3PLM	0.07	0.27	0.10	0.28	0.38	0.43	0.27	0.29	0.40
ÇOK B.	1000	0.1	%20	0.5	2PLM	0.10	0.38	0.10	0.47	0.50	0.40	0.35	0.37	0.42
ÇOK B.	1000	0.5	%20	0.5	2PLM	0.11	0.49	0.11	0.50	0.52	0.40	0.36	0.38	0.41
ÇOK B.	1000	0.9	%20	0.5	2PLM	0.12	0.54	0.11	0.56	0.50	0.39	0.40	0.36	0.42
ÇOK B.	2000	0.1	%20	0.5	2PLM	0.10	0.32	0.08	0.43	0.40	0.40	0.32	0.33	0.43
ÇOK B.	2000	0.5	%20	0.5	2PLM	0.10	0.34	0.08	0.46	0.42	0.40	0.34	0.33	0.43
ÇOK B.	2000	0.9	%20	0.5	2PLM	0.10	0.35	0.08	0.45	0.42	0.40	0.34	0.33	0.43
ÇOK B.	1000	0.1	%40	0.5	2PLM	0.10	0.45	0.09	0.49	0.37	0.42	0.38	0.41	0.47
ÇOK B.	1000	0.5	%40	0.5	2PLM	0.16	0.47	0.22	0.55	0.36	0.37	0.39	0.39	0.41
ÇOK B.	1000	0.9	%40	0.5	2PLM	0.16	0.45	0.23	0.51	0.34	0.37	0.36	0.38	0.41
ÇOK B.	2000	0.1	%40	0.5	2PLM	0.15	0.38	0.22	0.50	0.30	0.36	0.35	0.38	0.41
ÇOK B.	2000	0.5	%40	0.5	2PLM	0.14	0.39	0.22	0.50	0.31	0.36	0.36	0.38	0.41
ÇOK B.	2000	0.9	%40	0.5	2PLM	0.14	0.44	0.22	0.52	0.32	0.36	0.38	0.40	0.42
TEK B.	1000	0.1	%20	0.05	3PLM	0.11	0.29	0.09	0.53	0.76	0.26	0.38	0.40	0.15
TEK B.	1000	0.5	%20	0.05	3PLM	0.10	0.21	0.06	0.18	0.25	0.09	0.15	0.21	0.08
TEK B.	1000	0.9	%20	0.05	3PLM	0.10	0.22	0.05	0.15	0.19	0.07	0.13	0.18	0.07
TEK B.	2000	0.1	%20	0.05	3PLM	0.14	0.28	0.08	0.55	0.82	0.23	0.39	0.41	0.13
TEK B.	2000	0.5	%20	0.05	3PLM	0.09	0.17	0.03	0.16	0.20	0.05	0.14	0.18	0.05
TEK B.	2000	0.9	%20	0.05	3PLM	0.08	0.14	0.03	0.10	0.12	0.04	0.09	0.12	0.04
TEK B.	1000	0.1	%40	0.05	3PLM	0.14	0.25	0.07	0.54	0.75	0.15	0.32	0.43	0.10
TEK B.	1000	0.5	%40	0.05	3PLM	0.11	0.13	0.04	0.14	0.16	0.06	0.09	0.12	0.05
TEK B.	1000	0.9	%40	0.05	3PLM	0.09	0.13	0.03	0.11	0.12	0.04	0.07	0.10	0.05
TEK B.	2000	0.1	%40	0.05	3PLM	0.17	0.28	0.11	0.51	0.75	0.15	0.30	0.45	0.12
TEK B.	2000	0.5	%40	0.05	3PLM	0.10	0.13	0.03	0.13	0.16	0.04	0.09	0.12	0.04
TEK B.	2000	0.9	%40	0.05	3PLM	0.09	0.10	0.02	0.09	0.10	0.03	0.06	0.08	0.03
TEK B.	1000	0.1	%20	0.05	2PLM	0.05	0.25	0.06	0.58	1.05	0.31	0.44	0.60	0.18

TEK B.	1000	0.5	%20	0.05	2PLM	0.06	0.16	0.04	0.15	0.21	0.08	0.12	0.20	0.07
TEK B.	1000	0.9	%20	0.05	2PLM	0.06	0.15	0.04	0.11	0.16	0.06	0.10	0.18	0.07
TEK B.	2000	0.1	%20	0.05	2PLM	0.04	0.31	0.07	0.79	1.14	0.52	0.64	0.64	0.36
TEK B.	2000	0.5	%20	0.05	2PLM	0.04	0.11	0.02	0.10	0.15	0.04	0.08	0.16	0.04
TEK B.	2000	0.9	%20	0.05	2PLM	0.04	0.09	0.02	0.07	0.10	0.04	0.06	0.11	0.04
TEK B.	1000	0.1	%40	0.05	2PLM	0.06	0.17	0.06	0.70	1.04	0.29	0.39	0.59	0.15
TEK B.	1000	0.5	%40	0.05	2PLM	0.06	0.10	0.02	0.13	0.18	0.05	0.08	0.16	0.05
TEK B.	1000	0.9	%40	0.05	2PLM	0.06	0.10	0.02	0.10	0.13	0.04	0.07	0.12	0.04
TEK B.	2000	0.1	%40	0.05	2PLM	0.04	0.16	0.03	0.59	0.92	0.16	0.35	0.50	0.10
TEK B.	2000	0.5	%40	0.05	2PLM	0.04	0.08	0.01	0.10	0.15	0.03	0.06	0.14	0.03
TEK B.	2000	0.9	%40	0.05	2PLM	0.04	0.07	0.01	0.06	0.08	0.03	0.04	0.08	0.03
TEK B.	1000	0.1	%20	0.5	3PLM	0.11	0.27	0.11	1.10	1.28	0.81	1.02	1.11	0.78
TEK B.	1000	0.5	%20	0.5	3PLM	0.09	0.18	0.05	0.58	0.62	0.60	0.57	0.60	0.59
TEK B.	1000	0.9	%20	0.5	3PLM	0.10	0.18	0.05	0.51	0.54	0.52	0.52	0.54	0.53
TEK B.	2000	0.1	%20	0.5	3PLM	0.13	0.23	0.08	1.09	1.18	0.79	1.02	1.08	0.79
TEK B.	2000	0.5	%20	0.5	3PLM	0.08	0.16	0.04	0.57	0.59	0.58	0.56	0.59	0.57
TEK B.	2000	0.9	%20	0.5	3PLM	0.07	0.13	0.03	0.53	0.54	0.53	0.53	0.54	0.52
TEK B.	1000	0.1	%40	0.5	3PLM	0.09	0.27	0.10	1.26	1.46	0.84	1.16	1.19	0.82
TEK B.	1000	0.5	%40	0.5	3PLM	0.11	0.13	0.05	0.63	0.63	0.61	0.60	0.63	0.59
TEK B.	1000	0.9	%40	0.5	3PLM	0.10	0.12	0.04	0.53	0.54	0.52	0.53	0.55	0.53
TEK B.	2000	0.1	%40	0.5	3PLM	0.08	0.18	0.05	1.11	1.19	0.75	1.00	1.13	0.73
TEK B.	2000	0.5	%40	0.5	3PLM	0.10	0.12	0.05	0.62	0.62	0.59	0.60	0.63	0.57
TEK B.	2000	0.9	%40	0.5	3PLM	0.08	0.09	0.03	0.54	0.54	0.53	0.53	0.54	0.52
TEK B.	1000	0.1	%20	0.5	2PLM	0.08	0.25	0.11	0.92	1.26	0.69	0.93	1.12	0.79
TEK B.	1000	0.5	%20	0.5	2PLM	0.08	0.14	0.06	0.56	0.59	0.54	0.64	0.65	0.62
TEK B.	1000	0.9	%20	0.5	2PLM	0.11	0.14	0.07	0.49	0.51	0.47	0.58	0.58	0.55
TEK B.	2000	0.1	%20	0.5	2PLM	0.08	0.25	0.14	1.11	1.43	0.81	1.16	1.25	0.93
TEK B.	2000	0.5	%20	0.5	2PLM	0.08	0.11	0.06	0.56	0.57	0.53	0.64	0.65	0.61
TEK B.	2000	0.9	%20	0.5	2PLM	0.09	0.12	0.06	0.50	0.51	0.48	0.57	0.58	0.55
TEK B.	1000	0.1	%40	0.5	2PLM	0.08	0.19	0.15	1.36	1.64	0.92	1.27	1.47	0.97
TEK B.	1000	0.5	%40	0.5	2PLM	0.09	0.12	0.05	0.60	0.62	0.53	0.68	0.73	0.61
TEK B.	1000	0.9	%40	0.5	2PLM	0.11	0.14	0.07	0.51	0.51	0.46	0.59	0.63	0.55
TEK B.	2000	0.1	%40	0.5	2PLM	0.09	0.18	0.12	1.09	1.36	0.72	1.07	1.38	0.86
TEK B.	2000	0.5	%40	0.5	2PLM	0.08	0.12	0.05	0.60	0.61	0.52	0.67	0.76	0.60
TEK B.	2000	0.9	%40	0.5	2PLM	0.09	0.12	0.06	0.52	0.53	0.47	0.60	0.64	0.55

Ölç: Ölçek dönüştürme yöntemi, Örn: Örneklem büyüklüğü, Kor: Boyutlar arası korelasyon,  
Ort: Ortak madde oranı, Yet: Gruplar arası yetenek dağılımı farkı, Par: Parametre Kestirim Modeli,  
OO: Ortalama-ortalama, OS: Ortalama-sigma, SL: Stocking- Lord

**EK-B: Arařtırma Etik Komisyon İzin Muafiyeti Formu/ Arařtırma Etik Komisyonu Onay  
Bildirimi**



T.C.  
**HACETTEPE ÜNİVERSİTESİ**  
Rektörlük

Tarih: 02/06/2020 13:05  
Sayı: E-35853172-755.02.06  
00001097842



00001097842

Sayı : 35853172-755.02.06  
Konu : Yaşar Mehmet ZOR (Etik Komisyon İzni)

**EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE**

İlgi: 21.04.2020 tarih ve 1080389 sayılı yazınız.

Enstitünüz Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı Doktora programı öğrencisi Yaşar Mehmet ZOR'un Prof. Dr. Hülya KELECİOĞLU danışmanlığında yürüttüğü "Çok Boyutlu Testlerin Tek Boyutlu ve Çok Boyutlu Yöntemlere Göre Eşitlenmesi" başlıklı tez çalışması Üniversitemiz Senatosu Etik Komisyonunun 12 Mayıs 2020 tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini saygılarımla rica ederim.

e-imzalıdır  
Prof. Dr. Rahime Meral NOHUTCU  
Rektör Yardımcısı

**EK-C: Etik Beyanı**

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

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

beyan ederim.

31/07/2023

Yaşar Mehmet ZOR

**EK-Ç: Doktora Tez Çalışması Orijinallik Raporu**

31/07/2023

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

Tez Başlığı: Çok Boyutlu Testlerin Tek Boyutlu ve Çok Boyutlu Yöntemlere Göre Eşitlenmesi

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
31/07/2023	91	116,284	13/07/2023	%6	2139479461

Uygulanan filtreler:

- Kaynaklar hariç
- Alıntılar dâhil
- %1'den daha az örtüşme içeren metin kısımları hariç

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

**Ad Soyadı:** Yaşar Mehmet ZOR

**Öğrenci No.:** N15245021

**Ana Bilim Dalı:** Eğitim Bilimleri

**Programı:** Eğitimde Ölçme ve Değerlendirme

**Statüsü:**  Y.Lisans  Doktora  Bütünleşik Dr.

**DANIŞMAN ONAYI**

UYGUNDUR.

(Prof. Dr. Hülya KELECİOĞLU)



**EK-D: Dissertation Originality Report**

31/07/2023

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

Thesis Title: Equating Multidimensional Tests According to Unidimensional And Multidimensional Methods

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
31/07/2023	91	116,284	13/07/2023	%6	2139479461

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to %1 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:** Yaşar Mehmet ZOR  
**Student No.:** N15245021  
**Department:** Educational Sciences  
**Program:** Educational Measurement and Evaluation  
**Status:**  Masters  Ph.D.  Integrated Ph.D.

**ADVISOR APPROVAL**

APPROVED  
(Prof. Dr. Hülya KELECİOĞLU)

## EK-E: Yayınlama ve Fikrî Mülkiyet Hakları Beyanı

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

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

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

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

04 /08 /2023

Yaşar Mehmet ZOR

---

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

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

