

**ÇOK BOYUTLU MADDE TEPKİ KURAMININ FARKLI
MODELLERİNDEN ÇEŞİTLİ KOŞULLAR ALTINDA
KESTİRİLEN PARAMETRELERİN İNCELENMESİ**

**INVESTIGATION OF PARAMETERS ESTIMATED UNDER
VARIOUS CONDITIONS BY DIFFERENT
MULTIDIMENSIONAL ITEM RESPONSE THEORY
MODELS**

Derya ÇAKICI ESER

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı İçin
Öngördüğü

Doktora Tezi olarak hazırlanmıştır.

2015

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne,

Derya AKICI ESER'in hazırladıđı "ok Boyutlu Madde Tepki Kuramının Farklı Modellerinden eřitli Kořullar Altında Kestirilen Parametrelerin İncelenmesi" bařlıklı bu alıřma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde Ölme ve Deđerlendirme Bilim Dalı'nda Doktora Tezi** olarak kabul edilmiřtir.

<i>Bařkan</i>	Prof. Dr. řener B¼Y¼KÖZT¼RK	_____
<i>¼ye (Danıřman)</i>	Prof. Dr. Selahattin GELBAL	_____
<i>¼ye</i>	Prof. Dr. H¼lya KELECİOđLU	_____
<i>¼ye</i>	Prof. Dr. Mehtap AKAN	_____
<i>¼ye</i>	Do. Dr. Burcu ATAR	_____

ONAY

Bu tez Hacettepe ¼niversitesi Lisans¼st¼ Eđitim-Öđretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca řekle j¼ri ¼yeleri tarafından 18 / 11 / 2014 tarihinde uygun gör¼lm¼ř ve Enstit¼ Yönetim Kurulunca / / tarihinde kabul edilmiřtir.

Prof. Dr. Berrin AKMAN
Eđitim Bilimleri Enstitüsü M¼d¼r¼

ÇOK BOYUTLU MADDE TEPKİ KURAMININ FARKLI MODELLERİNDEN ÇEŞİTLİ KOŞULLAR ALTINDA KESTİRİLEN PARAMETRELERİN İNCELENMESİ

Derya ÇAKICI ESER

ÖZ

Bu çalışmada farklı madde sayısı ve örneklem büyüklüğü koşullarına sahip değişen boyutluluk özelliklerindeki veri setlerinin madde ve birey parametreleri tek boyutlu ve çok boyutlu madde tepki kuramı ile kestirilmiş; kestirime ilişkin olarak elde edilen standart hata ortalamalarının karekökü (RMSE), yanlılık ve kestirilen parametreler ile gerçek parametreler arasındaki korelasyon değerleri incelenmiştir. Sonuçta daha kararlı kestirim yapabilmek için gerekli test uzunluğu ve örneklem büyüklüğüne ilişkin öneriler geliştirilmeye çalışılmıştır.

Çalışmanın verileri simülasyon ile bir boyutlu, iki boyutlu karmaşık, iki boyutlu basit, üç boyutlu karmaşık ve üç boyutlu basit yapıda olmak üzere beş farklı yapı gösterecek biçimde üretilmiştir. Veri setlerini üretmek için önce madde parametreleri üretilmiştir. Madde parametrelerinin elde edilmesinde a parametresi için ortalaması 0,5 ve standart sapması 0,4 olan lognormal dağılımdan; d parametresi için standart normal dağılımdan faydalanılmıştır. Elde edilen madde parametrelerinden faydalanarak iki kategorili veri setleri ve yetenek parametreleri üretilmiştir. Yetenek parametrelerinin üretilmesinde standart normal dağılımdan faydalanılmıştır. Veri setlerinin üretiminde test uzunluğu 12 ve 48 madde; örneklem büyüklüğü 1000, 2000 ve 4000 olacak şekilde değiştirilmiştir. Bu şekilde ele alınan beş yapı ile 2 test uzunluğu ve 3 örneklem büyüklüğü koşulu çaprazlanarak 30 desen oluşturulmuştur. Oluşturulan her desene ilişkin olarak, madde ve birey parametreleri sabit tutularak 25 tekrar yapılmıştır. Sonuçta 30 x 25 olmak üzere toplam 750 veri seti elde edilmiştir. Her bir veri seti sahip olduğu yapıya uygun betikle parametreleri kestirmek üzere IRTPRO 2.1 yazılımının BA-EM algoritması ile analiz edilmiştir.

Elde edilen sonuçlara göre madde parametrelerine ait RMSE değerlerinin ve korelasyonların hem madde sayısındaki hem de örneklem büyüklüğündeki değişimden etkilendiği görülmüştür. Buna göre, test uzunluğunun ve birey sayısının

artması madde parametrelerinin daha düşük hatalı kestirilmesini, gerçek parametrelerle daha yüksek korelasyon vermesini sağlamıştır. Ele alınan tüm koşullarda d parametreleri aynı desendeki a parametrelerinden daha kararlı kestirilmiştir. Birey parametresi kestirimleri örneklem büyüklüğünün değişiminden etkilenmezken, madde sayısının artışıyla etkilenmiştir. Madde sayısının artması birey parametreleri kestirimlerine ait RMSE değerlerinin düşmesini, korelasyon değerlerinin artmasını sağlamıştır. Ancak madde sayısı sabit tutularak örneklem büyüklüğünün arttığı koşullarda birey parametrelerine ilişkin RMSE ve korelasyon değerlerinde önemli bir değişiklik gözlenmemiştir. Ele alınan istatistiklerden RMSE ve korelasyonlar aynı anlama gelen sonuçlar oluşturmuş, RMSE'nin düşük olduğu koşullarda korelasyonlar yüksek değerler almıştır. Buna karşılık, yanlılıklar RMSE veya korelasyonlardan bağımsız değişim göstermiş, madde ve birey sayısının artmasıyla düzenli bir artış ya da azalma göstermemiştir. Tek boyutlu yapılarda çok boyutlu yapılardan daha kararlı kestirimler yapılmıştır. Tek boyutlu testlerden elde edilen yanlılık değerleri sifıra yakındır. Madde parametreleri çok boyutlu basit yapılarda çok boyutlu karmaşık yapılardan daha düşük hata ile kestirilmiştir. İki boyutlu karmaşık yapılarda birey parametreleri basit yapılardan daha düşük RMSE değerleri almıştır. Yanlılıklar iki boyutlu basit ve karmaşık yapılarda benzer değerlerde olmakla beraber 0,1'in altındadır. İki boyutlu yapılardan kestirilen birey parametreleri ile gerçek parametrelerin korelasyonları kısa testlerde orta, uzun testlerde iyi derecede ilişkiye işaret etmektedir. Üç boyutlu yapılarda hem madde hem de birey parametresi kestirimleri iki boyutlu yapılardan daha düşük kararlılıkla yapılmıştır. Üç boyutlu karmaşık yapının kısa test koşulunda madde parametreleri için oldukça yüksek ($>1,0$) RMSE değerleri hesaplanmıştır. Üç boyutlu karmaşık yapının kısa test koşulunda madde parametreleri için elde edilen korelasyonlar düşük değerlerdedir. Üç boyutlu karmaşık yapıda üç boyutlu basit yapıdan daha yanlı kestirimler yapılmıştır.

Elde edilen sonuçlardan yola çıkarak araştırmacılara boyut sayısı arttıkça madde parametresi kestirimlerinin iyileştirilmesi için madde ve birey; birey parametresi kestirimlerinin iyileştirilmesi için madde sayısını arttırmaları önerilmiştir. RMSE değerleri göz önünde bulundurularak madde parametresi kestirimlerini iyileştirmek için tek boyutlu testlerde 12 madde ve 1000 kişilik örneklem büyüklüğü, iki boyutlu karmaşık yapılarda 48 madde ve 1000 kişilik örneklem büyüklüğü, iki boyutlu basit

yapılı testlerde 12 madde ve 2000 kişilik örneklem büyüklüğü önerilmiştir. Üç boyutlu testlerle çalışırken kararlı kestirimler elde etmek için karmaşık yapılarda en az 48 madde ve 4000'den büyük örneklem büyüklüğü, basit yapılarda ise 12 madde ve 4000 örneklem büyüklüğü ile çalışılması önerilmiştir. Korelasyon değerlerini göz önünde bulundurarak; birey parametresi kestirimlerini iyileştirmek için tek boyutlu yapılar, iki boyutlu (basit ve karmaşık) yapılar ve üç boyutlu basit yapılar için 12 maddelik, üç boyutlu karmaşık yapılar için 48 maddelik testlerin kullanılması önerilmiştir.

Anahtar sözcükler: Çok boyutlu madde tepki kuramı, madde tepki kuramı, parametre kestirimi

Danışman: Prof. Dr. Selahattin GELBAL, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

INVESTIGATION OF PARAMETERS ESTIMATED UNDER VARIOUS CONDITIONS BY DIFFERENT MULTIDIMENSIONAL ITEM RESPONSE THEORY MODELS

Derya ÇAKICI ESER

ABSTRACT

In this study, the item and ability parameters of the data sets which have changing dimensionality features with different item numbers and sampling size were estimated with the item response theory and multidimensional item response theory; related to these estimations root mean square error (RMSE), bias, the correlation values between estimated and real parameters were examined. At the end, in order to have accurate estimations, the suggestions about the required test length and sample size were developed.

The data of this study were generated in a way to show five different structures being one dimensional, two dimensional complex, two dimensional simple, three dimensional complex and three dimensional simple structure by means of simulation. To generate data sets, firstly item parameters were developed. In order to obtain item parameters, lognormal distribution having 0.5 mean and 0.4 standard deviation value were used for the parameter a ; standard normal distribution was used for the parameter d . Using the obtained item parameters, two-categorized data sets and ability parameters were generated. The standard normal distribution was used for the generation of ability parameters. In the generation of the data sets, test length was changed to include 12 and 48 items and the sample size was changed to be 1000, 2000 and 4000. Thus 2 test length and 3 sampling size conditions were intercrossed with five structures and a total of 30 patterns were generated. Each generated pattern was replicated for 25 times through fixed item and ability parameters. As a result, a total of 750 data sets (30×25) were obtained. Each data set was analyzed using suitable syntax for the structure with BA-EM algorithm of IRTPRO 2.1 software to estimate its parameters.

According to the obtained results, it was observed that RMSE values and correlations of the item parameters were affected by the changes in both item number and sample size. Accordingly, the increase in test length and examinee

number led to the estimation of item parameters with lower error and resulted in higher correlations with real parameters. In all examined conditions, the parameters d were estimated more consistently than the parameters a in the same design. The ability parameter estimations were not affected by the changes in the sample size while they were influenced by increase in the number of items. The increase in item number led to decrease in RMSE values of ability parameter estimations and increase in correlation values. However, no significant difference was observed in RMSE and correlation values of ability parameters under conditions which item number was fixed and sample size increased. Of the examined statistics, RMSE and correlations gave results that mean the same; in the conditions with low RMSE values, high correlation values were observed. However, biases were observed to change independently from RMSE and correlation values; they did not generate regular increase or decrease with the increase in the item number and the number of examinees. In the one dimensional structures, more accurate estimations were made than the multidimensional structures. Bias values obtained from one dimensional tests are close to zero. Item parameters were estimated with lower error in multidimensional simple structures compared to multidimensional complex structures. In the two dimensional complex structures, ability parameters had lower RMSE values than they did in the simple structures. Bias values were similar in both two dimensional simple and complex structures and these values were below 0.1. In two dimensional structures, correlations of estimated ability parameters and real parameters refer to medium level relationship in short tests; good level in long tests. In three dimensional structures, estimations of both item and ability parameters were made less accurately compared to two-dimensional structures. In short test condition of three-dimensional complex structure, very high (>1.0) RMSE values were calculated. In short test condition of the three dimensional complex structure, correlations obtained for item parameters had lower values. In three-dimensional complex structure, more biased estimations were made compared to the three-dimensional simple structure.

Based on the obtained results, researchers were suggested to increase the number of items and examinees for item parameter recovery, and they were suggested to increase the number of items for ability parameter recovery, as the number of dimensions increases. Considering the RMSE values; for one dimensional tests 12

items and 1000 sample size; for two dimensional complex structures 48 items and 1000 sample size and for two dimensional simple structures 12 items and 2000 sample size was suggested in order to improve parameter recovery. In order to have accurate estimations while studying with three dimensional tests it was suggested to conduct study with minimum 48 items and more than 4000 sample size in complex structures; 12 items and 4000 sample size in simple structures. Considering the correlation values, it was suggested that 12-item tests should be used for one-dimensional, two-dimensional (simple and complex) and three-dimensional structures; 48-item tests should be used for three-dimensional complex structures in order to improve ability parameter estimations.

Keywords: Multidimensional item response theory, item response theory, parameter estimation.

Advisor: Prof. Dr. Selahattin GELBAL, Hacettepe University, Department of Educational Sciences, Division of Educational Measurement and Evaluation

ETİK BEYANNAMESİ

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

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

beyan ederim.

Derya ÇAKICI ESER

TEŞEKKÜR

Ölçme ve Değerlendirme bilim dalında eğitime başladığım günden bu yana bilgisini ve emeğini esirgmeden önerileri ve fikirleriyle beni yönlendiren, tez sürecindeki zor zamanlarımda sorularıma ve karşılaştığım sorunlara önerdiği çözümler ile yol gösteren, çalışmamın tamamlanmasında büyük katkıları olan değerli hocam ve danışmanım Prof. Dr. Selahattin GELBAL'a içten teşekkürlerimi sunarım.

Yüksek lisans ve doktora eğitimim boyunca aldığım dersler aracılığıyla bilgisinden ve deneyiminden faydalandığım, bugüne gelmemde büyük katkıları olan değerli hocam Prof. Dr. Hülya KELECİOĞLU'na teşekkür ederim.

Tez jürimde yer alarak değerli görüş ve önerilerini esirgemeyen, tezimin son halini almasında katkıları bulunan Prof. Dr. Şener BÜYÜKÖZTÜRK, Prof. Dr. Mehtap ÇAKAN ve Doç. Dr. Burcu ATAR'a teşekkürlerimi sunarım.

Tezimin analizlerini yaparken karşılaştığım sorunlarla ilgili tüm sorularıma içtenlikle cevap veren, uzmanlığından ve bilgisinden faydalandığım değerli Prof. Dr. Terry ACKERMAN'a teşekkür ederim.

Tez süreci boyunca Skype aracılığıyla karşılıklı ders çalıştığım, moralimin bozuk olduğu zamanlarda beni destekleyen, tezimdiki her nokta ile tek tek ilgilenen değerli arkadaşım Sakine GÖÇER ŞAHİN'e teşekkür ederim.

Tezimi tamamlayabilmem için bana destek olan annem Zahide ÇAKICI, babam Şenol ÇAKICI ve ablam Pınar ÇAKICI SOYBEKAR başta olmak üzere tüm aileme teşekkür ederim.

Akademik hayatım boyunca beni destekleyen, cesaretlendiren, tezimin her aşamasıyla ilgilenip zor zamanlarımda bana olan inancıyla güç veren, yapabileceği hiçbir şeyi esirgemeyip yanımda olan, yol arkadaşım, sevgili eşim, Hasan ESER'e sonsuz teşekkürlerimi sunarım.

Eşim Hasan'a

İÇİNDEKİLER

ÖZ	iii
ABSTRACT	vi
ETİK BEYANNAMESİ.....	ix
TEŞEKKÜR	x
İÇİNDEKİLER.....	xii
TABLolar DİZİNİ	xiv
ŞEKİLLER DİZİNİ	xv
SİMGELEr VE KISALTMALAR DİZİNİ	xvii
1. GİRİŞ	1
1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı ve Önemi:	21
1.3. Problem Cümlesi:.....	21
1.3.1. Alt Problemler:	22
1.4. Sayıtlılar:	23
1.5. Sınırlılıklar:.....	23
1.6. Tanımlar:.....	23
2. İLGİLİ ARAŞTIRMALAR	24
3. YÖNTEM	31
3.1. Araştırmanın Türü	31
3.2. Araştırmanın Deseni	31
3.3. Verilerin Analizi	34
4. BULGULAR VE TARTIŞMA	37
4.1. Birinci alt problem için bulgular.....	37
4.1.1. Tek boyutlu testlere ilişkin bulgular	37
4.2. İkinci alt problem için bulgular	42
4.2.1. İki boyutlu karmaşık yapıli testlere ilişkin bulgular	42
4.2.2. İki boyutlu basit yapıli testlere ilişkin bulgular	48
4.3. Üçüncü alt problem için bulgular	53
4.3.1. Üç boyutlu karmaşık yapıli testlere ilişkin bulgular.....	53
4.3.2. Üç boyutlu basit yapıli testlere ilişkin bulgular.....	59
4.4. Tartışma.....	64
4.4.1. Madde parametrelerine ilişkin tartışma	64
4.4.2. Birey parametrelerine ilişkin tartışma	67
5. SONUÇ ve ÖNERİLER	69
5.1. Sonuçlar	69
5.1.1. Birinci Alt Probleme İlişkin Sonuçlar	69
5.1.2. İkinci Alt Probleme İlişkin Sonuçlar	69
5.1.3. Üçüncü Alt Probleme İlişkin Sonuçlar	70
5.2. Öneriler	71

5.2.1. Arařtırmacılara Yönelik Öneriler	71
5.2.2. İleri Arařtırmalara Yönelik Öneriler	72
KAYNAKÇA.....	74
EKLER DİZİNİ	78
EK 1. SIMUMIRT İLE ÖRNEK VERİ ÜRETME KODU	79
EK 2. VERİ ÜRETMEDE KULLANILAN MADDE PARAMETRELERİ.....	81
EK 3. IRTPRO 2.1 İLE ÖRNEK KESTİRİM KODLARI.....	84
EK 4. MADDE PARAMETRESİ KESTİRİMLERİNE İLİŐKİN ORTALAMANIN GÜVEN ARALIĐI TESTİ SONUÇLARI.....	86
EK 4. ETİK KURUL İZİNİ MUAFİYET FORMU.....	88
EK 5. ORJİNALLİK RAPORU	89
ÖZGEÇMİŐ	90

TABLULAR DİZİNİ

Tablo 1.1:Basit Yapıda Parametreler	10
Tablo 4.1: Tek Boyutlu Testler İçin Madde Parametrelerine Ait Bulgular	37
Tablo 4.2: Tek Boyutlu Testler İçin Birey Parametrelerine Ait Bulgular	40
Tablo 4.3: İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Bulgular	43
Tablo 4.4: İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Bulgular	43
Tablo 4.5: İki Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Bulgular	49
Tablo 4.6: İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Bulgular	49
Tablo 4.7: Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Bulgular	55
Tablo 4.8. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Bulgular	55
Tablo 4.9: Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Bulgular	60
Tablo 4.10: Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Bulgular	60

ŞEKİLLER DİZİNİ

Şekil 1.1. Madde Karakteristik Eğrisi	5
Şekil 1.2. Çok Boyutlu Madde Tepki Kuramında Modeller	7
Şekil 1.3. Basit yapı (solda) ve Karmaşık Yapı (sağda)	8
Şekil 1.4. Telafisel Modele Ait Madde Karakteristik Yüzeyleri	15
Şekil 1.5. Telafisel Olmayan Modele İlişkin Madde Karakteristik Yüzeyi	15
Şekil 1.6. Telafisel Model (sağda) ve Telafisel Olmayan Modele (solda) İlişkin Çizilen Eş Olasılık Konturları	17
Şekil 1.7. Madde Vektörleri.....	19
Şekil 4.1. Tek Boyutlu Testler İçin Madde Parametrelerine Ait RMSE Değerleri ...	37
Şekil 4.2. Tek Boyutlu Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri .	38
Şekil 4.3. Tek Boyutlu Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri	39
Şekil 4.4. Tek Boyutlu Testler İçin Birey Parametrelerine Ait RMSE Değerleri	40
Şekil 4.5. Tek Boyutlu Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri	41
Şekil 4.6. Tek Boyutlu Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri	41
Şekil 4.7. İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait RMSE Değerleri	42
Şekil 4.8. İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri	44
Şekil 4.9. İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri	45
Şekil 4.10. İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri	46
Şekil 4.11. İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri	46
Şekil 4.12. İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri	47
Şekil 4.13. İki Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait RMSE Değerleri	48
Şekil 4.14. İki Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri	50
Şekil 4.15. İki Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri	50
Şekil 4.16. İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri	51
Şekil 4.17. İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri	52
Şekil 4.18. İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri	52
Şekil 4.19. Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait RMSE Değerleri	53
Şekil 4.20. Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri	54
Şekil 4.21. Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri	56

Şekil 4.22. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri	57
Şekil 4.23. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri	58
Şekil 4.24. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri	58
Şekil 4.25. Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait RMSE Değerleri	59
Şekil 4.26. Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri	61
Şekil 4.27. Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri	61
Şekil 4.28. Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri	62
Şekil 4.29. Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri	63
Şekil 4.30. Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri	63

SİMGELER VE KISALTMALAR DİZİNİ

KTK: Klasik test kuramı

MTK: Madde Tepki Kuramı

ÇBMTK: Çok Boyutlu Madde Tepki Kuramı

1PLM: Bir parametrelili lojistik model

2PLM: İki parametrelili lojistik model

3PLM: Üç parametrelili lojistik model

MKE: Madde Karakteristik Eğrisi

RMSE: Standart Hata Ortalamalarının Karekökü

1. GİRİŞ

Bu bölümde araştırmanın problem durumuna, problem cümlesine, alt problemlere, araştırmanın amacı ve önemi başlıklarına yer verilmiştir.

1.1. Problem Durumu

Yapılacak ölçme işleminin hatasız veya en az hatalı olması için uyulması gereken kurallar, izlenmesi gereken belli başlı bir süreç vardır. Bu süreç, kurallar, sayıtlar, tanımlar ve teoremler bir teori başlığı altında ele alınabilir. Klasik test kuramı (KTK) bu amaçla geliştirilmiş ve üzerinde çok çalışılmış ilk teoridir. Her kuram gibi KTK da çeşitli varsayımlara dayanır ve bu varsayımlar pek çok veri setinde kolaylıkla karşılanabilir. Ancak yapılan işlemler her ne kadar iyi ifade edilmiş ve doğru olarak yapılmış olsa da kuramın doğasından kaynaklı bazı sınırlılıklar mevcuttur (Hambleton ve Swamington, 1985; Baykul, 2000).

Bu sınırlılıklardan bazıları şu şekilde sıralanabilir: KTK'da madde istatistikleri gruba bağımlı olarak kestirilir. Sadece aynı formu veya paralel formları alan bireylerin puanları karşılaştırılabilir. Grubun dağılımı fark etmeksizin tek bir standart hata değeri hesaplanır. Tüm maddelerin aynı cevap formatına sahip olması gereklidir. Testi alan bir bireyin bir maddeye ilişkin olarak doğru cevaplama olasılığı test uygulanmadan önce bilinemez. Ayrıca KTK modelleri test geliştirme, test eşitleme ve madde/test yanlılığını belirlemede yetersiz kalır, tatmin edici çözümler sunmaz (Hambleton ve Swamington, 1985; Embretson and Reise, 2000).

Tüm bu nedenlerden dolayı araştırmacılar daha uygun çözümler sunan bir ölçme kuramının geliştirilmesine yoğunlaşmışlardır. KTK'nın bu sınırlılıklarını karşılamak üzere Örtük Özellikler Kuramı ismi ile de bilinen Madde Tepki Kuramı ortaya atılmıştır.

MADDE TEPKİ KURAMI (MTK)

Madde tepki kuramı ilk olarak Lord ve Novick (1968)'in *Statistical Theories Of Mental Test Scores* kitabında Allan Birnbaum tarafından ele alınmıştır. Tüm psikometri yöntemleri için bir yapı taşı olan bu kitap madde tepki kuramının geliştirilmesinde önemli rol oynamıştır. Takip eden yıllar boyunca pek çok bilim insanı kurama katkı sağlamıştır (Embretson and Reise; 2000).

MTK'da "yetenek" kavramı ile ilgilenilir ve yetenek θ (teta) simgesi ile gösterilir. Kurama göre bireyin bir testte gösterdiği performansın altında o test ile ölçülmeye çalışılan yeteneği yatar. Kuramın amacı bireyin testle ölçülen yetenek düzeyini tahmin etmek veya kestirmek için bir temel oluşturmaktır. Buna göre bir bireyin yetenek düzeyi, bireyin maddelere verdiği cevaplardan kestirilmeye çalışılır. Kuram ile bireyin bir testte ortaya koyduğu performans ve bu performansın altında yatan gözlenemeyen yeteneği veya özelliği arasında bir ilişki ortaya konur. Gözlenen ve gözlenemeyen bu özellikler arasındaki ilişki matematiksel fonksiyonlar ile açıklanır. Bu sebeple madde tepki modelleri matematiksel modellerdir. Bu modeller oldukça güçlü varsayımlara dayanır (Hambleton ve Swamington, 1985; Embretson ve Reise, 2000).

MTK'nın varsayımları tek boyutluluk ve yerel bağımsızlık başlıkları altında ele alınır. Bu varsayımlar aşağıda açıklanmıştır:

Tek boyutluluk: Test ile tek bir örtük özelliğin ölçülüyor olmasıdır. Pratikte bunu karşılamak güçtür. Bu sebeple testin baskın bir faktörü ölçtüğünün gösterilmesi tek boyutluluğun sağlanması için yeterlidir (Hambleton ve Swamington, 1985; Embretson ve Reise, 2000).

Yerel Bağımsızlık: Aynı yetenek düzeyindeki bireylerin farklı maddelere verdiği tepkilerin istatistiksel olarak bağımsız olmasıdır. Bu varsayım pratikte bireyin bir maddeye verdiği cevabın diğer bir maddeyi olumlu ya da olumsuz etkilememesi anlamına gelir (Hambleton ve Swamington, 1985; Embretson ve Reise, 2000).

Yukarıda yer alan varsayımlara ek olarak tüm MTK modellerinin sağlaması gereken koşul kullanılan testin hız testi olmamasıdır. Bir testin hız testi olması testin ilgilenilen yeteneğin yanında performans hızını da ölçtüğü anlamına gelir. Bu durumda da testin tek boyutluluk varsayımı ihlal edilmiş olur (Hambleton ve Swamington; 1985).

MTK'da bireyin yetenek düzeyi ile maddeye doğru cevap verme olasılığı arasındaki matematiksel fonksiyonu veren bir madde karakteristik fonksiyonu ve fonksiyona dayalı olarak elde edilen madde karakteristik eğrisi (MKE) mevcuttur. MKE'nin şekli yetenek düzeyindeki değişim ile doğru cevaplama olasılığı arasındaki ilişkiyi ortaya koyar. Bu eğri; madde puanının yetenek üzerine regresyonudur.

MTK varsayımlarının karşılanmış olması doğrudan analize geçilebileceği anlamına gelmez. Bunun için öncelikle hangi modele göre analiz yapılacağına karar verilmesi gereklidir. MTK'da modeller cevap kategorisi sayısına göre belirlenir. Buna göre MTK modelleri iki kategorili modeller ve çok kategorili modeller başlıkları altında toplanır. Bu çalışmada sadece iki kategorili MTK modellerine değinilmiştir.

İki Kategorili MTK Modelleri

İki kategorili MTK modelleri maddelerin doğru (1) ve yanlış (0) şeklinde ikili olarak puanlandığı durumlar için kullanılan modellerdir. Burada bir maddeye ilişkin olarak verilen tepkiler başarı testlerinde doğru ya da yanlış olabildiği gibi; anketlerde evet-hayır, tutum ölçeklerinde katılıyorum-katılmıyorum şeklinde de olabilir.

İki kategorili MTK modellerinde maddelere ait madde güçlüğü, madde ayırt ediciliği ve şans (tahmin) parametresi olmak üzere üç parametre söz konusudur. Kullanılan model, olasılık fonksiyonunda yer alan parametre sayısına bağlı olarak, 1, 2 veya 3 parametrelilik şeklinde tanımlanır. Olasılık fonksiyonları lojistik ve ogive modeller şeklinde iki çeşittir. Modellerin matematiksel formları farklılık göstermesine karşılık, her iki modelden birbirine çok yakın sonuçlar elde edilir. Lojistik modeller ogive modellere göre hesaplamada kolaylık sunar ve bu sebeple daha çok tercih edilir (Embretson ve Reise; 2000). Bu sebeple burada lojistik modellere değinilmiştir.

- **1 parametrelilik lojistik model (1PLM) ve Rasch modeli**

1PLM'de testte yer alan tüm maddelerin ayırt ediciliği birbirine eşit olup maddeler güçlük bakımından farklılaşmaktadır. Bu modelde bireyin doğru cevap verme olasılığı ile yetenek düzeyi arasındaki ilişki madde güçlüğü parametresi üzerinden tanımlanır. Bu modelde bir s bireyinin i maddesini doğru cevaplama olasılığı eşitlik 1.1'deki gibi tanımlanır:

$$P(X_{is} = 1 | \theta_s, \beta_i) = \frac{\exp(\alpha(\theta_s - \beta_i))}{1 + \exp(\alpha(\theta_s - \beta_i))} \quad (\text{Eşitlik 1.1})$$

X_{is} = s bireyin i maddesine verdiği cevap (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesinin güçlüğü

α_i = i maddesinin ayırtıcılığı

1PLM'ye uygun bir testte; yetenek düzeyi arttıkça doğru cevaplama olasılığı artar, maddeler sadece güçlük değeri açısından farklılaşır. Madde ayırt ediciliklerinin birbirine eşit olması dolayısıyla çizilen MKE'ler birbirini kesmez. Şans parametresi ihmal edilebilir düzeydedir. Madde güçlüğünün değeri, madde karakteristik eğrilerinin 0,50 olasılığa denk gelen yetenek düzeyi eksenindeki değerdir.

Rasch modeli, 1PLM'nin özel bir durumudur. Rasch modelinde madde güçlükleri farklılaşır, madde ayırtıcılıkları ise 1'e eşittir. Rasch modeli eşitlik 1.2'de yer alan fonksiyon ile özetlenebilir. Formülde yer alan terimler eşitlik 1.1 ile aynı anlamları taşımaktadır.

$$P(X_{is} = 1 | \theta_s, \beta_i) = \frac{\exp(\theta_s - \beta_i)}{1 + \exp(\theta_s - \beta_i)} \quad (\text{Eşitlik 1.2})$$

- **2 parametrelili lojistik model (2PLM)**

2PLM'de tüm maddelerin madde güçlüklerinin yanı sıra madde ayırt edicilikleri de farklılaşır. 2PLM'de bireylerin doğru cevaplama olasılığı yetenek düzeyinin yanında madde güçlüğü ve madde ayırt ediciliği parametreleri üzerinden tanımlanır. Bu modelde bir s bireyinin i maddesini doğru cevaplama olasılığı eşitlik 1.3'te verilmiştir.

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i) = \frac{\exp[\alpha_i(\theta_s - \beta_i)]}{1 + \exp[\alpha_i(\theta_s - \beta_i)]} \quad (\text{Eşitlik 1.3})$$

X_{is} = s bireyin i maddesine verdiği cevap (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

β_i = i maddesinin güçlüğü

α_i = i maddesinin ayırtıcılığı

2PLM, maddelerin yetenek düzeyleri ile eşit şekilde ilişkili olmadığı durumlarda kullanılır. Ayırt edicilik parametreleri madde için çizilen madde karakteristik eğrisinin eğiminden elde edilebilir. Güçlük parametresi 2PLM'de de Rasch modelindeki gibi 0,50 olasılığa denk gelen θ eksenindeki değerdir.

- **3 parametrelili lojistik model (3PLM)**

3PLM'de bireyin doğru cevap verme olasılığını etkileyen parametreler kişinin yetenek düzeyi ile beraber maddenin güçlüğü, ayırtıcılığı ve şans parametresidir. Bu modele göre bir maddenin doğru cevaplanma olasılığı tüm yetenek düzeyleri için her zaman sıfırın üzerindedir. Tahmin parametresinin değeri madde karakteristik

eğrisinin en düşük asimptotu olarak kestirilir. 3PLM'ye ilişkin olasılık fonksiyonu eşitlik 1.4'teki gibidir:

$$P(X_{is} = 1 | \theta_s, \beta_i, \alpha_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \frac{\exp[\alpha_i(\theta_s - \beta_i)]}{1 + \exp[\alpha_i(\theta_s - \beta_i)]} \quad (\text{Eşitlik 1.4})$$

X_{is} = s bireyin i maddesine verdiği cevap (1 veya 0)

θ_s = s bireyinin yetenek düzeyi

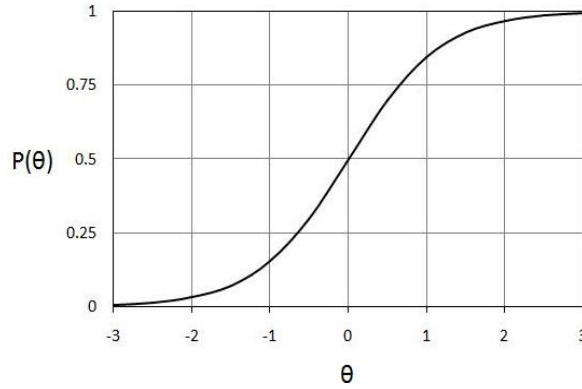
β_i = i maddesinin güçlüğü

α_i = i maddesinin ayırtıcılığı

γ_i = i maddesinin en düşük asimptotu (şans parametresi)

3PLM için çizilen MKE, şans parametresinin sıfırdan büyük olması sebebiyle orijinden başlamaz. Bu sebeple güçlük parametresi 0,50 olasılığa denk gelen nokta olarak belirlenmez. Bu modelde madde güçlüğü şans parametresinden etkilenir ve x-ekseninde $\frac{1+\gamma_i}{2}$ olasılığa denk gelen nokta olarak kestirilir.

İki kategorili bir maddeye ilişkin çizilen bir MKE'nin şekli genel itibari ile şekil 1.1'deki gibi verilebilir.



Şekil 1.1. Madde Karakteristik Eğrisi

Grafiğin yatay eksenini θ yetenek düzeylerini, dikey eksenini ise doğru cevaplama olasılığını ($P(\theta)$) verir. Bir maddeye ait MKE'ye bakılarak madde parametreleri tanımlanabilir. Buna göre bir MKE'nin bükülme noktası maddenin güçlüğü, eğimi maddenin ayırt ediciliğini ve en düşük asimptotu (eğrinin y-eksenini kestiğini nokta); maddenin şans parametresini verir.

Bir bireyin bir maddeyi doğru cevaplama olasılığı sadece madde karakteristik eğrisine bağlıdır. Bu olasılık bireyin içinde bulunduğu evrenin yetenek dağılımından

bağımsızdır. Dolayısıyla bir bireyin doğru cevap verme olasılığı aynı yetenek düzeyindeki birey sayısına da bağlı değildir. Bu değişmezlik özelliği madde parametrelerinin değişmezliğini de getirir (Hambleton ve Swaminogton; 1985).

Yukarıda kısaca değinildiği gibi her ne kadar MTK'da farklı özellikteki durumlara ilişkin birbirinden farklı modeller yer alsa da, varsayımlar karşılanmadığı takdirde bu modellerin kullanılması uygun olmayacaktır. Uygulamalarda en sık karşılaşılan durum, kullanılan ölçme araçlarının tek boyutlu olmamasıdır. Bu durumda MTK'nın temel varsayımı olan tek boyutluluk ihlal edilmektedir. Bu ve buna benzer durumlarda tek boyutlu MTK modelleri yerine çok boyutlu veri setleri için geliştirilmiş olan çok boyutlu madde tepki kuramının (ÇBMTK) kullanılması daha uygundur.

ÇOK BOYUTLU MADDE TEPKİ KURAMI (ÇBMTK)

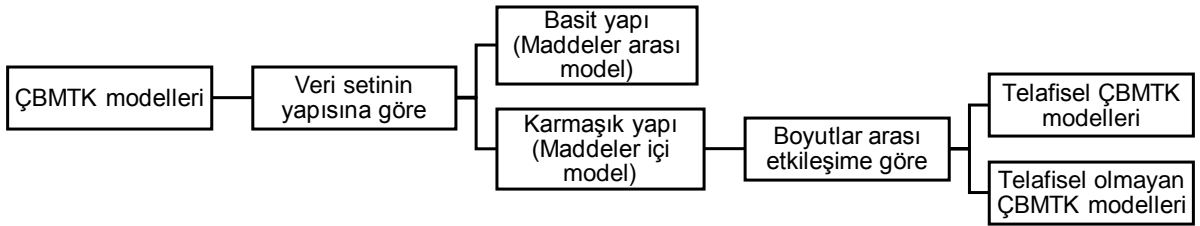
Çok boyutlu madde tepki kuramı eğitim, psikoloji, psikometri ve istatistik gibi pek çok bilim dalındaki gelişmelerden esinlenmektedir. Bu kuramın temelinde iki düşünce yer alır. İlki, gerçekte karşılaştığımız durumlar düşündüğümüzden daha karmaşıktır. İkincisi ise bu karmaşık durumlar farklı model ve teoriler ile açıklanabilir. ÇBMTK gerçek durumların idealize edildiği ve sadeleştirildiği bir model veya teoridir. Bireylerin kapasiteleri ile maddelere verdikleri cevaplar arasındaki ilişkiye dair bir yakınsama verir. Kuram, bireylerin test maddelerine verdikleri cevapların olasılıkları ile çok boyutlu uzaydaki yerleri arasındaki ilişkinin matematiksel olarak ifade edilebileceği temeline dayanır (Reckase, 2009).

Tek boyutlu madde tepki kuramı yapı veya boyut olarak adlandırılan gözlenemeyen *bir* değişken ile bireyin belli bir test maddesine doğru cevap verme olasılığı arasındaki ilişkiyi modellemek için kullanılır. Buna karşılık ÇBMTK *iki ya da daha* fazla gözlenemeyen değişken ile bireyin belli bir test maddesine doğru cevap verme olasılığı arasındaki ilişkiyi modellemek için kullanılır (Ackerman, Gierl ve Walker, 2003). Tek boyutlu MTK modellerinde doğru cevaplama olasılığı tek bir yeteneğe dayanırken, ÇBMTK'da doğru cevaplama olasılığı ölçülmeye çalışılan çok sayıda yeteneğin bir fonksiyonudur. Buna bağlı olarak da ÇBMTK'da ölçme modeli test sonucu olarak tek bir puan yerine bir yetenek profili ortaya koyar (Hartig ve Höhler, 2009).

Bir veri setine ÇBMTK'yı uygulamak için veri setinin çok boyutlu olmasının yanında başka varsayımlara da ihtiyaç duyulur. Bu varsayımlardan biri yerel bağımsızlık

diğeri monotonluktur. Yerel bağımsızlık varsayımı tek boyutlu MTK'daki ile aynı anlamı taşıdığından burada tekrar anlatılmamıştır. Monotonluk varsayımı ise bireyin doğru cevap verme olasılığının, θ -vektöründeki herhangi bir elemanın daha yüksek değer alması ile artmasıdır (Reckase, 2009). Diğer bir ifade ile maddenin ilişkili olduğu yeteneklerden herhangi birinin artmasıyla maddeyi doğru cevaplama olasılığının artmasıdır. Buna göre bireyin doğru cevap verme olasılığı ile θ 'sı arasında monoton artan bir fonksiyon ortaya konmalıdır. Monotonluk varsayımı tek boyutlu MTK'da da karşılanabilmesine rağmen bu varsayıma ihtiyaç duymayan tek boyutlu MTK modelleri de mevcuttur. Bu sebeple monotonluk tek boyutlu MTK'da bir varsayım olarak ele alınmaz (Roberts ve ark 2000; akt: Reckase, 2009).

ÇBMTK'da kullanılan farklı modeller, yetenek ile doğru cevaplama olasılığı arasında farklı istatistiksel ilişkileri varsayar. Buna ek olarak boyutlar ve maddeler arasındaki ilişki veri setinin yapısına göre tanımlanabilir. ÇBMTK'da modeller veri setinin karmaşıklığına göre ve boyutlar arası etkileşime göre sınıflandırılmıştır (Hartig ve Höhler, 2009). Bu sınıflandırma şekil 1.2'de ki gibi verilebilir.

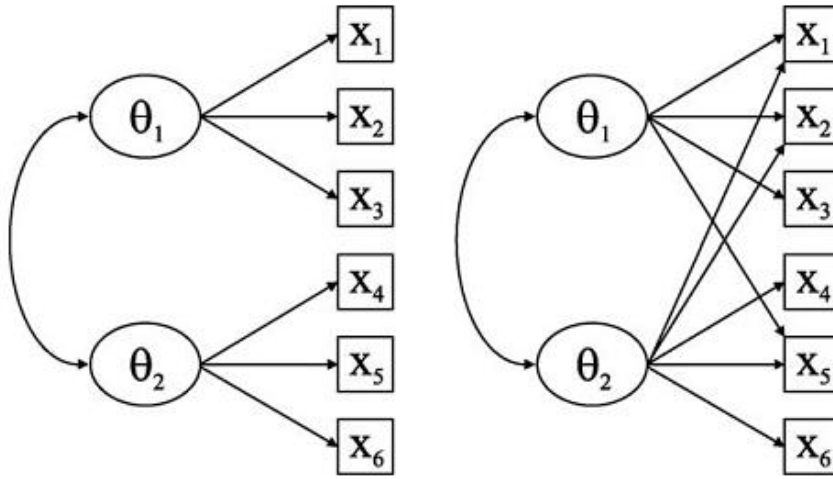


Şekil 1.2. Çok Boyutlu Madde Tepki Kuramında Modeller

ÇBMTK'da bir test birden fazla yeteneği içerir. Ancak bu durum iki farklı yapı ile elde edilebilir. Birinci yapı basit yapı adını alır. Basit yapı bir test birden fazla alt testten oluşur. Her bir alt test tek bir yeteneği veya özelliği ölçer ve kendi içinde homojendir. Her bir madde yer aldığı alt test ile ölçülmeye çalışılan tek bir özellik ile ilişkilidir. Basit yapı testlerin ortaya koyduğu modele *maddeler arası model* adı verilir. Bu model genelde geniş ölçekli testlerde karşımıza çıkar. Örneğin TIMSS (Trends in International Mathematics and Science Study) sınavı matematik ve fen bilimleri alt testlerinden oluşmaktadır. Her bir alt test çeşitli konu ve bilişsel alanları içermektedir. Matematik alt testinde cebir, geometri, sayılar ve veri gösterimi gibi farklı konu alanları ile bilme, uygulama ve mantık gibi farklı bilişsel alanlar ölçülmektedir. Her

bir konu alanı ham puanların rapor edildiği tek boyutlu bir alt test olarak ele alınır. Bunun yanında testin tamamının matematik yapısını ölçtüğü varsayılır. Bu türdeki çok boyutlu testlere çoklu-tek boyutlu model (multi-unidimensional) adı verilmektedir (Sheng 2005; Sheng ve Wikle, 2007; Bulut, 2013).

ÇBMTK'da ele alınan ikinci yapı ise karmaşık yapıdır. Karmaşık yapı, diğer bir ifade ile basit olmayan yapı, maddeler içi model olarak da bilinir. Bu modelde maddeler tek bir yetenek ile ilişkili değildir. Hem testin bütünü hem de testteki maddeler birden çok yetenek ile ilişkilidir. Bunun bir sonucu olarak maddeler ölçülmek istenen birden fazla yeteneğe faktör yükü verir (Sheng ve Wikle, 2007; Bulut, 2013). Basit yapı ve karmaşık yapıya ilişkin bir görsel şekil 1.3'te yer almaktadır (Hartig ve Höhler, 2009).



Şekil 1.3. Basit yapı (solda) ve Karmaşık Yapı (sağda)

Şekil 1.3'te θ_1 ve θ_2 ölçülen örtük özellikler olmak üzere, X_1 - X_6 maddeleri göstermektedir. Sol taraftaki gösterim maddeler arası model, sağ taraftaki gösterim maddeler içi modele işaret etmektedir (Hartig ve Höhler, 2009).

Boyutlar arası etkileşime dayalı yapılan sınıflandırmada öncelikli koşul, ele alınacak maddelerin maddeler içi modele sahip olmasıdır. Bir madde birden fazla özelliği ölçüyorsa, ölçtüğü özellikler arasındaki etkileşime dayalı olarak telafisel veya telafisel olmayan (kısmi telafisel) model başlığı altında ele alınabilir. Telafisel modelde; bir bireyin bir yetenekteki zayıflığı, maddenin ilişkili olduğu diğer boyutta sahip olduğu yüksek yetenek tarafından telafi edilir ve bu şekilde bireyin maddeyi

doğru cevaplama olasılığı artar. Bu modelin matematiksel alt yapısında, θ -koordinatlarının doğrusal kombinasyonuna ogive veya lojistik fonksiyon eklemesi yapılır. Telafisel olmayan modelde bireyin maddeyi doğru cevaplayabilmesi için maddenin ilişkili olduğu tüm boyutlarda belli bir yetenek düzeyine sahip olması gereklidir. Buna göre bu modelde, maddeyi doğru cevaplamak için bir boyutta sahip olunması gereken minimum düzeyden yüksek yetenek, diğer boyutta sahip olunması gereken minimum düzeyden düşük yeteneği telafi edemez. Bir bireyin bu tipteki bir maddeyi doğru cevaplama olasılığı maddeyi cevaplamak için gerekli yeteneklerin olasılıklarının çarpımıdır (Ackerman, 1996; Ackerman, Gierl ve Walker, 2003; Reckase, 2009).

Telafisel modele örnek olarak yabancı dil sınavında yer alan astronomi ile ilgili bir paragraf sorusu verilebilir. Böyle bir soruda hem astronomi bilgisi hem de yabancı dil bilgisi yoklanmaktadır. Ancak astronomi bilgisi yüksek olan bir birey yabancı dile ilişkin yeteneği düşük olsa da maddeyi doğru cevaplayabilir. Ya da bireyin astronomi bilgisi yeterli olmasa dahi sahip olduğu yüksek yabancı dil bilgisi ile maddeyi doğru cevaplayabilir. Burada maddenin ilişkili olduğu bu iki yetenek arasında –astronomi ve yabancı dil- telafi edicilik ilişkisi söz konusudur. Buna karşılık kimya dersindeki bir laboratuvar uygulamasını ele alalım. Burada öğrencinin hem kimya bilgisine sahip olması hem de deneyi nasıl yapacağını bilmesi gerekir. Öğrencinin kimya alanında bilgisi ne kadar fazla olursa olsun, başarılı olması için ilgili deneyin nasıl yapılacağına ilişkin olarak da belli bir bilgi düzeyine sahip olmalıdır. Bu durum da telafisel olmayan modele bir örnektir.

Basit yapı gösteren çoklu tek boyutlu model ile karmaşık model ortaya koyan telafisel ve telafisel olmayan model aşağıda matematiksel fonksiyonları ile birlikte ele alınmıştır.

Basit Yapı - Çoklu Tek Boyutlu Model (Multi-Unidimensional Model)

Çoklu-tek boyutlu modelde test çeşitli alt testlerden oluşur. Model, testte yer alan her bir alt testin belirli bir özelliği maddeler aracılığı ile ölçtüğü varsayımına dayanır. Bu modelde testteki her madde tek bir özelliği ölçer. Test içerisindeki her bir alt test tek boyutlu, her bir alt test farklı bir özelliği ölçtüğü için de testin tamamı çok boyutludur. Çoklu-tek boyutlu model ÇBMTK'nın özel bir durumudur. Modelde madde parametrelerinden güçlük ve şans parametreleri her madde için bir tane

kestirilirken; ayırt edicilik parametresi, her madde için testte yer alan boyut sayısı kadar elemana sahip bir vektör şeklinde kestirilir. Buna göre m alt testten oluşan bir testteki j maddesinin ayırt ediciliği $\alpha_j = (\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj})$ vektörüne eşittir. Ancak her bir maddenin ayırt ediciliği ilişkili olmadığı alt testlerde sıfırdır. Bu sebeple vektör $\alpha_j = (0, 0, \dots, \alpha_{vj}, 0, \dots, 0)$ şeklinde sadeleştirilir. Basit yapı gösteren 3 alt testten oluşan 3PLM'deki bir testin madde parametrelerinin genel görünüşü tablo 1.1'deki gibidir.

Tablo 1.1: Basit Yapıda Parametreler

Alt Test (i)	Madde (j)	α_{1j}	α_{2j}	α_{3j}	β_{ij}	γ_{ij}
1	1	α_{11}	0	0	β_{11}	γ_{11}
1	2	α_{12}	0	0	β_{12}	γ_{12}
1	3	α_{13}	0	0	β_{13}	γ_{13}
2	4	0	α_{24}	0	β_{24}	γ_{24}
2	5	0	α_{25}	0	β_{25}	γ_{25}
2	6	0	α_{26}	0	β_{26}	γ_{26}
3	7	0	0	α_{37}	β_{37}	γ_{37}
3	8	0	0	α_{38}	β_{38}	γ_{38}
3	9	0	0	α_{39}	β_{39}	γ_{39}

Çoklu-tek boyutlu 3 parametrelili lojistik modelde i bireyinin v alt testinde yer alan j maddesine doğru cevap verme olasılığı eşitlik 1.5'teki gibidir.

$$P(U_{vij} = 1 | \theta_{vi}, \alpha_{vj}, \beta_{vj}, \gamma_{vj}) = \gamma_{vj} + \frac{(1 - \gamma_{vj})}{1 + e^{[(\sum_{l=1}^v \beta_{jl} \theta_{il}) + \alpha_{vj}]}} \quad (\text{Eşitlik 1.5})$$

U_{vij} = i bireyinin v alt testindeki j maddesine verdiği cevap (1 veya 0)

θ_{vi} = i bireyinin v alt testine ilişkin yetenek düzeyi

β_{vj} = v alt testindeki j maddesinin güçlüğü

α_{vj} = v alt testindeki j maddesinin ayırıcılığı

γ_{vj} = v alt testindeki j maddesinin en düşük asimptotu (şans parametresi)

Modelden şans parametresi çıkartılırsa çoklu-tek boyutlu 2PL modele ilişkin olasılık fonksiyonu elde edilir.

Karmaşık Yapı - Maddeler İçi Model

Karmaşık yapı bir testte bir madde birden fazla örtük özellik ile ilişkilidir. Model örtük özellikler arasındaki telafi edicilik ilişkisine göre telafisel ve telafisel olmayan modeller şeklinde ele alınır.

Telafisel Modeller İçin Genişletilmiş Tek Boyutlu Modeller

Bu modelde ortaya konan matematiksel fonksiyonlar yetenek (θ), ayırıcılık (α), güçlük (β) ve şans (γ) parametreleri üzerinden ele alınır ve içerdiği parametre sayısına göre adlandırılır. Bu bakımdan telafisel modelde tanımlanan modeller tek boyutlu MTK'ya benzemektedir. Ancak tek boyutlu MTK'dan farklı olarak telafisel modellerde boyut sayısı kadar α parametresi ve θ parametresi yer alır. Ortaya konan matematiksel fonksiyonlar tek boyutlu MTK'da yer alan olasılık fonksiyonlarının genişletilmiş halleridir. Telafisel modelde;

- Çok boyutluluk için genişletilmiş lojistik modeller
- Çok boyutluluk için genişletilmiş normal ogive modeller

tanımlanmıştır. Modelde maddenin ilişkili olduğu yetenekler arasında telafi edicilik ilişkisi bulunması sebebiyle bir maddeyi doğru cevaplama olasılığı maddenin ilişkili olduğu yeteneklerdeki olasılıkların toplamından yararlanarak hesaplanır (Reckase, 2009).

- **Çok boyutluluk için genişletilmiş lojistik modeller**

Çok boyutlu lojistik modeller arasında var olan ilişki tek boyutlu MTK'da var olan ilişki ile aynıdır. Burada ele alınan modeller;

- Çok boyutluluk için genişletilmiş Rasch modeli
- Çok boyutluluk için genişletilmiş iki parametrelili lojistik model
- Çok boyutluluk için genişletilmiş üç parametrelili lojistik model

şeklinde adlandırılır.

Buna göre, şans parametresinin sıfırdan büyük ve ihmal edilemez olduğu durumda çok boyutluluk için genişletilmiş üç parametrelili lojistik model kullanılır. Şans parametresinin sıfır veya ihmal edilebilir olduğu ve ayırt edicilik parametrelerinin maddeler boyunca farklılaştığı durumda çok boyutluluk için genişletilmiş iki parametrelili lojistik model kullanılır. Maddelerin sadece güçlük parametresi bakımından farklılaştığı ve tüm ayırıcılık parametrelerinin 1'e eşit olduğu durumda çok boyutluluk için genişletilmiş Rasch modeli kullanılır.

Çok boyutluluk için genişletilmiş üç parametrelili lojistik modelde j bireyinin i maddesine doğru cevap verme olasılığına ilişkin olasılık fonksiyonu aşağıda verilmiştir.

$$P(U_{ij} = 1 | \theta_j, \alpha_i, \gamma_i, d_i) = \gamma_i + (1 - \gamma_i) \frac{e^{\alpha_i \theta_j' + d_i}}{1 + e^{\alpha_i \theta_j' + d_i}} \quad (\text{Eşitlik 1.6})$$

U_{ij} = j bireyinin i maddesine verdiği cevap (1 veya 0)

θ_j = j bireyinin yetenek düzeyi

α_i = i maddesinin ayırtıcılığı

γ_i = i maddesinin en düşük asimptotu (şans parametresi)

d_i = i maddesinin kesim noktası

Denklemden yer alan d parametresi α ve β parametrelerinin etkileşimi ile elde edilen kesim noktasıdır. Buna göre tek boyutlu lojistik modellerde yer alan $(\alpha(\theta - \beta))$ çarpımı yapılırsa $(\alpha\theta - \alpha\beta)$ elde edilir. Burada yer alan $-\alpha\beta$ terimi d ile yer değiştirildiğinde $\alpha\theta + d$ ifadesi elde edilmiş olur. Tek boyutlu model çok boyutlu modele genişletildiğinde; m boyutlu bir test için α parametresi 1 x m elemanlı ayırt edicilik vektörünü; θ parametresi 1 x m elemanlı birey koordinatları vektörünü temsil eder. d parametresi $\alpha\theta + d$ denkleminin kesim noktasını verir.

Modelde yer alan “e” teriminin üssünde yer alan ifade α ve θ parametrelerinin etkileştiğini göstermek üzere aşağıdaki şekilde genişletilebilir:

$$\alpha_i \theta_j' + d_i = \alpha_{i1} \theta_{j1} + \alpha_{i2} \theta_{j2} + \alpha_{i3} \theta_{j3} + \dots \dots \dots + \alpha_{im} \theta_{jm} + d_i = \sum_{l=1}^m \alpha_{il} \theta_{jl} + d_i \quad (\text{Eşitlik 1.7})$$

(Reckase, 2009).

• **Çok boyutluluk için genişletilmiş normal ogive model**

ÇBMTK çalışmalarında bir maddeye doğru cevap verme olasılığı ile çok boyutlu uzaydaki konum arasındaki ilişkiyi tanımlamak için daha çok normal ogive modellerden yararlanılmaktadır. Günümüzde istatistiksel kestirim programları ogive modellere dayalı olarak da kestirim yapmaktadır. Çok boyutluluk için genişletilmiş normal ogive modelin üç parametrelili formu eşitlik 1.8'deki gibidir:

$$P(U_{ij} = 1 | \theta_j, \alpha_i, \gamma_i, d_i) = \gamma_i + (1 - \gamma_i) \frac{1}{\sqrt{2\pi}} \int_{-z_i(\theta_j)}^{\infty} e^{-\frac{t^2}{2}} dt \quad (\text{Eşitlik 1.8})$$

U_{ij} = j bireyinin i maddesine verdiği cevap (1 veya 0)

$\theta_j = j$ bireyinin yetenek düzeyi

$\alpha_i = i$ maddesinin ayırcılığı

$\gamma_i = i$ maddesinin en düşük asimptotu (şans parametresi)

$d_i = i$ maddesinin kesim noktası

Denklemden yer alan $-z_i(\theta_j)$ fadesi $\alpha_i\theta_j' + d_i$ ifadesine karşılık gelmektedir. Diğer terimler daha önce tanımlandığı gibidir. Bu denklemden yer alan γ şans parametresi sıfıra eşit ise model iki parametrelili forma dönüşür. Çok boyutluluk için genişletilmiş ogive model, çok boyutluluk için genişletilmiş lojistik model ile aynı grafiksel alanı tanımlamaktadır. Haley (1952) yaptığı çalışmada lojistik modeller ile ogive modelleri karşılaştırmış, elde edilen olasılıklarda 0,01'den daha az fark olduğunu göstermiştir (Reckase, 2009).

Telifisel Olmayan Modeller İçin Genişletilmiş Tek Boyutlu Modeller

Bir maddenin ilişkili olduğu yetenekler arasında telifisellik özelliğinin olmadığı durumlar mevcuttur. Sympson (1978), bu tip durumlarda daha önce ortaya atılan modelin kullanılmasının uygun olmadığını savunmuş ve kısmi telifisel model olarak da adlandırılan telifisel olmayan modeller için genişletilmiş tek boyutlu modeli ortaya konmuştur. Bu modelde, bir maddeyi doğru cevaplama olasılığı maddenin ilişkili olduğu yetenek değerlerindeki olasılıkların çarpımı ile hesaplanır. Modele ilişkin fonksiyon aşağıda verilmiştir.

$$P(U_{ij} = 1 | \theta_j, \alpha_i, \beta_i, \gamma_i) = \gamma_i + (1 - \gamma_i) \left(\prod_{l=1}^m \frac{e^{1,7\alpha_{il}(\theta_{jl} + \beta_{il})}}{1 + e^{1,7\alpha_{il}(\theta_{jl} + \beta_{il})}} \right) \quad (\text{Eşitlik 1.9})$$

$U_{ij} = j$ bireyinin i maddesine verdiği cevap (1 veya 0)

$\theta_j = j$ bireyinin yetenek düzeyi

$\alpha_i = i$ maddesinin ayırcılığı

$\beta_i = i$ maddesinin güçlüğü

$\gamma_i = i$ maddesinin en düşük asimptotu (şans parametresi)

$m = 1, 2, 3 \dots$ olmak üzere boyut sayısı

Görüldüğü gibi model üç parametre üzerinden tanımlanmıştır. Ancak şans parametresinin (γ) sıfıra eşit olduğu durumda fonksiyon iki parametrelili modele

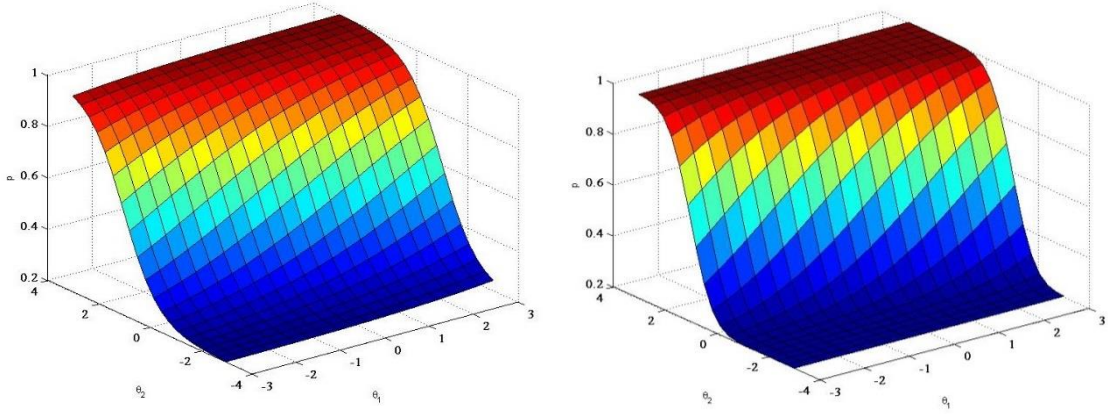
dönüşmektedir (Reckase, 2009). Telafisel olmayan çok boyutlu modellerde bir maddenin ilişkili olduğu θ kadar α ve β parametresi kestirilir. Örneğin üç boyutlu telafisel olmayan modeldeki bir testte her madde için üç tane α ve üç tane β parametresi kestirilir. Şans parametresi ise telafisel modeldeki gibi her madde için bir tane kestirilir.

Telafisel model ile telafisel olmayan model dayandıkları felsefe ve matematiksel form bakımından birbirinden ayrılır. Telafisel olmayan model, maddenin farklı beceri veya bilgilerle ilişkili olduğu durumlara uygun modeller sunar. Buna karşılık telafisel modelde birey ve maddeler arasındaki etkileşim bütüncül bir bakış açısı ile ele alınır. Yapılan az sayıda çalışmada modellerin kullanılabilirliği gerçek veri setlerini doğru biçimde ele alma ve temsil edebilme bakımından test edilmiştir (Bolt ve Lall, 2003; Babcock, 2009). Sonuçlar telafisel modelin daha iyi sonuçlar ürettiği ve bu bakımdan daha kullanışlı olduğunu göstermektedir (Reckase, 2009).

Madde Karakteristik Yüzeyleri ve Eş Olasılık Konturları

Çok boyutluluk için genişletilmiş modellerden faydalanarak madde karakteristik yüzeyi elde edilir. Madde karakteristik yüzeyi, MKE'ye benzer biçimde doğru cevaplama olasılığını ve madde ile yoklanan θ düzlemlerini içerir. Madde karakteristik yüzeyinde ilgili yeteneklerde bireyin sahip olduğu düzey ve maddeyi doğru cevaplama olasılığı arasındaki ilişki çok boyutlu grafiksel düzlemler üzerinde gösterilir.

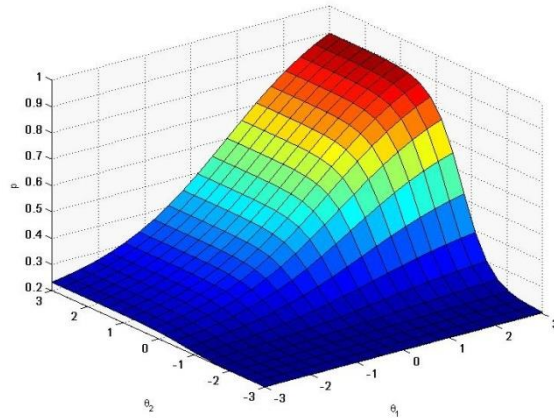
Şekil 1.4'te $\alpha_1=0,5$; $\alpha_2=1,5$; $\beta = 0$ ve $c=0,2$ olan iki boyutlu telafisel modeldeki bir maddenin lojistik ve normal ogive fonksiyona dayalı çizilen madde karakteristik yüzeyleri verilmiştir.



Şekil 1.4. Telafisel Modele Ait Madde Karakteristik Yüzeyleri (sol taraftaki şekil lojistik modelden, sağ taraftaki şekil ogive modelden elde edilmiştir.)

Şekil 1.4'te yer alan madde karakteristik yüzeyleri incelenirse; doğru cevaplama olasılığının, θ vektöründeki elemanların artışı ile monoton arttığı görülmektedir. Ayrıca yukarıda yer alan madde karakteristik yüzeylerinde θ_2 'deki birimlik artışın doğru cevaplama olasılığını θ_1 'deki aynı birimlik artıştan daha fazla arttırdığı görülmektedir. Buna göre θ_2 , bu madde için doğru cevaplama olasılığının artması üzerinde θ_1 'den daha etkilidir.

Telafisel olmayan modelde madde parametreleri $\alpha_1=0,5$; $\alpha_2=1,5$; $\beta_1=-0,80$; $\beta_2=-0,72$ ve $c=0,2$ olan bir maddenin madde karakteristik yüzeyi şekil 1.5'te verilmiştir.



Şekil 1.5. Telafisel Olmayan Modele İlişkin Madde Karakteristik Yüzeyi

Telafisel olmayan modele ilişkin olarak çizilen şekil 1.5'teki madde karakteristik yüzeyi incelenirse, yüzeyin şeklinin telafisel modelden oldukça farklı olduğu görülür.

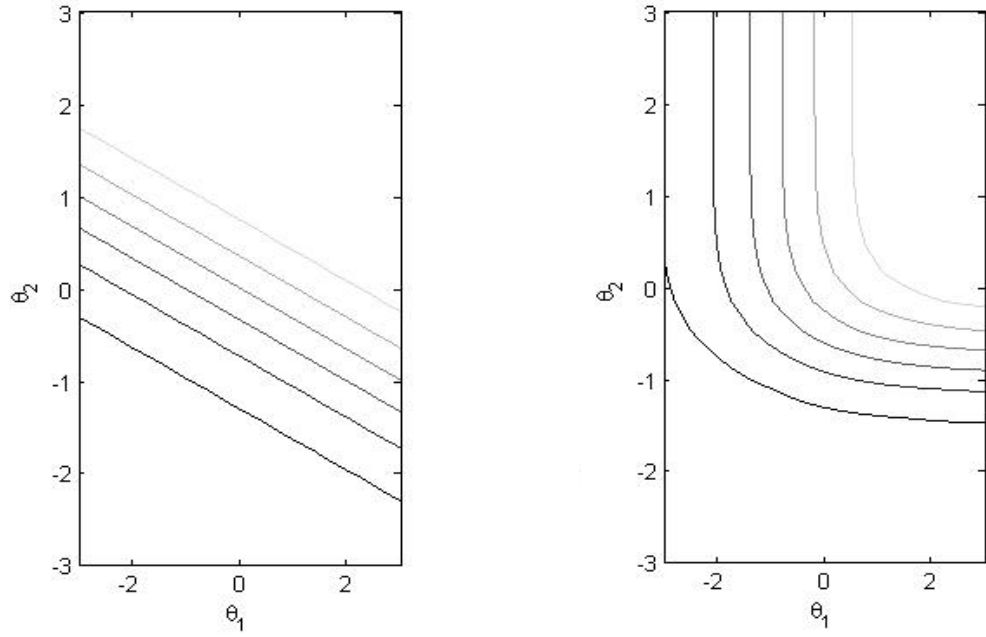
Olasılık θ_1 ve θ_2 yeteneklerinin artışı ile artmaktadır. Doğru cevaplama olasılığının en yüksek olduğu noktada hem θ_1 hem θ_2 yüksek değerler almıştır.

Bir maddenin gösterimi için madde karakteristik yüzeyi dışında madde eş olasılık konturlarından da yararlanılabilir. Madde eş olasılık konturları θ eksenlerine dayalı oluşturulan bir grafikdir. Grafikte yer alan konturlardan faydalanarak belli bir $\theta_1\theta_2$ bileşimine sahip bir bireyin doğru cevaplama olasılığı bulunabilir. Eş olasılık konturları bakımından telafisel ve telafisel olmayan model arasında çeşitli farklılıklar bulunmaktadır.

Telafisel modele ilişkin çizilen eş olasılık konturları eşit aralıklarla konumlanmış birbirine paralel düz çizgiler şeklindedir. Aynı kontur üzerinde yer alan bireylerin maddeyi doğru cevaplama olasılığı birbirine eşittir. Bu modelde madde karakteristik yüzeyinin eğimi dikleştikçe veya maddenin ayırt ediciliği arttıkça, konturlar birbirine yaklaşır. Bu sebeple telafisel modelde çizilen eş olasılık konturları ayırıcılık parametresinin bir fonksiyonudur.

Telafisel olmayan modelde çizilen eş olasılık konturları ise hiperboller şeklindedir. Ancak bu konturlar gerçek birer hiperbol değil, belirli β parametresi değerlerine çizilen asimptotlardır. Bu sebeple telafisel olmayan modelde çizilen eş olasılık konturlarının yönü madde güçlüğünün bir fonksiyonudur. Telafisel olmayan modelde konturlardan yararlanarak maddenin doğru cevaplama olasılığını hesaplamak için her bir bileşenin doğru cevaplama olasılığını bulmak gerekir. Bir bileşenin doğru cevaplama olasılığı, o bileşendeki yetenek düzeyine dik çizilen doğrunun asimptotu olduğu konturun değeridir. Maddeyi doğru cevaplama olasılığı maddenin her bir bileşenini doğru cevaplama olasılığının çarpımıdır (Ackerman, 1989; Ackerman, Gierl and Walker, 2003; Reckase, 2009).

Şekil 1.4 ve 1.5'te madde karakteristik yüzeyleri yer alan maddelere ait eş olasılık konturları şekil 1.6'da verilmiştir.



Şekil 1.6. Telafisel Model (sağda) ve Telafisel Olmayan Modele (solda) İlişkin Çizilen Eş Olasılık Konturları

Kontur grafikleri madde karakteristik yüzeylerinden daha bilgi vericidirler. Kontur grafiklerinden faydalanarak maddenin en iyi ölçtüğü yetenek birleşimi, maddelerin hangi yetenek düzeylerini daha etkili biçimde ayırdığı ve maddenin güçlüğü hakkında yorum yapılabilir. Ayrıca bu bilgilere ulaşmak amacıyla kontur grafiklerini yorumlamak madde karakteristik yüzeylerini yorumlanmaktan daha kolaydır. (Ackerman, 1996).

Çok Boyutlu Madde Tepki Kuramında Madde Parametreleri ve Madde Vektörleri

Maddeler için çizilen kontur grafikleri yorumlamayı kolaylaştırmak açısından genellikle bir madde vektörü ile birleştirilir. İki boyutlu bir madde vektör grafiği maddelerin dik açılı kartezyen koordinat sisteminde temsil edilmesini sağlar.

Madde vektörünün çizilebilmesi için vektörün uzunluğunun, yerinin ve yönünün tanımlanmış olması gereklidir. Çok boyutlu madde tepki kuramında bir madde vektörüne ait bu özellikler maddenin ayırıcılık, güçlük ve yön parametreleri ile tanımlanır.

Bir madde vektörünün uzunluğu, maksimum ayırıcılık parametresine denk gelir. Çok boyutlu madde ayırıcılığı (ÇBMA) olarak tanımlanan vektör uzunluğu, her bir yetenekteki ayırt edicilik parametrelerinin bileşiminden eşitlik 1.10 ile elde edilir:

$$\text{ÇBMA} = \sqrt{\sum_{k=1}^m \alpha_k^2} \quad (\text{Eşitlik 1.10})$$

Madde iki boyutlu ise eşitlik;

$$\text{ÇBMA} = \sqrt{\alpha_{i1}^2 + \alpha_{i2}^2} \quad (\text{Eşitlik 1.11})$$

şeklinde sadeleştirilebilir. ÇBMA tek boyutlu MTK'daki α parametresine karşılık gelir ve her zaman pozitif değer alır. Bu sebeple vektör her zaman koordinat düzleminin birinci veya üçüncü çeyreğinde yer alır.

Güçlük parametresi tek boyutlu MTK'daki madde güçlüğü parametresine karşılık gelir. Bu parametre madde vektörünün başlangıç noktasının orijine olan uzaklığıdır. Çok boyutlu madde güçlüğü (ÇBMG) olarak tanımlanan vektörün orijine uzaklığı (D) yer parametresi (d_i) ve ÇBMA'dan faydalanarak eşitlik 1.12 ile hesaplanır.

$$\text{ÇBMG} = D = \frac{-d_i}{\text{ÇBMA}} \quad (\text{Eşitlik 1.12})$$

Bu denkleme göre ÇBMG'nin değeri, pozitif ya da negatif olabilir. Negatif ÇBMG değerine sahip maddeler kolay olur ve kartezyen düzlemin üçüncü çeyreğinde yer alırlar. Buna karşılık pozitif ÇBMG değerine sahip maddeler nispeten daha zordur ve kartezyen düzlemin birinci çeyreğinde yer alırlar.

Tek boyutlu MTK'da ayırıcılık parametresi maddenin, testte elde edilen toplam puanın işaret ettiği yapı ile aynı yapıyı ölçüp ölçmediğini test etmek için kullanılır. Ayırıcılık parametresi bunun dışında θ ölçeğindeki farklı yetenek düzeylerinde, ölçmenin kararlılığı hakkında da bilgi verir. Tek boyutlu MTK'da güçlük parametresi, θ ölçeğindeki madde karakteristik eğrisinin (MKE) en dik olduğu noktadır. Bu nokta MKE'nin bükülme noktasına denk gelir. ÇBMG de buna benzer biçimde madde karakteristik yüzeyinin en dik olduğu bükülme noktasına denk gelir. Bu sebeple ÇBMG θ uzayında maddenin en ayırıcı olduğu bölgeye işaret eder (Reckase ve McKinley, 1999).

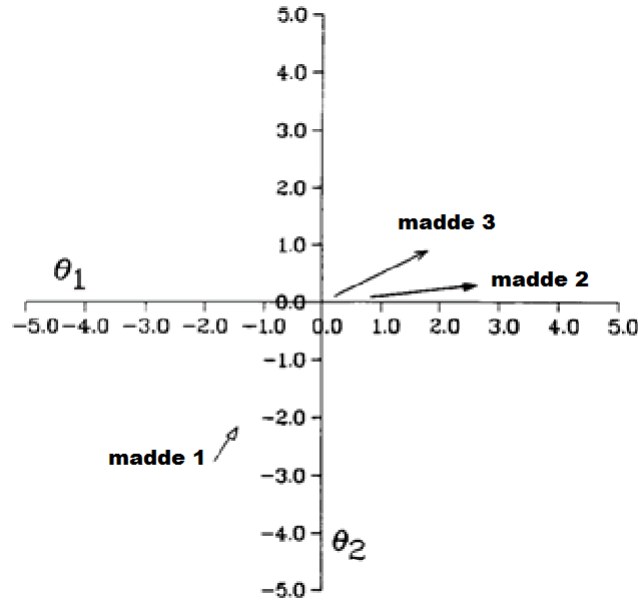
ÇBMTK'da maddeler kartezyen düzlem üzerinde madde vektörleri ile temsil edilebilirler. Her bir madde vektörü orijinden geçen bir doğru üzerindedir. Vektörün yönü pozitif θ_1 eksenine ile yaptığı açı olarak tanımlanır. Bir i maddesinin vektörünün yönü;

$$\alpha_i = \arccos \frac{\alpha_{i1}}{\text{ÇBMA}_i} \quad (\text{Eşitlik 1.13})$$

eşitliği ile hesaplanır. θ_1 eksenine daha yakın konumlanan maddeler öncelikli olarak θ_1 yeteneğini, θ_2 eksenine daha yakın konumlanan maddeler ise öncelikli olarak θ_2 yeteneğini ölçmektedir. İki yetenek eksenine ile 45° 'lik açı yapan maddeler ise her iki yeteneği de eşit düzeyde ölçmektedir (Ackerman, 1994; Ackerman, Gierl and Walker, 2003).

Vektörel gösterimde c parametresinin değeri, vektör okunun şekli ile verilmektedir. Tahmin parametresi 0,1'den küçük maddeler açık ok ile; 0,1 ile 0,2 arasında olan maddeler kapalı ok ile; 0,2'den büyük olan maddeler dolgulu kapalı ok ile gösterilir.

Bir kartezyen düzlem üzerinde birden fazla maddenin vektörleri çizilebilir. Bu özellik maddelerin yorumlanmasını ve karşılaştırılmasını kolaylaştırır. İki boyutlu telafisel modeldeki 3 maddenin (madde 1, madde 2 ve madde 3) vektörleri şekil 1.7'deki gibidir.



Şekil 1.7. Madde Vektörleri (Ackerman, 1996'dan alınmıştır).

Şekil 1.7'de yer alan vektörleri inceleyerek maddeler hakkında;

- ✓ Buldukları bölgeye göre en kolay maddenin madde 1 ve en zor maddenin madde 2 olduğu,
- ✓ Uzunluklarına göre en uzun vektöre sahip madde 3'ün en ayırıcı, en kısa vektöre sahip madde 1'in en az ayırıcı olduğu,
- ✓ Oklarının şekline göre c parametresinin madde 3'te 0,1'den küçük, madde 1'de 0,1 ile 0,2 arasında ve madde 2'de 0,2'den büyük kestirildiği yorumları yapılabilir.

Ayrıca madde vektörleri incelenerek hangi yeteneklerde daha ayırıcı maddelerin yer aldığı, madde güçlüğünün ilgililenen başat yetenek ile ilgili olup olmadığı, tahmin parametresinin güçlük, kapsam veya yetenek bileşiminin bir fonksiyonu olup olmadığı hakkında çıkarımlarda da bulunulabilir (Ackerman, 1996).

MTK farklı test yapılarına özgü farklı modeller ve çıktılar sunan zengin bir kuramdır. KTK ile aynı alanlarda uygulanabilmesine rağmen çoğu zaman KTK'dan daha iyi sonuçlar sunmaktadır (Hambleton ve Jones, 1993; Harvey ve Hammer, 1999; Sinar ve Zickar, 2002; McDonald ve Paunonen, 2002). Temelleri 1980'lerde ortaya konan (Sympson, 1978; Doody-Bogan ve Hattie, 1981; Yen, 1983; akt: Ansley ve Forsyth, 1985) ÇBMTK ile ilgili yurtdışında yapılan çeşitli çalışmalar mevcuttur. Ancak ÇBMTK çerçevesinde yapılan yerli çalışmaların sayısı oldukça azdır. Yurt içinde sadece doktora düzeyinde tamamlanmış birkaç tez mevcuttur (Köse, 2010; Sünbül, 2011; Özkan, 2012, Koğar, 2014). Bu çalışmalarda yer alan kuram uygulamaları telafisel modeldeki iki boyutlu testler ile sınırlı kalmıştır. Ancak gerçekte iki boyuttan daha fazla boyut sayısına sahip testler de kullanılmaktadır. Benzer biçimde maddeler içi model olan telafisel modelin yanında özellikle geniş ölçekli sınavlarda çoklu-tek boyutlu model (basit yapılu testler) ile ölçmeler yapılmaktadır. Bu tür çok boyutlu testler ile çalışırken doğru kestirimler yapabilmek için uygun örneklem büyüklüğü ve test uzunluğu koşullarının belirlenmesi gereklidir. Bu nedenle bu çalışma ile farklı boyutlardaki ve farklı yapı özelliklerindeki testlerin MTK ve ÇBMTK çerçevesinde ele alınmasına, ele alınan yapılara ilişkin farklı örneklem büyüklüğü ve test uzunluğu sınamalarının yapılarak kuramın değişen koşullar altında nasıl davrandığının incelenmesine gerek olduğu görülmüştür.

1.2. Araştırmanın Amacı ve Önemi:

Araştırmanın amacı, tek boyutlu testler ile basit ve karmaşık yapıda olan iki ve üç boyutlu testlerin farklı örneklem büyüklüğü ve farklı test uzunluğu koşullarında madde ve birey parametreleri kestirimlerine ilişkin hataların ortalama karekökünü (RMSE) ve yanlılık durumlarını belirlemek ve elde edilen parametre kestirimleri ile gerçek parametreler arasındaki ilişkiyi incelemektir. Elde edilen sonuçlara bağlı olarak çalışılan test yapılarında minimum hataya ve yanlılığa sahip kestirimler yapabilmek için gerekli örneklem büyüklüğü ve test uzunluğuna ilişkin önerilerin geliştirilmesi amaçlanmaktadır.

Çok boyutlu madde tepki kuramına ilişkin olarak yapılan çalışmalar genel olarak iki boyutlu testler üzerine yoğunlaşmıştır. Yapılan pek çok çalışmada basit ve karmaşık yapı ele alınmıştır. Fakat aynı çalışma içerisinde iki modeli de ele alan çalışmalara pek rastlanmamaktadır. Türkiye’de yapılan ÇBMTK çalışmaları iki boyutlu testleri ve karmaşık yapıyı ele almakta; üç boyutlu yapıları ele alan ÇBMTK çalışması bulunmamaktadır. Bu doğrultuda basit ve karmaşık yapılu üç boyutlu testler ile ilgili bir çalışma da henüz mevcut değildir. Bu çalışma MTK ve ÇBMTK kapsamında bir, iki ve üç boyutlu testleri ele alması bakımından önem taşımaktadır. Ele alınan çok boyutlu testlerin karmaşık yapının yanında basit yapıyı da kapsamasının araştırmayı önemli kıldığı ve basit yapı olarak adlandırılan çoklu-tek boyutlu modelin tanıtılmasının, farklı boyut sayısı, örneklem büyüklüğü, test uzunluğu koşullarında incelenmesinin alan yazına katkı sağlayacağı düşünülmektedir. Türkiye’de yapılan çalışmalarda farklı kestirim yöntemleri için farklı yazılımlar kullanılmıştır, ancak IRTPRO yazılımı kullanılarak yapılmış bir çalışma bulunmamaktadır. Bu çalışmada IRTPRO 2.1 kullanıldığından, çalışmanın sonucunda yazılım ile ilgili olarak araştırmacılara yönelik öneri geliştirilmesinin, araştırmanın önemini arttırdığı düşünülmektedir.

1.3. Problem Cümlesi:

Bu araştırmanın problem cümlesi aşağıdaki gibi ifade edilmiştir:

“Tek boyutlu testlerden madde tepki kuramıyla yapılan madde ve birey parametresi kestirimleri ile basit yapılu ve karmaşık yapılu iki ve üç boyutlu testlerden

çok boyutlu madde tepki kuramıyla yapılan madde ve birey parametresi kestirimleri örneklem büyüklüğü ve test uzunluğundan nasıl etkilenmektedir?”

Araştırma problemi aşağıda yer alan küçük problemlere (alt problemler) bölünerek çözümlenmeye çalışılacaktır.

1.3.1. Alt Problemler:

Bu problem cümlesine ilişkin oluşturulan alt problemler aşağıdaki gibidir;

1. Tek boyutlu testler için;

1.1. Örneklem büyüklüğü 1000, 2000 ve 4000; test uzunluğu 12 ve 48 olarak ele alındığında madde ve birey parametreleri kestirimlerine ilişkin standart hata ortalamasının karekökü, yanlılık, gerçek ve kestirilen parametreler arasındaki korelasyonlar nasıl değerler almaktadır?

2. İki boyutlu testler için;

2.1. Karmaşık yapıli testlerde örneklem büyüklüğü 1000, 2000 ve 4000; test uzunluğu 12 ve 48 olarak ele alındığında madde ve birey parametreleri kestirimlerine ilişkin standart hata ortalamasının karekökü, yanlılık, gerçek ve kestirilen parametreler arasındaki korelasyonlar nasıl değerler almaktadır?

2.2. Basit yapıli testlerde örneklem büyüklüğü 1000, 2000 ve 4000; test uzunluğu 12 ve 48 olarak ele alındığında madde ve birey parametreleri kestirimlerine ilişkin standart hata ortalamasının karekökü, yanlılık, gerçek ve kestirilen parametreler arasındaki korelasyonlar nasıl değerler almaktadır?

3. Üç boyutlu testler için;

3.1. Karmaşık yapıli testlerde örneklem büyüklüğü 1000, 2000 ve 4000; test uzunluğu 12 ve 48 olarak ele alındığında madde ve birey parametreleri kestirimlerine ilişkin standart hata ortalamasının karekökü, yanlılık, gerçek ve kestirilen parametreler arasındaki korelasyonlar nasıl değerler almaktadır?

3.2. Basit yapıli testlerde örneklem büyüklüğü 1000, 2000 ve 4000; test uzunluğu 12 ve 48 olarak ele alındığında madde ve birey parametreleri

kestirimlerine ilişkin standart hata ortalamasının karekökü, yanlılık, gerçek ve kestirilen parametreler arasındaki korelasyonlar nasıl değerler almaktadır?

1.4. Sayılılar:

Araştırmanın verilerinin üretilmesinde kullanılan madde ve birey parametrelerinin dağılım özelliklerinin gerçek durumları örneklediği varsayılmaktadır.

1.5. Sınırlılıklar:

Araştırma, ele alınan yapılar, koşullar ve kullanılan yazılımlar ile sınırlıdır.

1.6. Tanımlar:

Basit Yapı: Testin bütünü birden fazla yetenek ile ilişkili iken, testin her biri tek boyutlu farklı alt testlerden oluştuğu yapı.

Karmaşık yapı: Hem testin bütünü hem de testte yer alan maddelerin birden fazla yetenek ile ilişkili olduğu yapı.

2. İLGİLİ ARAŞTIRMALAR

Bu bölümde, ilgili alan yazında yapılmış çalışmaların özetlerine yer verilmiştir.

Ansley ve Forsyth (1985); yaptıkları çalışmada çok boyutlu bir veri seti üzerinde tek boyutlu kestirimler yapmışlardır. Bu doğrultuda boyutlar arası korelasyonları sırasıyla 0,0; 0,3; 0,6; 0,9 ve 0,95 olan telafisel olmayan modelde iki boyutlu veri setleri üretmiş ve parametrelerini kestirmişlerdir. Sonuçta tek boyutlu modelden kestirilen ayırıcılık parametrelerinin çok boyutlu modelden kestirilen ayırıcılık parametrelerinin ortalamasına yakın olduğu bulgusuna ulaşmışlardır. Ayrıca araştırmanın sonuçlarına göre güçlük parametresi tek boyutlu modelde çok boyutlu modeldekinden daha yüksek değerler almıştır.

Way, Ansley ve Forsyth (1988), çalışmalarında telafisel ve telafisel olmayan modele ilişkin iki boyutlu veri setlerini tek boyutlu MTK ile analiz etmişlerdir. Her iki model için de yetenekler arası korelasyonu 0,0'dan 0,95'e kadar değişen veri setleri üretmiş, üretilen veri setlerinin parametre kestirimlerini yapmışlardır. Tek boyutlu parametre kestirimlerinden elde edilen a parametresi, telafisel olmayan modeldeki a1 ve a2 parametrelerinin ortalaması olarak kestirilmiştir. Buna karşın telafisel modelde a parametresi a1 ve a2'nin toplamı olarak kestirilmiştir. Telafisel olmayan modelde tek boyutlu MTK ile kestirilen d parametresi d1'den daha büyük; telafisel modelde tek boyutlu MTK ile kestirilen d parametresi ise d1 ve d2'nin ortalaması olarak kestirilmiştir. Her iki model için de kestirilen θ parametreleri θ_1 ve θ_2 'lerin ortalamasına yakın değerlerdedir. Telafisel olmayan modelde bu ilişki yetenekler arasındaki korelasyon ile doğru orantılıdır. Ancak telafisel modelde benzer bir duruma rastlanmamıştır.

Ackerman (1989), yaptığı çalışmada telafisel ve telafisel olmayan çok boyutlu modellerdeki maddelerin tek boyutlu kalibrasyonunu yapmıştır. Bu amaçla telafisel ve telafisel olmayan model için boyutlar arası korelasyonları sırasıyla; 0,0; 0,3; 0,6 ve 0,9 olan iki boyutlu veri setleri oluşturulmuştur. Veri setlerinde yer alan maddelerin kalibrasyonları için BILOG ve LOGIST programlarından faydalanılmıştır. Sonuçta madde kalibrasyonunda madde güçlüğü ve boyutluluğun LOGIST programını daha az, BILOG programını daha fazla etkilediği görülmüştür. Ayrıca veri setlerindeki yetenekler arası korelasyonlar arttıkça cevapların tek boyutluluğa yaklaştığı görülmüştür.

Choi (1996), farklı kalibrasyon programlarını parametre kestirimleri açısından incelemiştir. Bu doğrultuda öncelikle çok boyutlu aşamalı tepki modeline dayalı veri seti üzerinde kestirimler yapmıştır. Üç ve beş kategorili maddelerden oluşturduğu farklı veri setlerini POLYFACT ile analiz etmiştir. Daha sonra iki kategorileştirme yöntemine dayalı olarak veri setindeki çok kategorili maddeleri iki kategorili maddelere dönüştürmüş ve elde ettiği veri setlerini çok boyutlu iki parametrelili lojistik modele uygun olarak NOHARM ve TESTFACT ile kalibre etmiştir. Doğrudan ve dolaylı üç kalibrasyon yöntemini parametre doğruluğu açısından farklı test uzunluğu, kategori sayısı, test yapısı, madde sayısı-örneklem büyüklüğü oranı koşullarında karşılaştırmalı olarak ele almıştır. Hem simülasyon veri seti hem de gerçek veri seti kullandığı çalışmanın sonucunda tek boyutluluk varsayımı ihlalinde çok kategorili puanlanan maddelerden oluşan testlerin parametre kestirimlerini yapmak için çok boyutlu aşamalı tepki modelinin kullanışlı olduğu ve bu modelin parametre kestirimini yüksek kararlılıkla yaptığı bulgusuna ulaşmıştır.

Bolt ve Lall (2003), örneklem büyüklüğü, boyutlar arası korelasyon ve madde sayısı farklılaşan veri setlerini çok boyutlu iki parametrelili lojistik model ve çok boyutlu gizil özellik modeli kapsamında karşılaştırmışlardır. Yapılan karşılaştırmalar sonunda madde sayısı ve örneklem büyüklüğü azalmasının RMSE değerinde artışa sebep olduğunu bulmuşlardır. Ayrıca çok boyutlu iki parametrelili lojistik modelden kestirilen a_1 , a_2 ve d parametrelerine ilişkin hata çok boyutlu gizil özellik modelinden elde edilen hatalara göre küçük değerler almıştır.

Zhang ve Stone (2004) çalışmalarında, iki boyutlu üç parametrelili telafisel modelin parametre kestirimini çeşitli yollarla yapmışlardır. Bu doğrultuda şans parametresinin sabitlenerek diğer parametrelerin kestirildiği dolaylı yöntemler ile tüm parametrelerin eş zamanlı ele alındığı doğrudan yöntemleri karşılaştırmışlardır. Kestirimlerde doğrusal olmayan faktör analizi için NOHARM, maksimum olabilirlik yöntemi için TESTFACT, maksimum olabilirlik genetik algoritması için ASSEST ve Markov Zinciri Monte Carlo yöntemi için Winbugs kullanmışlardır. Hem gerçek veri hem de simülasyon verisinin kullanıldığı çalışmada araştırmacılar dolaylı yöntemlerin şans parametresi kestirimi için MULTILOG'dan faydalanmıştır. Sonuçta doğrudan yöntemlerin tahmin ve eşik parametrelerini daha iyi kestirdiği bulgusuna ulaşmışlardır.

de la Torre ve Patz (2005) yaptıkları çalışmada boyut sayısının, her boyutta yer alan madde sayısının ve yetenekler arası korelasyonun, yetenek parametresinin kestirim

kararlılığına etkisini araştırmışlardır. Çalışmada yetenek kestirimleri için MCMC yöntemini kullanmışlardır. Çalışmanın koşullarını oluşturmak üzere boyut sayısı 2 ve 5; madde sayısı 10, 20 ve 30; yetenekler arası korelasyon 0,0; 0,4; 0,7 ve 0,9 olacak şekilde çaprazlamalar yapılmıştır. Tüm koşullar için çok boyutlu 3PLM'den yararlanarak kestirimler yapılmıştır. Kestirilen parametreler ile gerçek parametreler arasındaki uyuma Pearson korelasyonu ve göreceli etkililiğe RMSE ile bakılmıştır. Sonuç olarak yetenekler arası korelasyon sıfırdan farklı olduğu durumda eş zamanlı çok boyutlu kestirimlerin tek boyutlu kestirimlerden daha kesin ve değişmez sonuçlar ürettiği görülmüştür.

Sheng ve Wikle (2007), yaptıkları çalışmada çoklu tek boyutlu ÇBMTK modeli ile tek boyutlu MTK modelini karşılaştırmışlardır. Bu amaçla simülasyon yoluyla dört veri seti üretilmiştir. Veri setlerinden ilki tek örtük özelliği ölçmekte, diğer üçü de θ 'lar arası farklı korelasyon (0,0; 0,6 ve 0,8) ile iki örtük özelliği ölçmektedir. Araştırmacılar her bir veri setine üç model uygulamışlardır. Birinci modelde tüm kestirimleri tek boyutlu MTK'ya dayalı gerçekleştirmiş, ikinci modelde her alt testi tek boyutlu MTK ile kestirmiş ve üçüncü modelde çoklu-tek boyutlu model ile kestirim yapmışlardır. Ayrıca gerçek veri seti kullanarak da tek boyutlu ve çok boyutlu kestirimler yapmışlardır. Simülasyon veri setinden elde edilen sonuçlara göre tek boyutlu MTK, veri seti tek bir örtük özelliği ölçtüğü durumda iyi çalışmaktadır. Bunun dışındaki veri setlerinde çoklu-tek boyutlu MTK iyi sonuçlar üretmiştir. Gerçek veri setinden elde edilen sonuçlara göre çoklu-tek boyutlu MTK'dan elde edilen uyum indeksleri tek boyutlu MTK'ya benzer sonuçlar vermiştir. Bu durumun sebebi boyutlar arası korelasyonun yüksek olması olarak sunulmuştur.

Yao ve Boughton (2007) çalışmalarında alt test puanlarının kestiriminin ve sınıflandırma oranının doğruluğunu incelemek amacıyla MCMC yöntemi ile madde ve birey parametrelerini kestirmişlerdir. Çalışmada geniş ölçekli bir testin madde parametrelerini kullanarak iki ve çok kategorili, örneklem büyüklüğü 1000, 3000, 6000 olarak, boyutlar arası korelasyonu 0,0; 0,1; 0,2; 0,3; 0,5; 0,7 ve 0,9 olarak değişen veri setleri oluşturmuşlardır. Analizlerin sonucuna göre alt test düzeyinde doğru tanılayıcı bilgi sağlamak için alt testlerin yüksek korelasyon göstermesi gerektiği veya çok boyutlu yaklaşımın kullanılması gerektiği belirtilmiştir.

Hartig ve Höhler (2009), yaptıkları çalışmada farklı ÇBMTK modellerine değinmiş ve bu modellerin eğitimsel değerlendirmelerdeki farklılıklarını çeşitli uygulamalar ile ele almışlardır. Çalışmada ÇBMTK'nın esnekliğini ve yararlarını göstermek

amacıyla üç uygulama senaryosu ele alınmıştır. Buna göre; tek boyutlu yapıyı ölçmek isterken ortaya çıkan ve istenmeyen çok boyutluluk durumu, yetenek boyutları arasındaki örtük kovaryans yapıları ve belirli test maddelerini çözmek için gerekli birden çok θ arasındaki etkileşim ele alınmış ve ampirik örnekler verilmiştir. Finch (2010), yaptığı çalışmada basit yapılı iki boyutlu modeldeki testlerin madde parametreleri kestirimlerini doğrulayıcı faktör analiz yöntemlerinden olan ağırlıklandırılmamış en küçük kareler, ağırlıklandırılmış en küçük kareler yöntemleri ve tek boyutlu MTK yaklaşımı ile kestirmiş, elde ettiği yanlılık ve standart hata değerlerini farklı örneklem büyüklüğü, test uzunluğu ve boyutlar arası korelasyon koşullarında karşılaştırmıştır. Çalışmada Monte Carlo yöntemi ile oluşturduğu simülasyon veri setini kullanmıştır. Verilerin analizinde en küçük kareler yöntemi için NOHARM, ağırlıklandırılmış en küçük kareler yöntemi için MPlus ve tek boyutlu MTK yaklaşımı için BILOG-MG'den yararlanmıştır. Çalışmanın sonucunda, örtük özelliklerin çarpık dağılım gösterdiği durumlarda NOHARM ve MPlus'ın yanlı sonuçlar ürettiği, büyük örneklerde yine bu iki program ile elde edilen ayırıcılık parametrelerinin yüksek kararlık gösterdiği sonuçlarına ulaşmıştır.

Köse (2010), doktora tezi çalışmasında geliştirdiği ve uyguladığı telafisel tipteki iki boyutlu test verilerini tek ve çok boyutlu madde tepki kuramı modelleri altında araştırmıştır. Bu doğrultuda madde ve yetenek parametreleri ile model-veri uyum değerlerini farklı değişkenler altında karşılaştırmış, veri grubuna uyan en iyi modeli ortaya koymayı amaçlamıştır. Analizler sonunda, ÇBMTK'nın madde ve yetenek parametrelerinde daha az hata içerdiği, daha iyi bir model veri uyumu verdiğini göstermiştir. Ayrıca örneklem büyüklüğünün yetenek parametreleri üzerinde fazla etkili olmadığı, ancak madde parametreleri üzerinde etkili olduğu bulgusuna ulaşılmıştır. Çalışmada örneklem büyüklüğü ve test uzunluğu etkilerinin birlikte ele alındığı koşullarda sadece çok boyutlu kurama dayalı kestirilen madde parametreleri üzerine olumlu bir etki görülmüş diğer kestirimlerde bir farklılık görülmemiştir.

Finch (2011), yaptığı çalışmada NOHARM ve tek boyutlu MTK'ya dayalı olarak basit yapılı olmayan (birden fazla örtük özellik ile ilgili) maddelerin güçlük ve ayırıcılığını kestirmiştir. Kestirme işleminde örneklem büyüklüğü, boyutlar arası korelasyon ve basit yapıdan uzaklaşma derecesi koşullarını manipüle etmiştir. Sonuçta yanlılık ve

standart hata deęerlerini referans alarak yorumlar yapmıřtır. Sonulara gre yanlılık ve hata basit yapıdan uzaklařtıkaa daha yksek deęerler almıřtır.

Snbl (2011), iki boyutlu veri setlerinde parametre deęiřmezlięini KTK, MTK ve BMTK erevesinde sınamıřtır. Bu doęrultuda boyutlar arası korelasyon (0,0; 0,3; 0,6 ve 0,8 řeklinde), ikinci boyutta yer alan madde sayısı, a_2 parametre ortalaması gibi kořulları deęiřtirerek 96 deneysel hcre elde etmiřtir. Her deneysel hcre iin 90.000 bireylik veri seti oluřturmuř ve eřitli byklklerde rneklemler ekerek analiz etmiřtir. Analiz sonularını standart hata ortalamalarından faydalanarak yorumlamıřtır. Sonuta KTK iin parametre deęiřmezlięinin btn kořullarda saęlandıęı, MTK iin hem a hem b parametresinde parametre deęiřmezlięinin byk oranda saęlandıęı bulgularına ulařmıřtır. BMTK erevesinde incelenen parametre deęiřmezlięinin de byk oranda saęlandıęını belirtilmiřtir.

Lee (2012) MCMC simlasyonu kullanarak yaptıęı alıřmasında BMTK ile parametrelerin kestirim doęruluęu zerine faktrlerin etkileřim etkisini arařtırmıřtır. rtk zellik daęılımının korelasyonunun ve arpıklıęının gz nnde bulundurulmadıęı modelin bulgularına gre yksek boyut sayısında daha byk rneklemlerin gerekli olduęunu (6 boyut iin 2000, 3 boyut iin 1000) gstermiřtir. alıřma rtk zellik zerine korelasyon veya arpıklık gibi ek bir faktr eklendięinde rneklem byklęnn arttırılmasının parametrelerin kestirim kararlılıęını arttırmada yardımcı olmayacaęını gstermiřtir. rtk zelliklerin iliřkili olduęu durumda ve normal olmayan daęılımın normal daęılıma dnřtrlmesinde alternatif bir BMTK modelinin denemesi gerektięini nermiřtir. Genel olarak boyut sayısı arttıkaa kestirimlerin yanlılıęının arttıęını, rtk yapılar baęımsız ve normal daęıldıęı durumda boyut sayısı yksek olsa dahi parametre kalibrasyonunda daha dřk yanlılık oluřacaęını belirtmiřtir. rtk zelliklerin arpık veya normal olmayan daęılıma sahip veya iliřkili olduęundan řphelenildięinde yanlılıęı dřrmek iin rneklem sayısını arttırmanın yeterli olmayacaęını, eř zamanlı olarak madde sayısını da arttırmanın veya bireylerin geniř bir yetenek ranjından rneklenmesinin gerektięini ifade etmiřtir.

zkan (2012), doktora tez alıřmasında KTK, MTK ve BMTK'ya dayalı olarak BBS matematik ve Trke testlerinin madde parametrelerini, gvenirlik deęerlerini ve bařarı puanlarını kestirmiřtir. Her iki testten de elde ettięi sonulara gre BMTK'ya dayalı hesaplanan gvenirlik daha yksek deęer almıřtır. BMTK'ya

dayalı kestirilen yetenek parametreleri MTK ile kestirilen yetenek parametrelerinden ve KTK ile elde edilen test puanlarından daha düşük hataya sahip çıkmıştır. Ayrıca KTK'ya dayalı elde edilen test puanları ile MTK ve ÇBMTK'ya dayalı kestirilen yetenek puanları arasında yüksek ve anlamlı ilişki bulunmuştur.

Md Desa (2012) bi-faktör çok boyutlu madde tepki kuramını alt puanların kestirilmesi, güvenilirliği ve sınıflandırılması amaçlarıyla ele almıştır. Çalışmasında bi-faktör yapıyı telafisel ve kısmi telafisel modellerde ele almıştır. Monte Carlo çalışması yürüttüğü araştırmada 1500 birey, iki farklı test uzunluğu (30 ve 60 madde) ve 5 düzeyde madde güçlük değeri koşullarını ele almıştır. MCMC- Gibbs örnekleme yöntemi ile modellerden elde ettiği veri setlerinin alt puanlarını kestirmiştir. Elde edilen sonuçlar kestirimin kararlılığına, uyuma, alt puan güvenilirliğine dayalı olarak tartışılmıştır. Buna göre hem bi-faktör telafisel modele hem bi-faktör telafisel olmayan modele ilişkin alt test puanlarının güvenilir olduğu, düşük yanlışlık ve EAP alt puanlarında hata varyanslarında azalma sağlandığı belirtilmiştir. Bu bulgulara ek olarak bi-faktör telafisel modelin spesifik yetenekler ile başat yeteneği ayırmada daha iyi performans sergilediği sonucuna ulaşılmıştır.

Zhang (2012), çalışmasında marjinal maksimum olabilirlik metoduna dayalı tek boyutlu ve çok boyutlu yaklaşımları karşılaştırmıştır. Bu karşılaştırmada simülasyona dayalı olarak oluşturduğu veri setlerini kullanmış; örneklem büyüklüğü, madde sayısı ve boyutlar arası korelasyonlar açısından karşılaştırmalar yapmıştır. Sonuçlara göre madde sayısı az olduğunda çok boyutlu yaklaşımın madde parametrelerini daha kesin bir biçimde kestirdiğini ifade etmiştir. Tek boyutlu yaklaşımların ise madde sayısı yeterince fazla olduğunda benzer davrandığı bulgusuna ulaşmıştır. Ayrıca teorik ve pratik olarak basit yapı varsayımının ihlali de araştırılmıştır. Buna göre veri seti basit yapıya değilken basit yapıya gibi kalibrasyon yapıldığında model yanlış kestirilmekte ve θ 'lar arası korelasyon için de üst kestirimde bulunmaktadır.

Koçar (2014) yaptığı doktora tez çalışmasında örneklem büyüklüğü, test uzunluğu ve boyutlar arası korelasyonu değişen basit ve iki boyutlu yapıların madde parametreleri, maddelere ait model veri uyumları ve teste ait model veri uyumlarını tek boyutlu parametrik olmayan, tek boyutlu parametrik ve çok boyutlu madde tepki kuramı ile kestirmiş ve karşılaştırmıştır. Çalışmada simülasyon verisi kullanılmış ve

her bir desen için 20 tekrar yapılmıştır. Bulgulara göre birey ve madde sayısının artması üç MTK uygulamasında da model uyumlarını arttırmış, boyutlar arası korelasyon ile veri uyumu arasında örüntü saptanmamıştır. Ayrıca tüm simülasyon düzeneklerinde parametrik olmayan MTK'nın model veri uyumu oldukça düşük hata ile hesaplanmıştır. Simülasyon koşulları b parametresi üzerinde etkili olmamakla beraber, tek boyutlu ÇBMTK'daki a_1 ve a_2 parametreleri örneklem büyüklüğü; MTK'daki a parametresi ise hem örneklem büyüklüğü hem test uzunluğundaki artış ile azalmıştır.

İlgili çalışmalar incelendiğinde 1980'lerden beri MTK çalışmalarının yapıldığı görülmektedir. Son yıllarda çalışmalar ÇBMTK üzerine yoğunlaşmış, hem MTK ve ÇBMTK'nın birlikte ele alındığı hem de sadece ÇBMTK'yı konu edinen çalışmalar artmıştır. Yapılan parametre kestirim çalışmalarında birey ve madde parametrelerinin kestirim kararlığı örneklem büyüklüğü, test uzunluğu, boyutlar arası korelasyon, yetenek dağılımı gibi özelliklerin manipüle edilmesi ile belirlenmeye çalışılmıştır. Türkiye'de yapılan çalışmalar sayıca az olmakla beraber, çalışmalarda MTK ve ÇBMTK birlikte ele alınmıştır. Ancak bu çalışmalarda iki boyutlu yapıların üstüne çıkılmamıştır. Bu çalışmanın, hem MTK hem ÇBMTK'yı kapsamaması; ele alınan çok boyutlu testlerde iki boyutun yanında üç boyutlu testlerin de yer alması; basit ve karmaşık yapıları çok boyutlu testlerde parametre kestirim kararlığını örneklem büyüklüğü ve test uzunluğu bakımından incelemesi bakımından alan yazına katkı sağlaması düşünülmektedir.

3. YÖNTEM

Bu bölümde araştırmanın türü, araştırmanın deseni ve verilerin analizi başlıklarına yer verilmiştir.

3.1. Araştırmanın Türü

Bu çalışma, farklı örneklem büyüklüğü, test uzunluğu ve boyut sayısı koşulları altında tek ve çok boyutlu madde tepki kuramının farklı modellerinden yapılan kestirimlere ilişkin elde edilen hata, yanlılık ve yapılan kestirimler ile gerçek parametrelerin korelasyonlarını incelemesi, bu doğrultuda bir kuramın söz konusu modellerini test etme amacını taşımasından dolayı temel araştırma niteliğindedir.

3.2. Araştırmanın Deseni

Araştırmada simülasyon yoluyla oluşturulan veri setleri kullanılmıştır. Simülasyonlar 1990'lı yıllarda bilgisayar teknolojilerinde meydana gelen gelişmelerin sosyal bilimlerde nicel araştırmalar için yeni bakış açıları oluşturması amacıyla yapılmaya başlanan bilgisayar ortamındaki model deneyleridir. Simülasyonlar ile çok çeşitli sosyal kavramları temsil eden çeşitli bilgisayar modellerinin kurulması ve bu modellerin çeşitli koşullarda nasıl davranacağına, farklı değişkenlerin modeli nasıl etkileyeceğinin test edilmesi mümkün hale gelmiştir (Gilbert, 1999). Bu araştırmada yürütülen simülasyon çalışmasında MTK ve ÇBMTK ile yapılan parametre kestirimleri üzerine örneklem büyüklüğü ve test uzunluğunun etkileri sınanmıştır.

Araştırmanın verilerini üretmek için öncelikle alt problemlere uygun olarak madde ve birey parametreleri oluşturulmuştur. Bunun için tek boyutlu MTK'da madde sayısı kadar ayırt edicilik ve güçlük parametresi üretilmiştir. İki ve üç boyutlu basit yapıları ÇBMTK modellerinde her bir madde tek bir boyut ile ilişkili olduğu için madde sayısı kadar ayırt edicilik ve güçlük parametresi üretilmiştir. Üretilen her bir ayırt edicilik parametresi bir madde ve bir boyut ile ilişkilendirilmiş, ilişkilendirildiği maddenin diğer boyutlardaki ayırt ediciliği sıfır olarak alınmıştır. İki ve üç boyutlu karmaşık yapıları ÇBMTK modellerinde veri üretimi ve kestirimler telafisel modele göre yapılmıştır. Bu modelde her bir madde tüm boyutlar ile ilişkili olduğundan her bir madde için boyut sayısı kadar ayırt edicilik parametresi ile bir tane güçlük parametresi üretilmiştir. Böylece karmaşık yapılarda boyut sayısı x madde sayısı kadar ayırt edicilik parametresi ve madde sayısı kadar güçlük parametresi elde

edilmiştir. Yapılan bir çalışmada simülasyon verisi kullanmanın avantajı gerçek madde ve yetenek parametrelerinin bilinmesidir. Böylece araştırmacı gerçek değerler ile kestirilen değerler arasında sağlıklı karşılaştırmalar yapabilir. Ancak dikkat edilmesi gereken iki önemli koşul vardır. Bunlardan ilki araştırmacı veri setini oluşturmak için uygun bir yöntemi seçmelidir. İkinci koşul ise oluşturulan simülasyon veri setlerinin test maddelerine verilecek gerçek cevapları temsil ettiğinden emin olunması gerektiğidir (Way, Ansley ve Forstyh, 1988).

Yukarıda bahsedilen koşulları karşılamak amacıyla verilerin üretimi sırasında madde parametrelerinin gerçekte var olan bir durumu yansıtacak şekilde olmasına özen gösterilmiş, bunun için parametrelerin üretilmesinde bazı dağılım özellikleri göz önünde bulundurulmuştur. Buna göre ayırt edicilik parametrelerinin (a_i) üretilmesinde ortalaması 0,5 ve standart sapması 0,4 olan lognormal dağılım, güçlük parametrelerinin (d_i) üretilmesinde standart normal dağılım kullanılmıştır. Böylece istenen dağılım özelliklerinde ve istenen koşula ilişkin madde parametreleri üretilmiştir. Üretilen madde parametrelerinden faydalanarak her bir desene göre 2 parametrelili lojistik modelde iki kategorili veri seti ve her desen için yetenek parametresi seti oluşturulmuştur. Yetenek parametrelerinin oluşturulmasında standart normal dağılımdan faydalanılmıştır. Madde ve birey parametreleri ile iki kategorili verilerin üretilmesinde SimuMIRT (Yao, 2003) yazılımından faydalanılmıştır.

Çalışmada tek ve çok boyutlu yapılar sabit olarak ele alınmıştır. Buna göre sabitler: tek boyutlu, iki boyutlu karmaşık, iki boyutlu basit, üç boyutlu karmaşık ve üç boyutlu basit yapı şeklindedir. Test uzunluğu ve örneklem büyüklüğü araştırmacının bağımsız değişkenleri olarak ele alınmıştır.

Test uzunluğu: Bağımsız değişkenlerden biri olan test uzunluğunun parametre kestirimi üzerine etkisini test etmek amacıyla kısa ve uzun testleri temsil edecek şekilde iki test uzunluğu koşulu ele alınmıştır. Bu doğrultuda madde sayısının ele alınan boyut sayılarının tam katı olması, basit yapı testlerin her bir alt testinde en az 3 maddenin yer alması koşullarını yerine getirecek biçimde kısa testleri temsil etmek üzere 12, uzun testleri temsil etmek üzere 48 maddeli testlerden yararlanılmıştır. Basit yapı testlerin her bir alt testinde eşit sayıda madde yer almıştır.

Örneklem büyüklüğü: Kestirimler üzerine örneklem büyüklüğünün etkisini test etmek ve araştırmanın sonucunda örneklem büyüklüğü önerisinde bulunmak amacıyla farklı örneklem büyüklükleri ile çalışmıştır. Literatürde çok boyutlu yapılarda en az 1000 kişilik örneklem ile çalışılması önerilmektedir (Bolt ve Lall, 2003; Yao ve Boughton, 2007; Lee 2012). Bu sebeplerle çalışmada 1000, 2000 ve 4000 olmak üzere 3 örneklem büyüklüğü koşulu ele alınmıştır.

Sonuç olarak çalışmada tek boyutlu, iki boyutlu basit yapılı, iki boyutlu karmaşık yapılı, üç boyutlu basit yapılı ve üç boyutlu karmaşık yapılı olan, test uzunluğu 12 ve 48; örneklem büyüklüğü 1000, 2000, 4000 şeklinde değişen 5 MTK modeli * 2 test uzunluğu * 3 örneklem büyüklüğü olmak üzere 30 desen oluşturulmuştur.

Araştırmada her bir desene dair tekrar (replikasyon) yapılmıştır. Bu tip simülasyon çalışmalarında tekrar örneklem büyüklüğü ile eş anlamlıdır. MTK çalışmalarında tekrar sayısı kestirilen parametrelerin örnekleme hatasının varyansını ve simülasyon sonuçlarına dayalı istatistiklerin yeterli güce sahip olmasını etkiler. Stone (1993) çalışmasında 10, 25, 50 ve 100 tekrar ile farklı örneklem büyüklüğü, test uzunluğu ve dağılım özelliklerine sahip veri setleri üretmiş ve bu koşullardan hangilerinin hatalar üzerinde etkili olduğunu incelemiştir. Sonuçta tekrar arttıkça etkinin arttığını, birey sayısı etkisini test ettiğinde de 25 tekrar koşulunda etkinin 1.0 olduğunu belirtmiştir. Bu sebeple en az 25 tekrar kullanmanın gerekli olduğunu ifade etmiştir (Harwell vd. 1996). Bu çalışmada da alan yazın önerileri ve analizlerin uzun sürmesi gibi sınırlılıklar göz önünde bulundurularak her bir desene ait başlangıç parametreleri (madde ve birey parametreleri) sabit tutulmak üzere, 25 tekrar yapılmıştır. Sonuçta 30 desen * 25 tekrar olmak üzere 750 veri seti elde edilmiştir. Oluşturulan desenler tablo 3.1'de özet halinde verilmiştir.

Tablo 3.1: Araştırmanın Deseni

<i>Sabitler</i>	<i>Değişkenler</i>		<i>Veri seti</i>	
	<i>Yapı</i>	<i>Örneklem Büyüklüğü</i>		<i>Test Uzunluğu</i>
1		1000	12 madde	
			48 madde	
		2000	12 madde	
			48 madde	
		4000	12 madde	
			48 madde	
2 Boyutlu Karmaşık Yapı	1000	12 madde		
		48 madde		
	2000	12 madde		
		48 madde		
	4000	12 madde		
		48 madde		
2 Boyutlu Basit Yapı	1000	12 madde		
		48 madde		
	2000	12 madde		
		48 madde		
	4000	12 madde		
		48 madde		
3 Boyutlu Karmaşık Yapı	1000	12 madde		
		48 madde		
	2000	12 madde		
		48 madde		
	4000	12 madde		
		48 madde		
3 Boyutlu Basit Yapı	1000	12 madde		
		48 madde		
	2000	12 madde		
		48 madde		
	4000	12 madde		
		48 madde		

3.3. Verilerin Analizi

Araştırmada simülasyon yoluyla üretilen veri setleri madde ve birey parametrelerinin kestirilmesi amacıyla analiz edilmiştir. 30 desen için üretilen 750 veri setinin madde ve birey parametreleri test yapısına uygun betikle IRTPRO 2.1 yazılımından faydalanarak kestirilmiştir. IRTPRO 2.1 yazılımında kestirimler Bock-Aitkin Expectation-Maximization (BA-EM), Adaptive Quadrature ve Metropolis-Hastings Robbins-Monro (MH-RM) algoritmaları ile yapılabilmektedir. Ancak MH-RM

algoritmasının kullanımı boyut sayısının iki veya üçü aştığı durumlarda önerilmekte, Adaptive Quadrature algoritmasıyla yapılan birey parametresi kestirimlerinde her bir iterasyonda quadrature nod sayısı değişmektedir (IRTPRO guide, 2011). Çalışmada ele alınan veri setlerinde boyut sayısının üçü geçmemesi ve quadrature nod sayılarının sabit tutularak birey parametrelerinin kestirilmek istenmesi sebebiyle, analizler BA-EM algoritmasından faydalanarak yapılmıştır. Analizlerin seri biçimde yapılabilmesi için yazılım batch modunda çalıştırılmıştır. Tek boyutlu testlerin analizlerinin tamamlanması kısa sürede gerçekleşmiştir. Çok boyutlu testlerin analizlerinde analiz için gereken süre özellikle karmaşık yapılarda madde ve boyut sayısının artması ile artmış, 3 boyutlu karmaşık yapılarda en yüksek değerini almıştır. Buna göre IRTPRO 2.1, 3 boyutlu karmaşık yapıları 48 madde ve 4000 örneklem büyüklüğü koşulundaki bir veri setinin madde ve birey parametrelerini kestirmek için Intel Core i7-4500 CPU 1.80GHz 2.40 GHZ 8GB (RAM) 64 bit işletim sistemi olan bir bilgisayarda ortalama 22 saat çalıştırılmıştır.

Analizlerin tamamlanmasından sonra madde ve birey parametrelerinin kestirilen ve gerçek değerlerinden faydalanarak GOR istatistikleri (kestirim iyiliği - goodness of recovery) hesaplanmıştır. GOR istatistikleri farklı koşulların etkisini araştırmak ve kestirim kararlılığını test etmek amacıyla parametreler için hesaplanan istatistiklerdir (Maris, 1999; Turhan, 2006). Bu çalışmada her bir parametre için hataların ortalama karekökü (RMSE) ve yanlışlık olmak üzere iki GOR istatistiği hesaplanmıştır. RMSE gerçek ile kestirilen parametre değerleri arasındaki farkın karesinin ortalamasının kareköküdür. Bu istatistik eşitlik 3.1'de yer alan formülden faydalanarak hesaplanmıştır.

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (\bar{\tau}_{nj} - \tau_{nj})^2}{N}} \quad (\text{Eşitlik 3.1})$$

RMSE formülündeki terimlerden N terimi testte yer alan parametre sayısını temsil etmektedir. Parametre sayısı, RMSE'nin madde parametreleri için hesaplandığı durumda testteki madde sayısı; birey parametreleri için hesaplandığı durumda testteki birey sayısı olarak alınmıştır. τ_{nj} n . maddeye ait j parametresinin gerçek değeri iken, $\bar{\tau}_{nj}$ n . maddenin j parametresinin kestirilen değeridir. Her bir desende her bir tekrar için parametrelerin RMSE değeri hesaplanmıştır. Daha sonra desende

yer alan 25 tekrardan elde edilen RMSE değerlerinin ortalaması, ilgili parametrenin RMSE değeri şeklinde raporlanmıştır.

Yanlılık, parametrenin kestirilen ortalama değeri ile gerçek değeri arasındaki farktır. Çalışmada yanlılık değerleri her bir madde ve her bir birey parametresi için hesaplanmıştır. Bunun için eşitlik 3.2'den yararlanılmıştır.

$$Yanlılık(\tau_j) = \left(\frac{\sum_{r=1}^R \bar{\tau}_{jr}}{R} \right) - \tau_j \quad (Eşitlik 3.2)$$

Eşitlik 3.2'de yer alan R tekrar sayısını temsil etmektedir. $\bar{\tau}_{jr}$ j parametresinin r. tekrardaki kestirilen değeri, τ_j j parametresinin gerçek değeridir. Formüle göre her bir parametrenin her tekrarda kestirilen değerlerinin ortalamasının gerçek parametreden farkı bulunmuştur. Bu sayede madde parametreleri için madde sayısı kadar, birey parametreleri için birey parametresi kadar yanlılık değeri elde edilmiştir. Elde edilen yanlılıkların madde/birey sayısına bölünerek hesaplanan ortalaması ilgili desendeki parametrenin yanlılığı olarak raporlanmıştır.

Araştırmada gerçek parametreler ile kestirilen parametreler arasındaki ilişkiyi incelemek üzere korelasyonlara bakılmıştır. Bunun için her bir tekrardan kestirilen parametreler ile gerçek parametreler arasındaki Pearson Momentler Çarpımı Korelasyon katsayısı hesaplanmıştır. Desende yer alan 25 tekrar için bu işlem yapılmış ve ortalaması hesaplanarak rapor edilmiştir. Korelasyon katsayılarının hesaplanmasına dair izlenen bu süreç hem madde hem de birey parametreleri için aynı şekildedir.

4. BULGULAR VE TARTIŞMA

Bu bölümde alt problemlere ilişkin bulgular tablo ve grafik şeklinde sunulmuş, desen koşulları göz önünde bulundurularak yorumlanmıştır. Bulguların yorumlanmasının ardından tartışma alt başlığında çalışmanın bulguları ilgili literatür göz önünde bulundurularak tartışılmıştır.

4.1. Birinci alt problem için bulgular

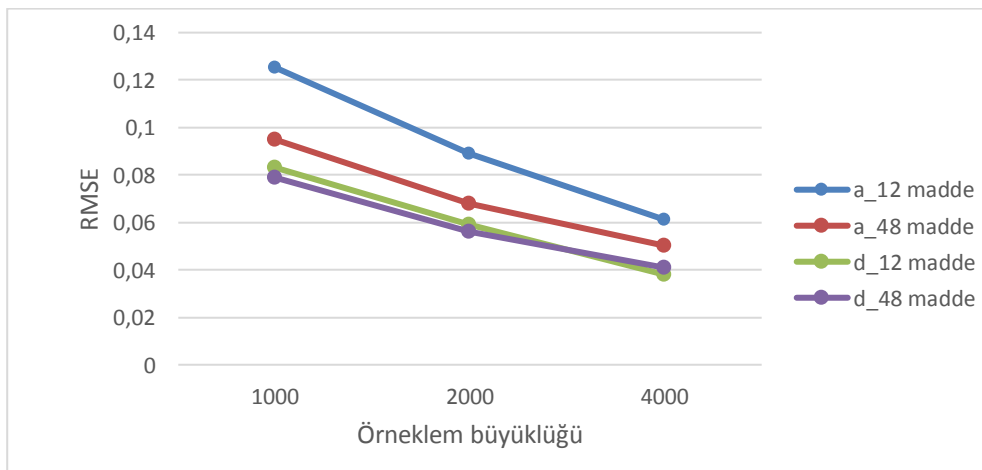
4.1.1. Tek boyutlu testlere ilişkin bulgular

Tek boyutlu testlerden elde edilen madde parametrelerine ait bulgular tablo 4.1'deki gibidir:

Tablo 4.1: Tek Boyutlu Testler İçin Madde Parametrelerine Ait Bulgular

Madde Sayısı	Örnekleme	a			D		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,125	-0,002	0,946	0,083	0,007	0,995
	2000	0,089	-0,012	0,970	0,059	-0,009	0,998
	4000	0,061	-0,013	0,986	0,038	-0,001	0,999
48 madde	1000	0,095	0,008	0,959	0,079	0,002	0,997
	2000	0,068	-0,006	0,978	0,056	-0,011	0,998
	4000	0,050	-0,010	0,999	0,041	-0,004	0,998

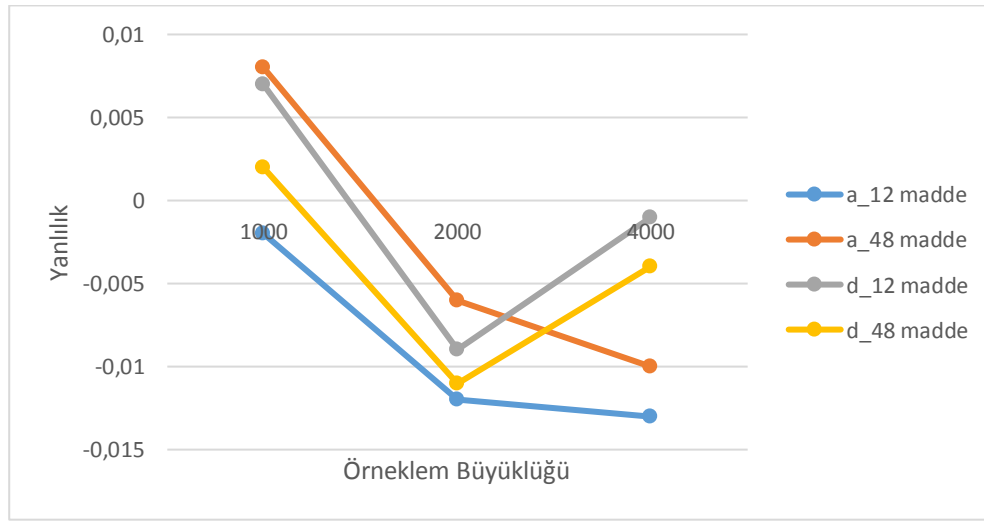
Tablo 4.1'de verilen değerler şekil 4.1-3'teki grafiklerde görselleştirilmiştir.



Şekil 4.1. Tek Boyutlu Testler İçin Madde Parametrelerine Ait RMSE Değerleri

Tek boyutlu testlerde madde parametrelerine ait RMSE değerlerine ilişkin sonuçlar tablo 4.1'de ve şekil 4.1'de verilmiştir. Bu sonuçlara göre tek boyutlu testlerden madde parametreleri için elde edilen RMSE değerleri 0,041 ile 0,125 arasında değerler almaktadır. Hem a hem d parametresine ait RMSE değerleri genel olarak birey sayısı ve madde sayısı artıkça azalış göstermektedir.

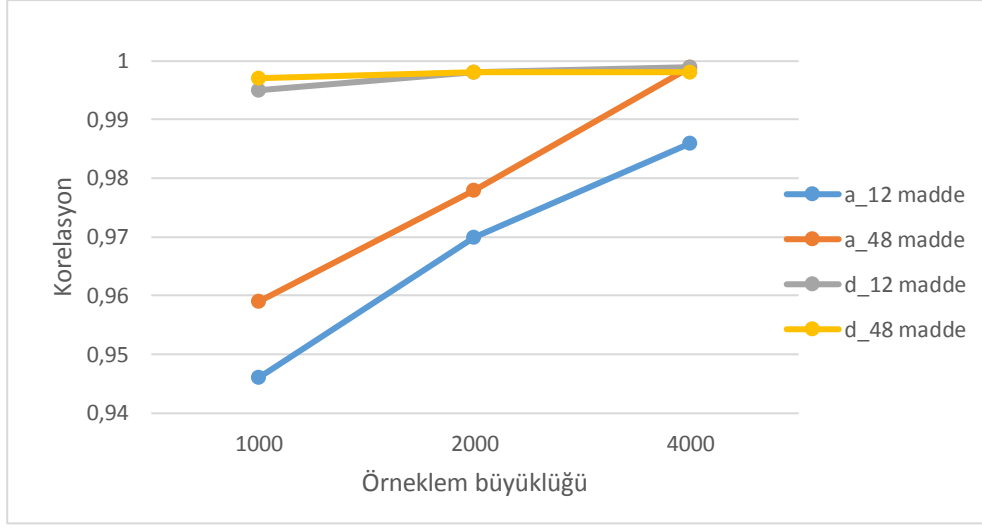
Grafik incelenecek olursa her iki test uzunluğu koşulunda da a parametresinin d parametresinden daha yüksek RMSE değerine sahip olduğu görülmektedir. Bu durum a parametresinin d parametresine göre daha yüksek hata ile kestirildiği, d parametresinin daha kararlı davrandığı şeklinde yorumlanabilir. Ancak d parametresine ait 48 maddelik test ve 4000 örneklem büyüklüğü koşulundaki RMSE değeri, 12 maddelik test ve 4000 örneklem büyüklüğü koşulundaki RMSE değerinden büyüktür. Buna rağmen iki RMSE arasındaki farkın 0,009 gibi oldukça küçük bir değer olması sebebiyle RMSE değerlerinin madde sayısı ve örneklem büyüklüğü arttıkça azaldığı söylenebilir.



Şekil 4.2. Tek Boyutlu Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri

Tek boyutlu testlerde madde parametrelerine ait yanlılık değerleri -0,013 ile 0,07 arasında değişmektedir. Hem a hem d parametresi için elde edilen yanlılık değerleri madde sayısı veya örneklem büyüklüğündeki artış ile düzenli bir örüntü göstermemiştir. 12 maddelik testte a parametresine ait yanlılık değerleri negatif yönlüdür ve örneklem büyüklüğü arttıkça artmıştır. 48 maddelik testte a parametresine ait yanlılık örneklem büyüklüğü arttıkça mutlak değerce önce azalmış daha sonra artmıştır. Her iki test uzunluğu koşulunda da d parametresine ait yanlılık

değerleri örneklem büyüklüğü 1000'den 2000'e çıktığında mutlak değerce önce artmış, örneklem büyüklüğü 4000 olduğu koşulda azalmıştır.



Şekil 4.3. Tek Boyutlu Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri

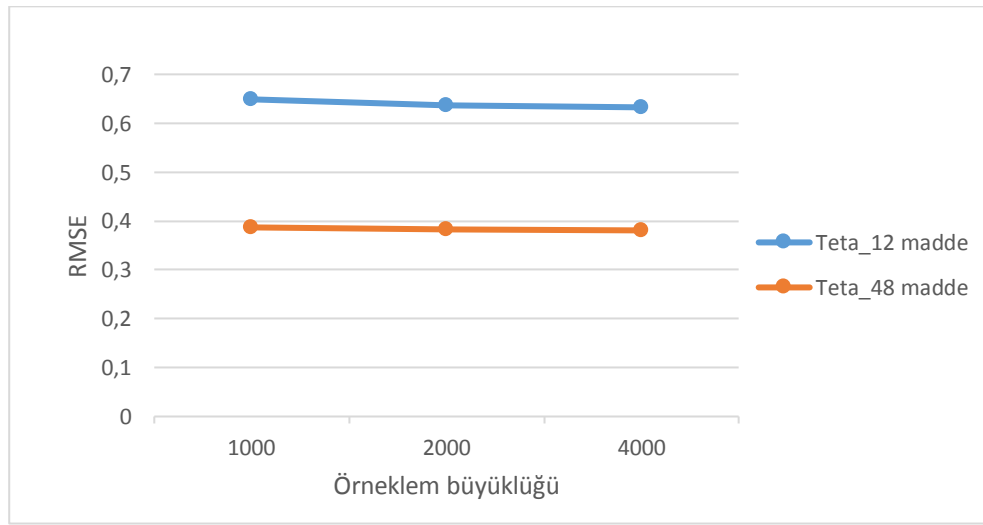
Tek boyutlu testlerde hem a hem d parametresi için kestirilen değerler gerçek parametre değerleri ile mükemmel düzeyde korelasyon göstermektedir. Tekrarlardan elde edilen korelasyon değerlerinin ortalaması 0,946 ve üzerindedir. Her iki test uzunluğunda da örneklem büyüklüğü arttıkça korelasyon değeri artış göstermiştir. d parametresi için elde edilen korelasyonlar aynı koşullardaki a parametresinden yüksek çıkmıştır. Ayrıca d parametresi için elde edilen korelasyon değerleri tüm örneklem büyüklüklerinde birbirine yakın değerlerdedir. Bu sonuca göre d parametresinin gerçek ve kestirilen değerlerinin korelasyonlarının örneklem büyüklüğünden ve test uzunluğundan çok fazla etkilenmediği söylenebilir. Bu bulgular d parametrelerinin daha kararlı kestirilmesinin bir sonucu olarak yorumlanabilir.

Tek boyutlu testlerden kestirilen birey parametrelerine ait RMSE, yanlılık ve korelasyon değerleri tablo 4.2'deki gibidir.

Tablo 4.2: Tek Boyutlu Testler İçin Birey Parametrelerine Ait Bulgular

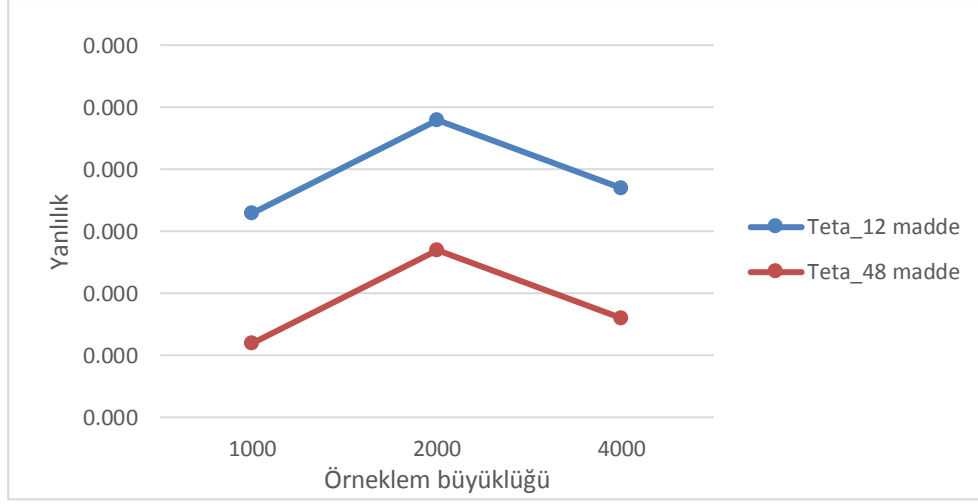
Madde Sayısı	Örneklem	θ		
		RMSE	YANLILIK	KORELAS.
12 madde	1000	0,649	0,033	0,764
	2000	0,638	0,048	0,765
	4000	0,632	0,037	0,767
48 madde	1000	0,387	0,012	0,923
	2000	0,383	0,027	0,922
	4000	0,381	0,016	0,922

Tabloda verilen değerlere ait grafikler şekil 4.4-6'da yer almaktadır.



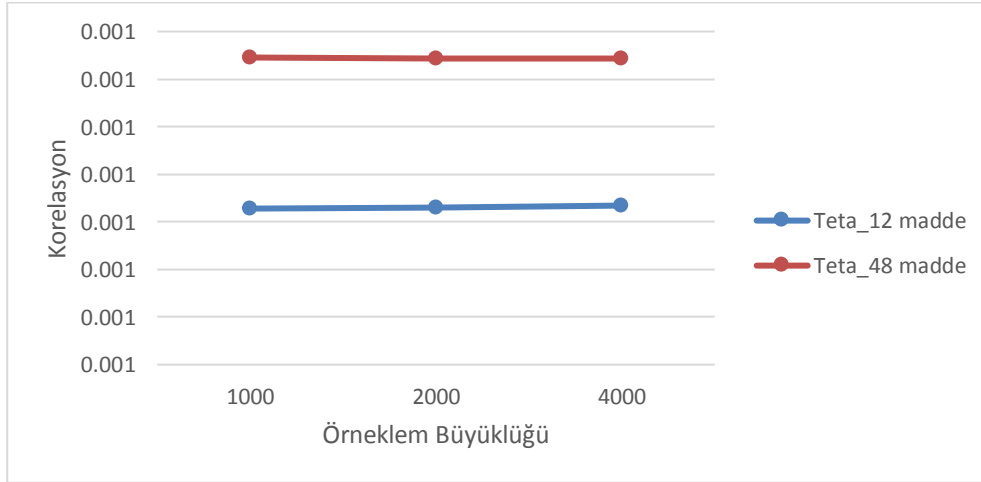
Şekil 4.4. Tek Boyutlu Testler İçin Birey Parametrelerine Ait RMSE Değerleri

Tek boyutlu testler için elde edilen birey parametrelerinin RMSE değerleri 12 maddelik test koşulunda 0,64 değeri civarındadır. Madde sayısı 12'den 48'e çıktığında bu değer 0,38'e düşmüştür. Buna göre test uzunluğunun artması ile birey parametrelerinin kestirimleri daha az hatalı hale gelmiştir. Ancak madde sayısı sabit tutularak örneklem büyüklüğü arttığında birey parametresi kestirimlerinin RMSE değeri önemli bir değişiklik göstermemiştir. Buna göre örneklem büyüklüğünün 1000'den 2000'e çıkarılması durumunda meydana gelen değişim, örneklem büyüklüğünün 2000'den 4000'e çıkarılması durumundaki değişimden daha büyüktür. Sonuç olarak tek boyutlu testlerde yetenek parametrelerinin örneklem büyüklüğünden çok fazla etkilenmediği, ancak test uzunluğundan etkilendiği söylenebilir.



Şekil 4.5. Tek Boyutlu Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri

Şekil 4.5'te de görüldüğü gibi birey parametresi kestirimlerine ilişkin olarak elde edilen yanlılık değerleri tüm koşullarda pozitif işaretlidir. Yanlılık, RMSE değerlerine paralel olarak kısa testten uzun teste geçildiğinde düşmüştür. Her iki test uzunluğu koşulunda da en küçük yanlılık değeri 1000 kişilik örneklem büyüklüğünde elde edilmiş, örneklem büyüklüğü arttıkça yanlılık önce artmış, daha sonra azalmıştır.



Şekil 4.6. Tek Boyutlu Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri

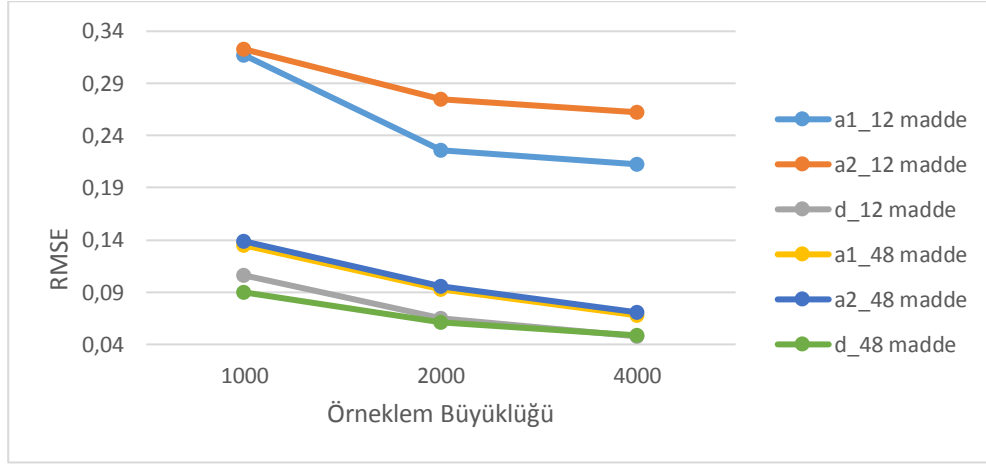
Şekil 4.6'ya göre tek boyutlu testler için kestirilen birey parametreleri ile gerçek birey parametreleri arasındaki korelasyonların ortalaması 0,764 ile 0,923 arasında değişmektedir. Korelasyonlar, örneklem büyüklüğünün artışından etkilenmeyip, testteki madde sayısının artması ile artmıştır. Bu bulgu, yetenek parametrelerinin RMSE'sinin örneklem büyüklüğünden etkilenmemesi bulgusu ile paralellik göstermektedir.

4.2. İkinci alt problem için bulgular

4.2.1. İki boyutlu karmaşık yapıli testlere ilişkin bulgular

İki boyutlu karmaşık yapıli testlerden elde edilen madde parametrelerine ait bulgular tablo 4.3'teki gibidir:

İki boyutlu karmaşık yapıli testler için elde edilen bulgulara ait grafikler şekil 4.7-9'da verilmiştir.



Şekil 4.7. İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait RMSE Değerleri

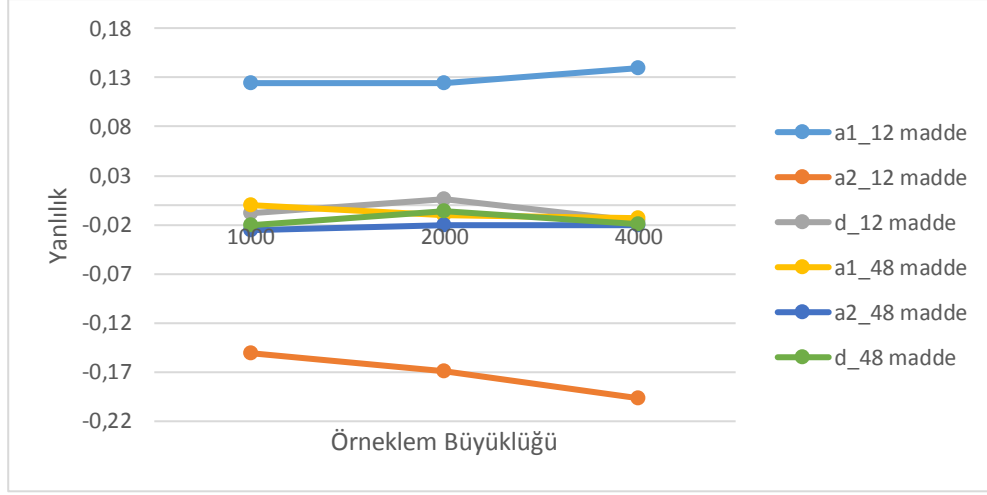
İki boyutlu karmaşık yapıli testlerde madde parametresi kestirim sonuçlarına göre d parametresi her iki test koşulunda da a parametrelerinden daha düşük hata ile kestirilmiştir. Kısa test koşulunda a₁ ve a₂ parametrelerinden elde edilen RMSE değerleri, uzun test koşulundaki RMSE değerlerinden belirgin biçimde yüksektir. Bu sonuçlara göre tüm parametreler hem test uzunluğunun artması hem de örneklem büyüklüğünün artması ile daha düşük hatalı kestirilmiştir. Ayrıca d parametresi, test uzunluğu ve örneklem büyüklüğüne ilişkin tüm koşullarda a parametrelerinden daha düşük RMSE değerleri ile kestirilmiş; d parametresinin a parametresine göre daha kararlı davrandığı gözlenmiştir.

Tablo 4.3: İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Bulgular

Madde sayısı	Örneklem	a_1			a_2			d		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,317	0,124	0,805	0,323	-0,151	0,905	0,106	-0,008	0,995
	2000	0,226	0,124	0,891	0,275	-0,169	0,945	0,065	0,006	0,998
	4000	0,212	0,140	0,917	0,262	-0,197	0,966	0,048	-0,016	0,999
48 madde	1000	0,135	0,000	0,946	0,139	-0,026	0,946	0,090	-0,020	0,997
	2000	0,093	-0,010	0,972	0,096	-0,020	0,983	0,061	-0,006	0,999
	4000	0,068	-0,013	0,985	0,071	-0,020	0,991	0,049	-0,019	0,999

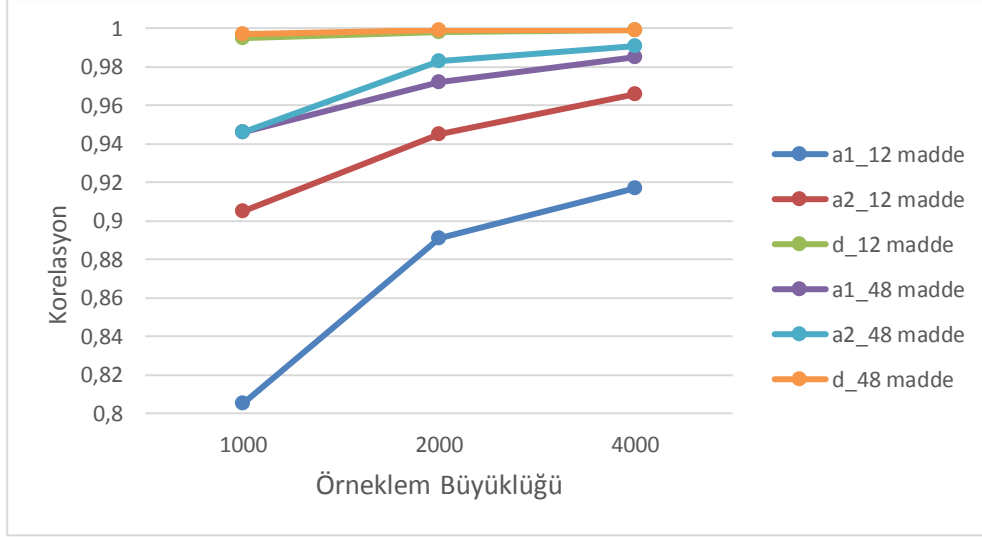
Tablo 4.4: İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Bulgular

Madde Sayısı	Örneklem	θ_1			θ_2		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,715	-0,016	0,686	0,634	0,040	0,774
	2000	0,704	-0,210	0,693	0,621	0,022	0,786
	4000	0,708	-0,008	0,696	0,623	0,030	0,788
48 madde	1000	0,540	-0,005	0,835	0,496	0,039	0,876
	2000	0,531	-0,012	0,838	0,489	0,022	0,880
	4000	0,533	0,005	0,840	0,491	0,030	0,880



Şekil 4.8. İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri

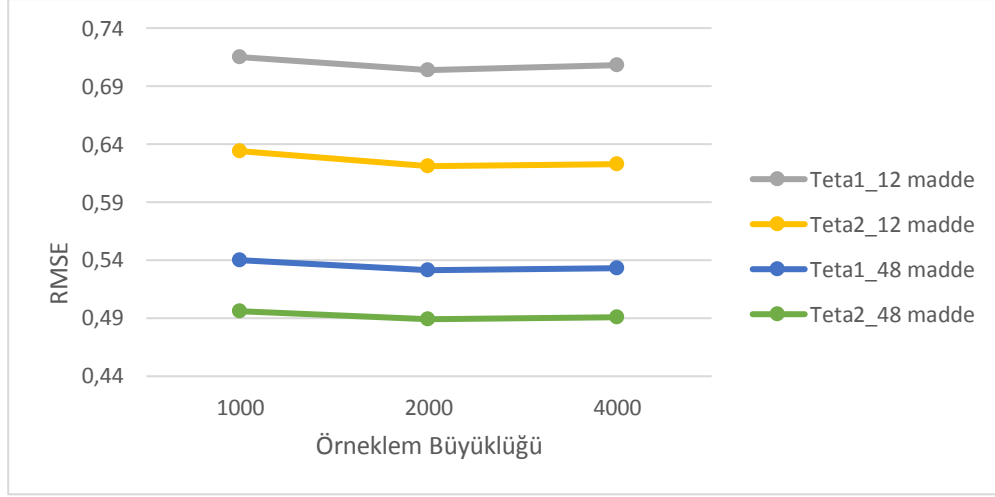
Tablo 4.3 ve Şekil 4.8'e göre; madde parametrelerine ait yanlılık değerlerinin -0,197 ile 0,140 arasında değiştiği görülmektedir. Mutlak değerce en yanlı parametreler 12 maddeli testten kestirilen a_1 ve a_2 parametreleridir. Kestirilen diğer parametrelere ait yanlılıklar birbirine yakın ve daha düşük değerdedir. Kısa test koşulunda d parametresi, hem a_1 hem a_2 parametresinden daha yansız kestirilmiştir. Uzun test koşulunda a_1 parametresi sifıra yakın bir yanlılıkla kestirilmiştir. Grafik incelendiğinde örneklem büyüklüğünün artması ile ne a parametrelerinin ne de d parametresinin yanlılığında düzenli bir azalma ya da artmanın olmadığı söylenebilir. Şekil 4.8'in geneli değerlendirilecek olursa uzun test koşulunda daha yansız kestirimlerin yapıldığı söylenebilir.



Şekil 4.9. İki Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri

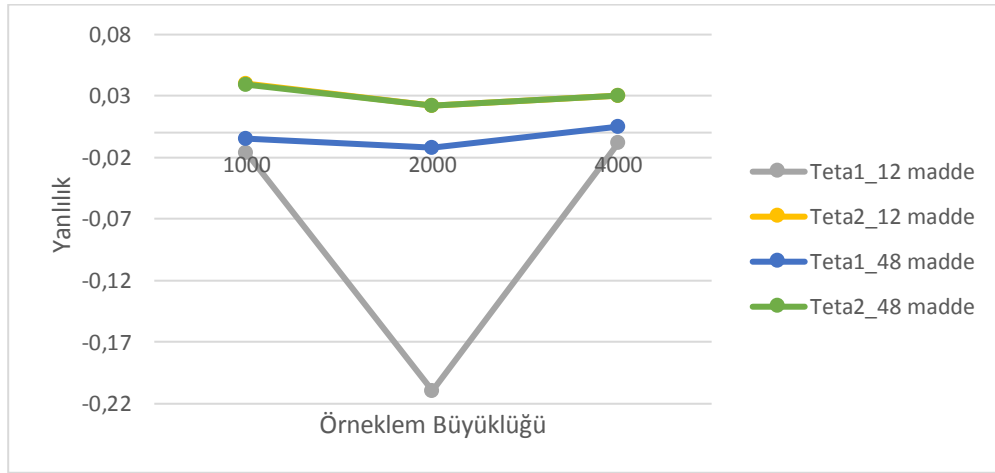
Simülasyon ile elde edilen veri setlerinden kestirilen parametreler ile gerçek parametrelerin korelasyonlarının ortalamasını gösteren şekil 4.9'a göre, tüm parametreler tüm koşullarda 0,805 ve üzerinde korelasyon değerlerine sahiptir. RMSE ve yanlılık bakımından en düşük değerlere sahip olan d parametresi, her bir koşulda en yüksek korelasyon değerlerine sahip parametre olmuştur. d parametresi tüm koşullarda gerçek parametrelerle 0,995 ve üzerinde pozitif yönde, mükemmel düzeyde ilişkili olup, örneklemin artmasından önemli biçimde etkilenmemiştir. a parametresinin korelasyon değerleri tablo 4.3'te de görüldüğü gibi 0,805 ile 0,991 arasında değişmektedir. a parametrelerine ilişkin korelasyonlar örneklem büyüklüğünün artması ile daha yüksek değere ulaşmıştır. RMSE ve yanlılık sonuçları ile çelişir biçimde, iki boyutlu karmaşık yapıdaki testlerin tüm koşullarında daha yüksek hata ve yanlılıkla kestirilen a₂ parametresi küçük bir farkla da olsa tüm koşullarda a₁ parametresinden daha yüksek korelasyonlar vermektedir.

İki boyutlu karmaşık yapıları testlerden kestirilen birey parametrelerine ilişkin RMSE, yanlılık ve korelasyon değerleri tablo 4.4'te verilmiştir. Tablo 4.4'te yer alan değerler, şekil 4.10-12'de grafiklerle verilmiştir.



Şekil 4.10. İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri

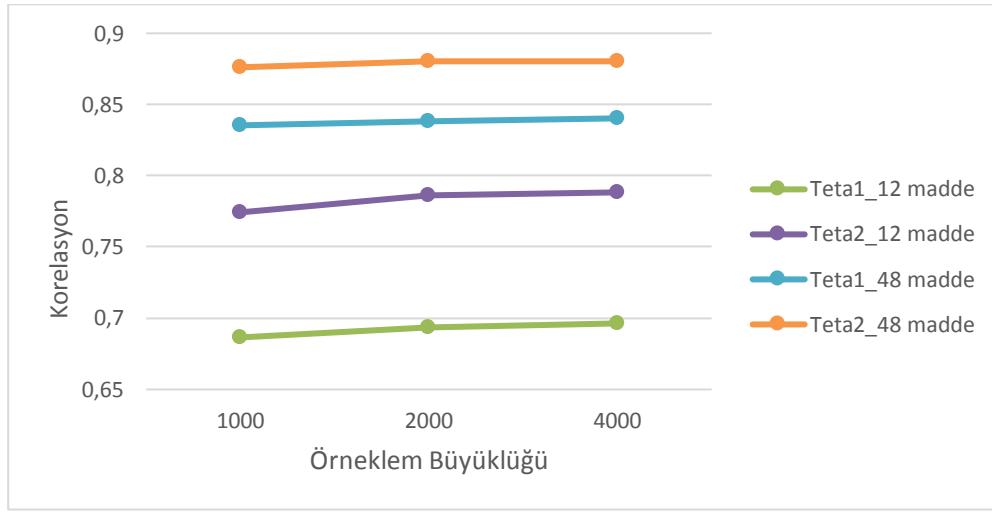
İki boyutlu karmaşık yapıları testlerden kestirilen yetenek parametrelerine ilişkin RMSE değerleri incelendiğinde, değerlerin örneklem büyüklüğünün artması ile ciddi bir değişime uğramadığı görülmektedir. Ancak kısa testten uzun teste geçilmesi ile her iki yetenek parametresi için RMSE değerlerinde azalış gözlenmiştir. Bu durum, birey parametrelerinin test uzunluğundan etkilendiği; örneklem büyüklüğünden etkilenmediği şeklinde yorumlanabilir. Ayrıca her iki test uzunluğu koşullarında da teta2 parametresi, teta1 parametresinden daha kararlı kestirilmiş, RMSE değerleri daha düşük çıkmıştır.



Şekil 4.11. İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri

Şekil 4.11'deki yanlılık sonuçlarına göre parametreler, -0,210 ile 0,040 değerleri arasında değişen bir yanlılıkla kestirilmiştir. Değerler tek boyutlu veri setlerindeki

gibi örneklem büyüklüğü ve test uzunluğundaki artış ile orantılı olmayan bir örüntüye sahiptir. Teta1 için hem uzun test hem kısa testte hesaplanan yanlışlık değerleri örneklem büyüklüğünün 1000'den 2000'e çıkması ile önce artmış; 4000 örnekleme gelindiğinde tekrar azalmıştır. Teta2 için ise tam tersi durum gerçekleşmiştir. Mutlak değerce en büyük yanlışlık, teta1 parametresinin 12 madde ve 2000 örneklem koşulunda ortaya çıkmıştır. Her iki test uzunluğunda da teta2 parametreleri, teta1 parametrelerinden daha düşük ve her iki test uzunluğunda da birbirine çok yakın yanlışlıklarla kestirilmiştir.



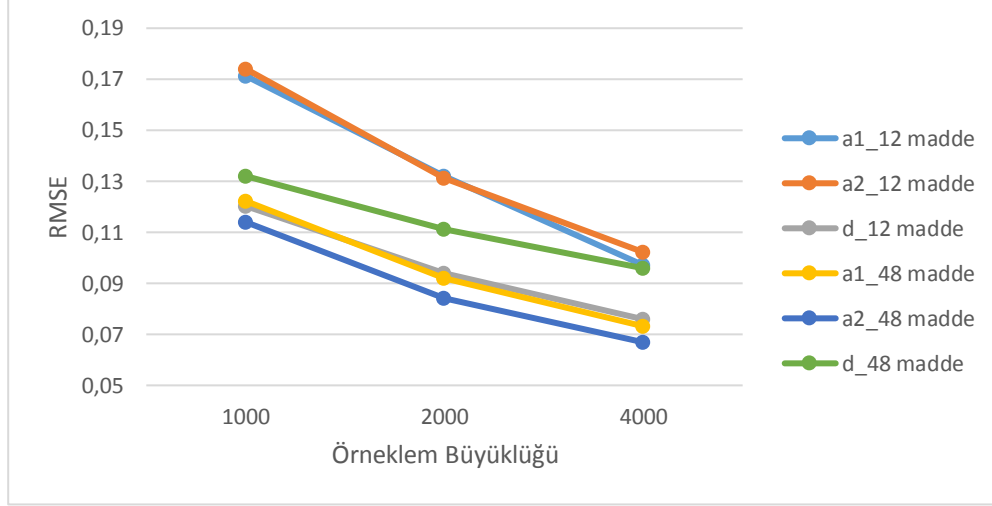
Şekil 4.12. İki Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri

Her bir tekrarda kestirilen birey parametreleri ile gerçek birey parametrelerinin korelasyonlarının ortalamalarını veren şekil 4.12'de görüldüğü gibi yetenek parametrelerine ait korelasyon değerleri 0,686 ile 0,880 arasında değişmektedir. Genel olarak testte yer alan madde sayısı arttıkça teta parametrelerine ait korelasyon değeri artmıştır. RMSE'si daha düşük değer alan teta2 parametresi her iki test uzunluğu koşulunda da gerçek parametrelerle daha yüksek ilişki göstermektedir. Madde sayısı sabit kalmak üzere örneklem büyüklüğü arttırıldığında korelasyon değerinde ciddi bir değişiklik meydana gelmemiştir. Bu bulgu RMSE değerleri ile birlikte ele alındığında birey parametrelerinin örneklem büyüklüğünden etkilenmediği bulgusunu destekler niteliktedir.

4.2.2. İki boyutlu basit yapıli testlere ilişkin bulgular

İki boyutlu basit yapıli testlerden kestirilen madde parametrelerine ait RMSE, yanlılık ve korelasyon deęerleri tablo 4.5'teki gibidir.

Tabloda yer alan istatistiklere ait grafikler Őekil 4.13-15'te yer almaktadır.



Őekil 4.13. İki Boyutlu Basit Yapılı Testler İin Madde Parametrelerine Ait RMSE Deęerleri

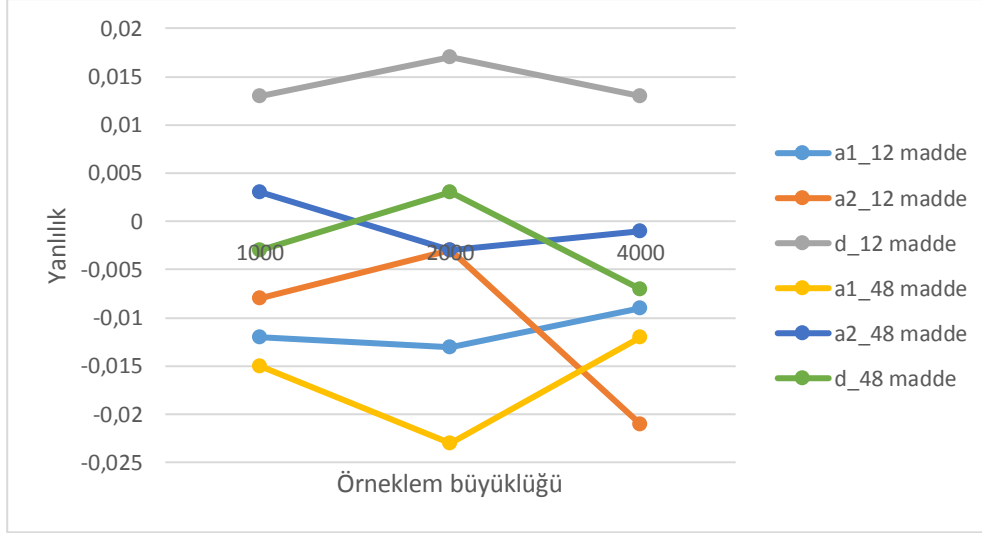
Őekil 4.13 incelenirse, karmaŐık yapıli testlerde olduęu gibi basit yapıli testlerde de hem örneklemin büyümesi hem de madde sayısının artması RMSE deęerlerinin düşmesine sebep olmuŐtur. Tüm parametreler için RMSE deęerleri 0,067 ile 0,174 arasında deęişen deęerler almaktadır. Kısa test koşulunda d parametresine ait RMSE'ler aynı koşulda kestirilen a parametrelerinden daha düşük deęerdedir. Bu sebeple bu yapıda da en kararlı kestirilen parametrenin d parametresi olduęu söylenebilir. Tüm parametreler için en düşük RMSE'ler madde ve birey sayısının en yüksek olduęu koşul olan 48 madde ve 4000 örnekleme koşulunda ortaya çıkmıŐtır.

Tablo 4.5: İki Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Bulgular

Madde Sayısı	Örneklem	a_1			a_2			d		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,171	-0,012	0,959	0,174	-0,008	0,967	0,120	0,013	0,983
	2000	0,132	-0,013	0,969	0,131	-0,003	0,976	0,094	0,017	0,985
	4000	0,097	-0,009	0,978	0,102	-0,021	0,981	0,076	0,013	0,987
48 madde	1000	0,122	-0,015	0,974	0,114	0,003	0,979	0,132	-0,003	0,956
	2000	0,092	-0,023	0,981	0,084	-0,003	0,985	0,111	0,003	0,957
	4000	0,073	-0,012	0,982	0,067	-0,001	0,988	0,096	-0,007	0,958

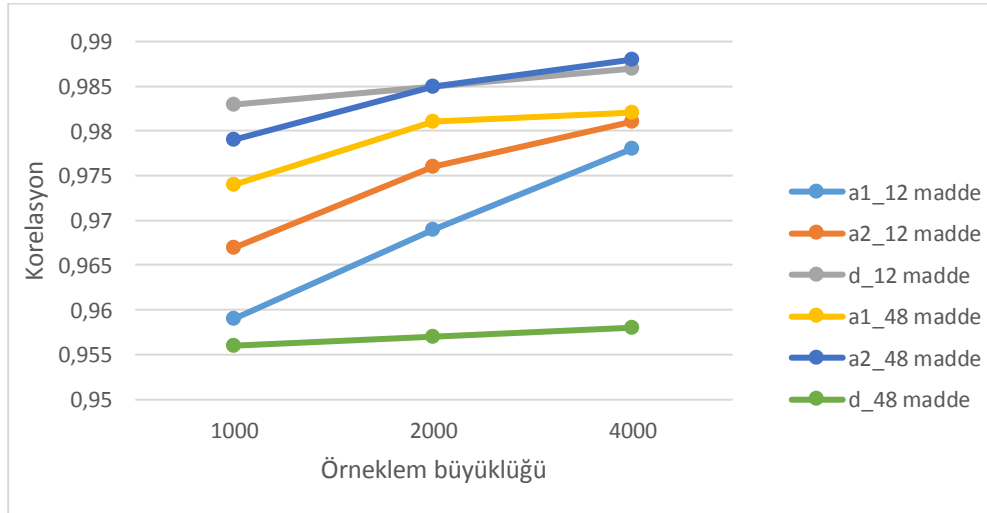
Tablo 4.6: İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Bulgular

Madde Sayısı	Örneklem	θ_1			θ_2		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,761	0,027	0,631	0,749	0,072	0,668
	2000	0,756	0,022	0,630	0,741	0,054	0,676
	4000	0,761	0,040	0,635	0,744	0,061	0,677
48 madde	1000	0,499	0,022	0,865	0,526	0,043	0,865
	2000	0,492	0,015	0,863	0,523	0,025	0,862
	4000	0,495	0,033	0,865	0,524	0,033	0,864



Şekil 4.14. İki Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri

Şekil 4.14'te yer alan yanlılık değerleri incelendiğinde, yanlılığın daha önce ele alınan yanlılık grafiklerindeki gibi düzensiz dağılım sergilediği görülmektedir. Elde edilen sonuçlara göre yanlılık -0,023 ile 0,017 arasında değişen değerler almıştır. Yanlılığın ranjının ve aldığı değerlerin oldukça düşük olması, parametrelerin genel olarak benzer biçimde, sifıra yakın bir yanlılık ile kestirildiği anlamına gelmektedir.

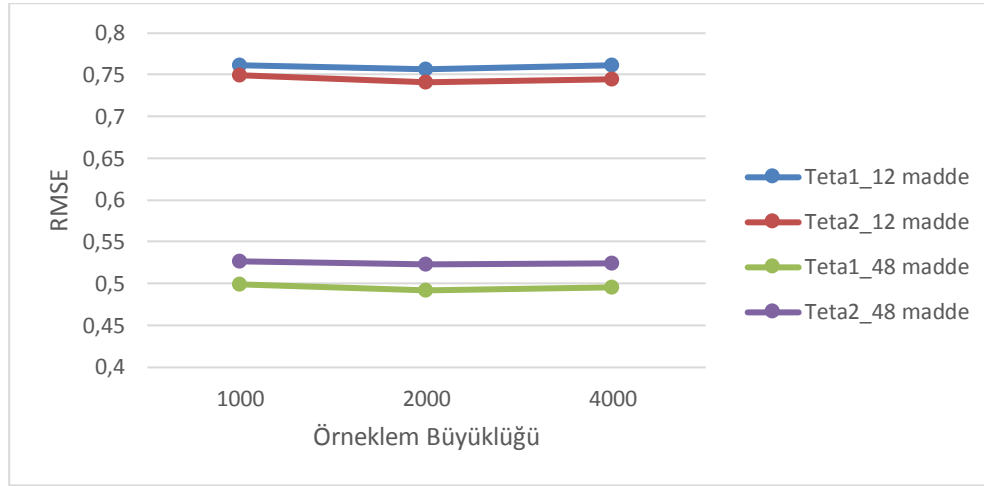


Şekil 4.15. İki Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri

İki boyutlu basit yapıları testlerden kestirilen madde parametrelerinin gerçek parametrelerle korelasyonlarının ortalamasını veren şekil 4.15 incelenirse, uzun testlerde kestirilen d parametresinin en düşük korelasyon değerine sahip olduğu

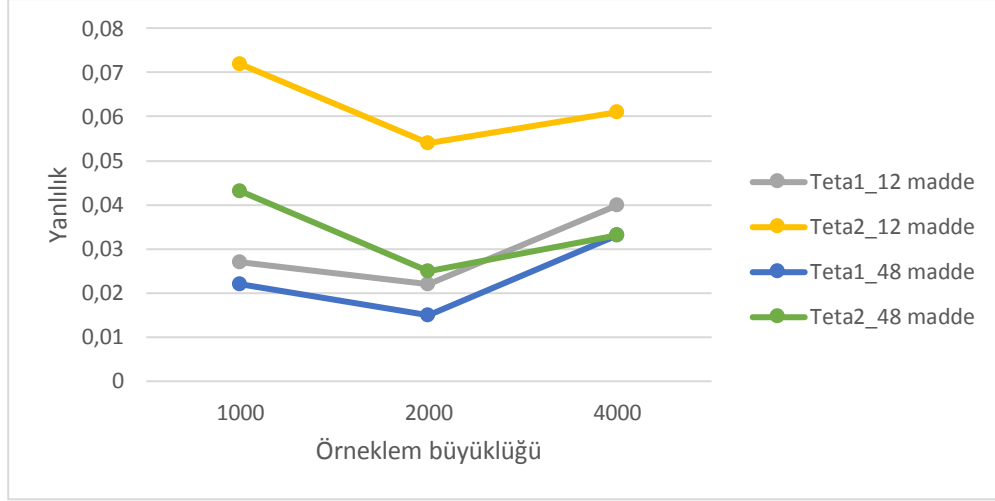
görülmektedir. Ancak değerlerin 0,956 ile 0,988 arasında değişmesi sebebiyle ranjin çok düşük, parametrelerin hepsinin mükemmel düzeyde korelasyon değerlerine sahip olduğu söylenebilir.

İki boyutlu basit yapıli testlerden kestirilen birey parametrelerinin RMSE, yanlılık ve korelasyon değerleri tablo 4.6'da verilmiştir. Tablo 4.6'ya ilişkin grafikler şekil 4.16-18'de yer almaktadır.



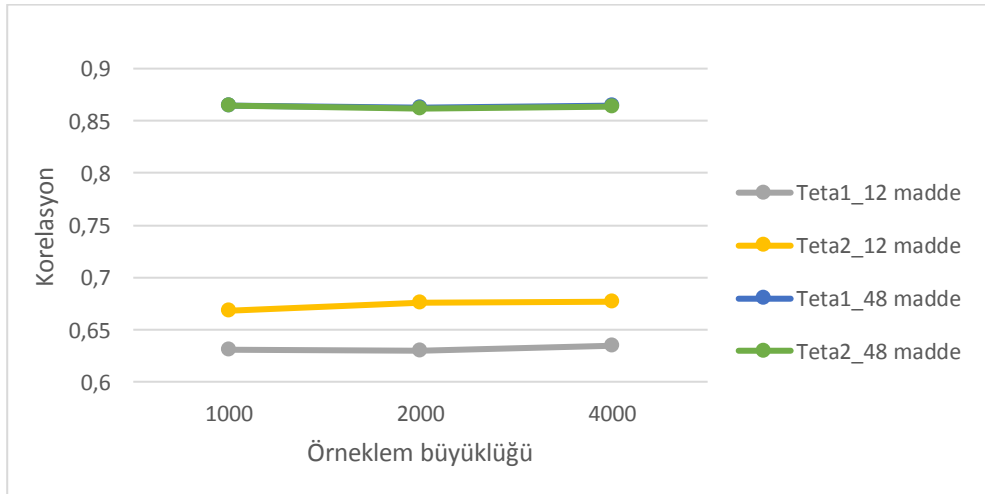
Şekil 4.16. İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri

Şekil 4.16'da da görüldüğü gibi 2 boyutlu basit yapıli testlerin birey parametrelerine ilişkin RMSE değerleri madde sayısının artması ile hem teta1 hem teta2 parametresi için 0,7 civarından 0,5 civarına gerilemiştir. Kısa test koşulunda teta2'ye ilişkin RMSE değerleri teta1'den daha düşükken, uzun testte teta1'e ilişkin RMSE değeri teta2'den daha düşüktür. Madde sayısı sabit olmak koşulu ile birey sayısı arttırıldığında birey parametrelerinin RMSE değerlerinde önemli bir değişiklik görülmemiştir. Bu durum daha önceki yapılarda da elde edilen birey parametrelerinin örneklem büyüklüğünden etkilenmediği bulgusuna paralellik göstermektedir.



Şekil 4.17. İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri

Şekil 4.17'deki yanlılık değerleri grafiğine göre ele alınan tüm koşullarda parametrelerin yanlılığı pozitifdir ve 0,015 ile 0,072 arasında değişen değerler almaktadır. Ayrıca tüm koşullarda birey sayısı arttıkça yanlılık değerleri önce azalmış sonra artmıştır. Buna göre hem uzun test hem kısa testte teta1 parametreleri için en yanlı kestirim 4000 kişilik örnekleme, en yansız kestirim 2000 kişilik örnekleme yapılmıştır. Her iki test uzunluğu koşulunda da teta2 parametreleri için yapılan en yanlı kestirim 1000 örnekleme yapılırken, en yansız kestirim 2000 kişilik örnekleme yapılmıştır. Grafiğin bütünü göz önüne alınırsa teta1 parametrelerinin teta2 parametrelerinden daha yansız kestirildiği söylenebilir.



Şekil 4.18. İki Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri

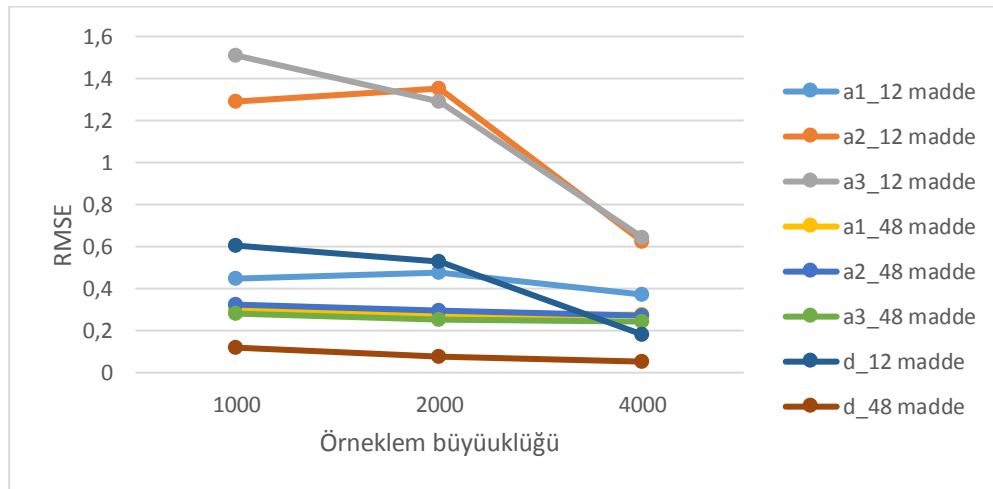
Şekil 4.18’de iki boyutlu basit yapıli testlerden kestirilen birey parametreleri ile gerçek birey parametrelerinin korelasyonlarının ortalaması görölmektedir. Buna göre kısa test koşulu her iki teta parametresi için de en düşük korelasyonların elde edildiği koşuldur. Bu koşulda teta1 0,63 civarında korelasyon değerleri ile teta2’den daha düşük değerlere sahiptir. Uzun test koşulunda ise hem teta1 hem teta2 0,86 etrafında ve birbirine çok yakın korelasyon değerlerine sahiptir. Bu sebeple grafikte 48 maddelik testten elde edilen teta1 ve teta2 parametrelerinin korelasyon değerlerine ilişkin çizilen çizgiler çakışık çıkmıştır. Daha önceki bulgulara benzer şekilde burada da birey parametrelerine ait korelasyonlar örneklem büyüklüğünden etkilenmemiştir.

4.3. Üçüncü alt problem için bulgular

4.3.1. Üç boyutlu karmaşık yapıli testlere ilişkin bulgular

Üç boyutlu karmaşık yapıli testlerden elde edilen madde parametrelerine ait RMSE, yanlılık ve korelasyon değerleri tablo 4.7’deki gibidir.

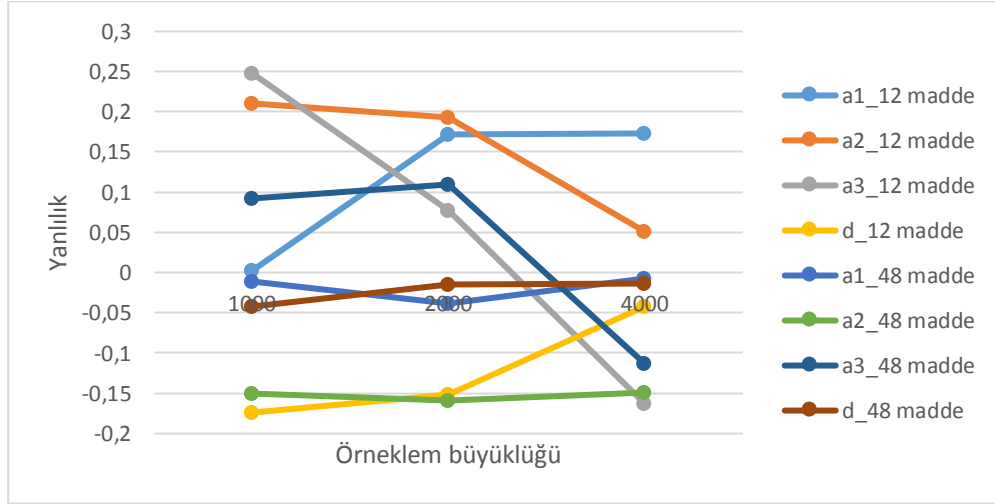
Tabloda yer alan istatistiklere ilişkin grafikler şekil 4.19-21’de verilmiştir.



Şekil 4.19. Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait RMSE Değerleri

Üç boyutlu karmaşık yapıli testlerden kestirilen madde parametrelerine ilişkin RMSE değerleri, tek ve iki boyutlu desenlerde elde edilen RMSE değerlerinden farklılık göstermiştir. Bu desende hesaplanan RMSE değerleri 0,054 ile 1,508 arasında değişen geniş bir aralığa sahiptir. Üç boyutlu karmaşık yapılarda en yüksek hata ile kestirilen parametre 12 maddeli test ve 1000 örneklem büyüklüğü koşulunda

kestirilen a_3 parametresi olmuştur. Örneklem büyüklüğünün artması ile a_3 parametresinin RMSE değerinde ciddi bir düşme gözlenmiştir. Kısa test koşulunda a_3 parametresinden sonra en yüksek RMSE değerleri a_2 parametresine aittir. Bu parametre en yüksek RMSE değerini 2000 örneklem büyüklüğü koşulunda almıştır. Diğer desenlerin aksine bu desende kısa testin 1000 ve 2000 örneklem büyüklüğü koşulunda a_1 parametresinin RMSE değeri d parametresinin RMSE değerinden düşük değer almıştır. Bu yapıda da kısa testten uzun teste geçildiğinde madde parametrelerine ait RMSE değerlerinde önemli bir düşme gözlenmiştir. Uzun test koşulunda a_1 , a_2 ve a_3 parametrelerinin RMSE değerleri birbirine çok yakındır. Uzun test koşulunda en düşük RMSE değerlerine sahip parametre d parametresi olmuştur. Bu bulgu önceki desenlerde elde edilen bulgulara paralellik göstermektedir.



Şekil 4.20. Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri

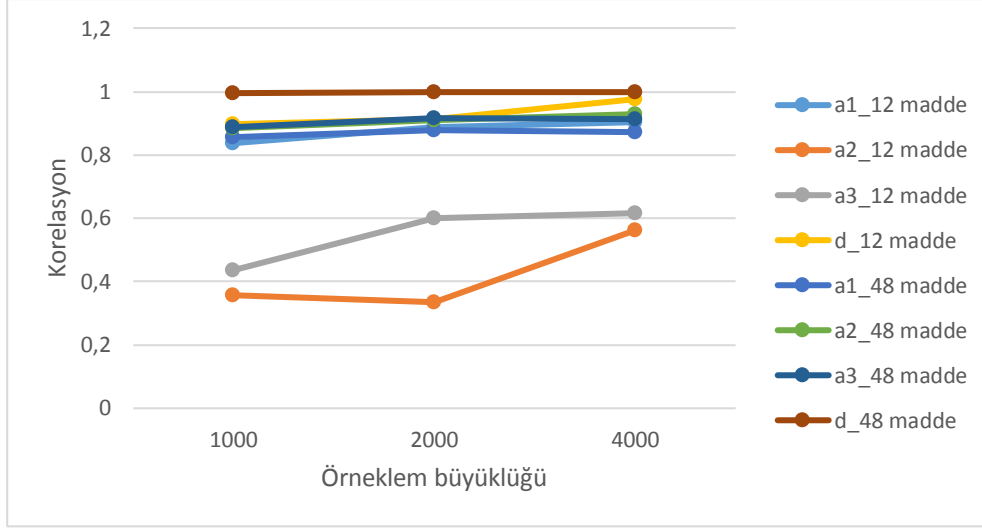
Şekil 4.20'de yer alan grafiğe göre madde ve örneklem büyüklüğündeki artış ile yanlılık değerlerinde düzenli bir örüntü elde edilememiştir. En yanlı kestirim testin kısa ve örneklem büyüklüğünün 1000 olduğu koşulda a_3 parametresine (0,247); en yansız kestirim ise testin kısa ve örneklem büyüklüğünün 1000 olduğu koşulda a_1 parametresine (0,002) aittir.

Tablo 4.7: Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Bulgular

Madde sayısı	Örnekleme	a1			a2			a3			d		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,446	0,002	0,836	1,289	0,210	0,358	1,508	0,247	0,435	0,602	-0,174	0,899
	2000	0,474	0,172	0,888	1,350	0,193	0,336	1,290	0,077	0,601	0,530	-0,152	0,912
	4000	0,371	0,173	0,903	0,624	0,051	0,562	0,640	-0,163	0,615	0,181	-0,042	0,977
48 madde	1000	0,298	-0,011	0,857	0,325	-0,150	0,884	0,279	0,092	0,889	0,118	-0,042	0,996
	2000	0,273	-0,039	0,880	0,294	-0,159	0,911	0,254	0,110	0,915	0,078	-0,015	0,998
	4000	0,274	-0,008	0,872	0,273	-0,149	0,928	0,245	-0,113	0,913	0,054	-0,014	0,999

Tablo 4.8. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Bulgular

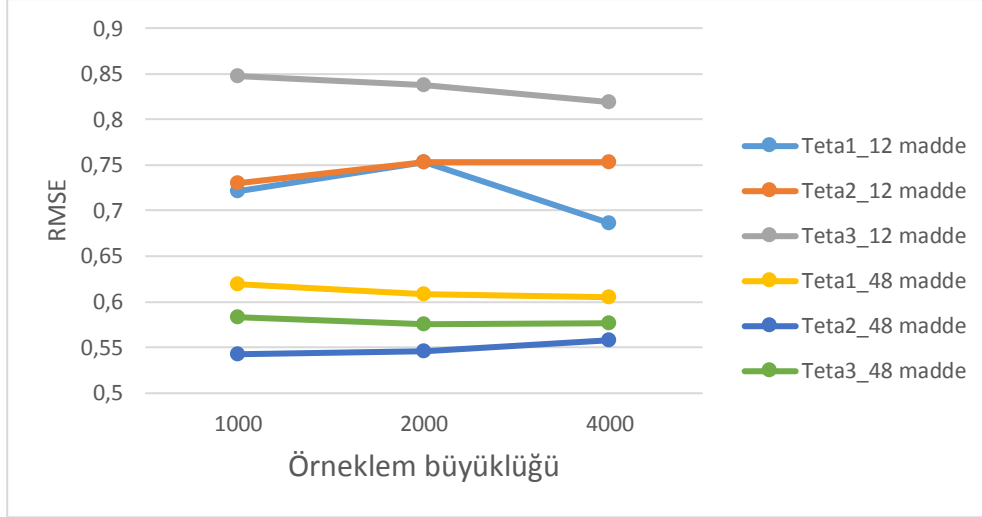
Madde sayısı	Örnekleme	θ1			θ2			θ3		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,721	0,043	0,695	0,730	0,038	0,626	0,848	0,067	0,526
	2000	0,753	0,007	0,718	0,753	0,034	0,603	0,838	0,070	0,531
	4000	0,686	0,010	0,731	0,753	0,036	0,649	0,819	0,046	0,556
48 madde	1000	0,619	0,009	0,785	0,542	0,006	0,821	0,583	0,046	0,823
	2000	0,608	-0,029	0,793	0,546	-0,001	0,825	0,575	0,044	0,830
	4000	0,605	-0,021	0,797	0,558	0,003	0,833	0,576	0,029	0,823



Şekil 4.21. Üç Boyutlu Karmaşık Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri

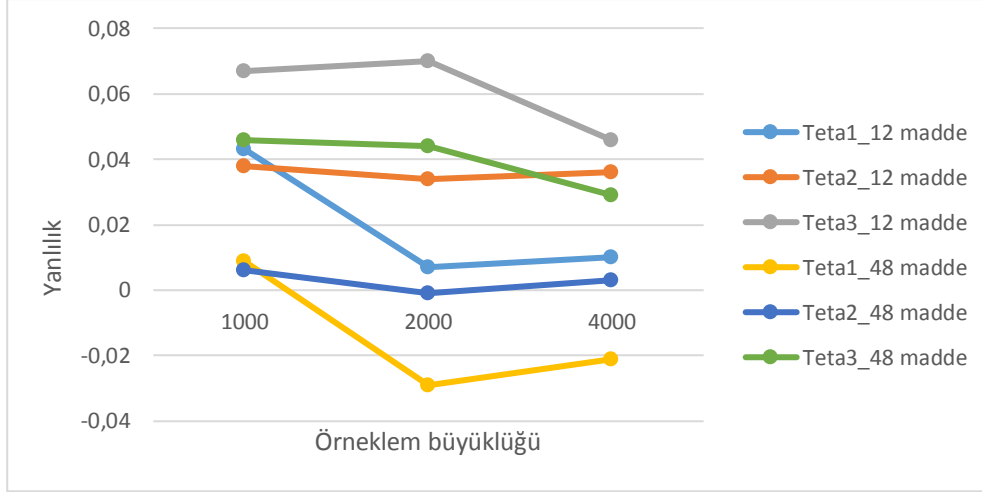
Şekil 4.21’de kestirilen madde parametreleri ile gerçek madde parametreleri arasındaki korelasyonların ortalaması görülmektedir. Korelasyon ortalamaları 0,336-0,999 aralığındadır. Tüm madde parametreleri için ilgili desenlerdeki korelasyon değerleri, örneklem büyüklüğü ve test uzunluğu arttıkça artmıştır. En düşük korelasyonlar, kısa test koşulunda kestirilen a_2 ve a_3 parametrelerine aittir. Bu iki parametre 0,336 ile 0,615 arasında değişen düşük ve orta düzeyde korelasyon değerlerine sahiptir. Uzun test koşulunda tüm a parametrelerinin korelasyon değerlerinin 0,857 ve üzerinde çıktığı görülmektedir. d parametresi hem 12 hem 48 maddeli testte tüm örneklem büyüklüklerinde 0,899 ve üzerindeki değerlerle gerçek parametrelerle mükemmel düzeyde ilişkiye sahiptir.

Üç boyutlu karmaşık yapıli testlerden elde edilen birey parametrelerine ait RMSE, yanlılık ve korelasyon değerleri tablo 4.8’deki gibidir. Tablo 4.8’de yer alan bulgulara ilişkin grafikler 4.22-24’te verilmiştir.



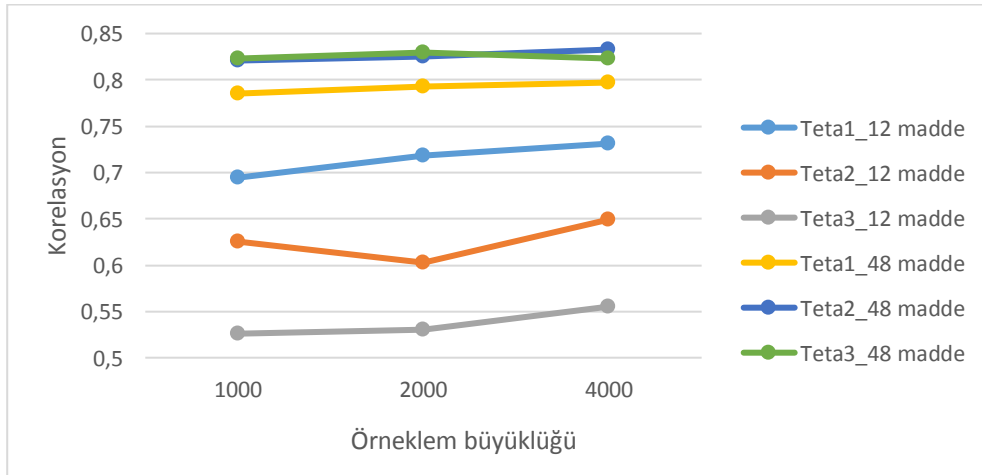
Şekil 4.22. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri

Üç boyutlu karmaşık yapılı testlerden kestirilen birey parametrelerine ilişkin RMSE değerleri 0,542 ile 0,848 arasında değişmektedir. En yüksek hata ile kestirilen parametre, kısa testlerde kestirilen teta3 parametresi olmuştur. Örneklem büyüklüğünün artması parametrelerin RMSE değerlerinde tek yönlü bir değişime sebep olmamış, bazı yetenek parametrelerinde örneklem büyüklüğünün artmasıyla RMSE azalmış, bazı parametrelerde ise artmıştır. Madde sayısının 12'den 48'e çıkmasıyla teta1 ve teta3'ün RMSE değerlerinde düşüş gerçekleşmiştir. 48 maddeli testte ise RMSE'nin örneklem büyüklüğüne bağlı değişimi üç yetenek parametresinde de küçük miktarlarda gerçekleşmiştir. Uzun test koşulunda tüm örneklem büyüklüklerinde elde edilen en düşük RMSE değeri teta2 parametresine aittir.



Şekil 4.23. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri

Şekil 4.23'e göre yetenek parametrelerine ait yanlılık -0.029 ile 0.070 arasında değişen değerler almaktadır. Üç boyutlu karmaşık yapıli testlerde birey parametrelerine ait yanlılık, örneklem büyüklüğü ve madde sayısından ciddi biçimde etkilenmemiştir. Kestirilen yanlılık değerlerinin oldukça küçük değerlere sahip olması sebebiyle yanlılığın genel olarak birbirine yakın olduğu söylenebilir. Buna göre yetenek parametreleri tüm koşullarda benzer yanlılıkla kestirilmiştir.



Şekil 4.24. Üç Boyutlu Karmaşık Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Değerleri

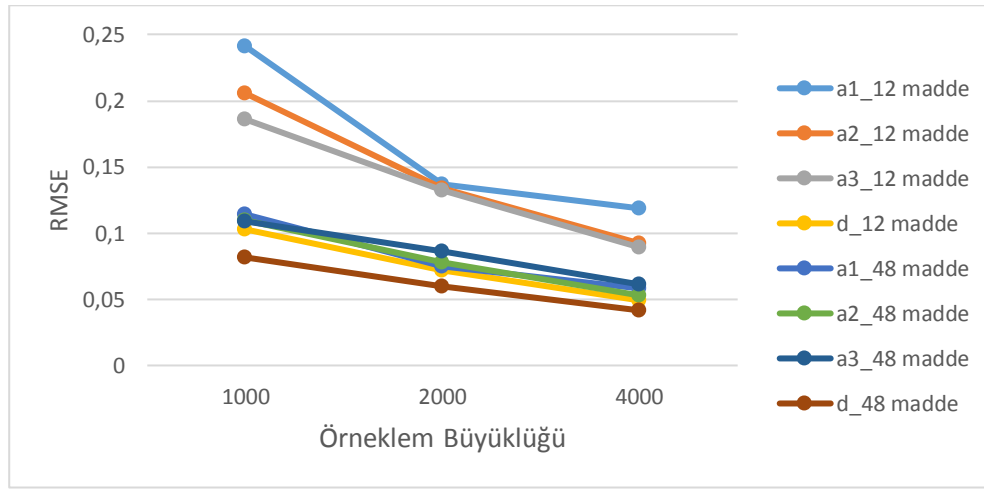
Şekil 4.24'e göre kısa testten elde edilen korelasyon ortalamaları 0,526 ile 0,731; uzun testten elde edilen korelasyon ortalamaları 0,785 ile 0,833 arasında değişmektedir. Testte yer alan madde sayısının artması korelasyon değerlerinin artmasını sağlamıştır. 12 maddeli testte örneklem büyüklüğünün artışı, 48 maddeli testteki örneklem büyüklüğü artışından daha etkili olmuştur. Korelasyonlar en

yüksek değerini 48 maddeli test ve 4000 kişilik örneklem büyüklüğü koşulunda teta2 parametresinde, en düşük değerini 12 maddeli test ve 1000 örneklem büyüklüğü koşulunda teta3 parametresinde almıştır.

4.3.2. Üç boyutlu basit yapıli testlere ilişkin bulgular

Üç boyutlu basit yapıli testlerden elde edilen madde parametrelerine ait RMSE, yanlılık ve korelasyon değerleri tablo 4.9'daki gibidir.

Tablo 4.9'da yer alan bulgulara ilişkin grafikler şekil 4.25-27'de yer almaktadır.



Şekil 4.25. Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait RMSE Değerleri

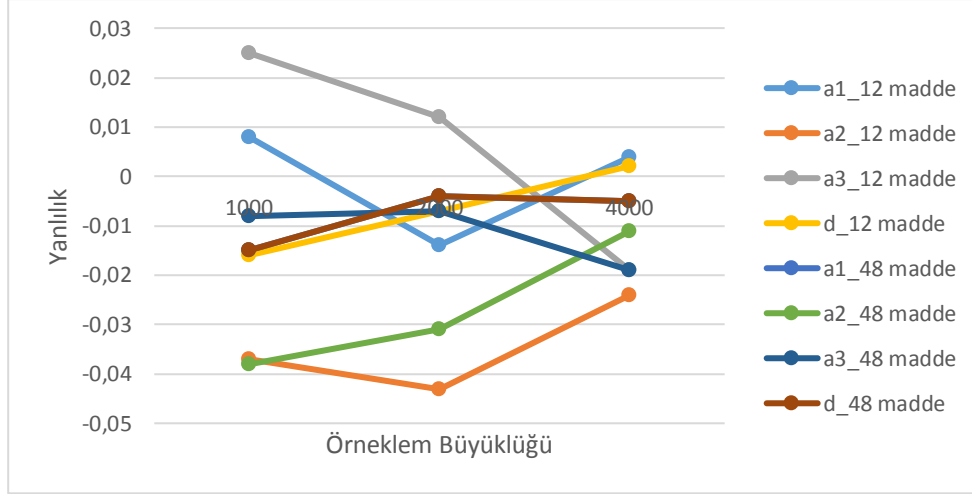
Şekil 4.25'e göre, madde parametrelerinin kestirimine ait RMSE değerleri 0,241 ile 0,042 arasında değerler almaktadır. Bu yapıda da örneklem büyüklüğünün artması ve madde sayısının 12'den 48'e çıkmasıyla RMSE değerleri azalmıştır. Hem uzun hem kısa test koşulunda elde edilen sonuçlara göre en düşük RMSE değerine sahip parametre d parametresi olmuştur. Bu sonuca dayalı olarak bu yapıda da en kararlı kestirilen madde parametresinin d parametresi olduğu yorumuna gidilebilir. Kısa test koşulunda a parametrelerinin RMSE değerleri arasında değişim gözlenirken uzun test koşuluna gelindiğinde a₁, a₂ ve a₃ parametrelerinin RMSE değerleri genel olarak birbirine çok yakın değerler almıştır.

Tablo 4.9: Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Bulgular

Madde sayısı	Örneklem	a_1			a_2			a_3			d		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,241	0,008	0,955	0,206	-0,037	0,942	0,186	0,025	0,980	0,103	-0,016	0,993
	2000	0,137	-0,014	0,984	0,134	-0,043	0,972	0,132	0,012	0,990	0,072	-0,007	0,996
	4000	0,119	0,004	0,987	0,092	-0,024	0,986	0,089	-0,019	0,995	0,049	0,002	0,998
48 madde	1000	0,114	0,004	0,990	0,110	-0,038	0,987	0,109	-0,008	0,987	0,082	-0,015	0,996
	2000	0,075	-0,001	0,996	0,078	-0,031	0,994	0,086	-0,007	0,992	0,060	-0,004	0,998
	4000	0,058	0,005	0,997	0,053	-0,011	0,997	0,061	-0,019	0,996	0,042	-0,005	0,999

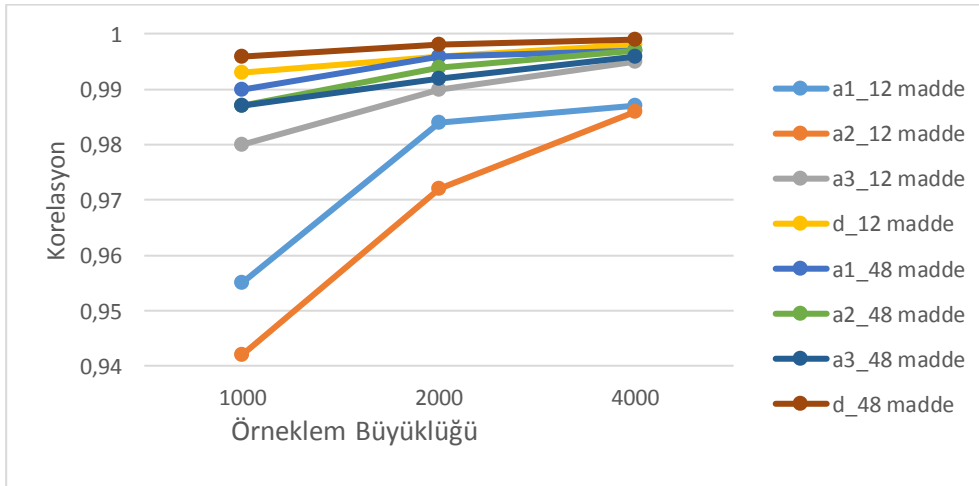
Tablo 4.10: Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Bulgular

Madde sayısı	Örneklem	θ_1			θ_2			θ_3		
		RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.	RMSE	YANLILIK	KORELAS.
12 madde	1000	0,827	0,044	0,542	0,832	0,045	0,480	0,768	0,085	0,650
	2000	0,817	0,015	0,574	0,838	0,037	0,493	0,763	0,084	0,655
	4000	0,819	0,024	0,578	0,819	0,042	0,515	0,752	0,068	0,655
48 madde	1000	0,567	0,054	0,828	0,580	0,018	0,803	0,612	0,064	0,803
	2000	0,553	0,016	0,833	0,587	0,010	0,804	0,611	0,062	0,805
	4000	0,554	0,024	0,835	0,592	0,014	0,815	0,605	0,047	0,802



Şekil 4.26. Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Yanlılık Değerleri

Üç boyutlu basit yapıli veri setlerinden kestirilen madde parametrelerinin yanlılıkları şekil 4.26'da verilmiştir. Madde yanlılıkları daha önce incelenen iki ve üç boyutlu yapılara benzer biçimde madde sayısının ve örneklem büyüklüğünün artması ile düzenli bir artma ya da azalma göstermemiştir. Diğer taraftan madde yanlılıkları -0,043 ile 0,025 arasında değişen, birbirine ve sıfıra yakın değerler almıştır. Tüm koşullar göz önünde bulundurulduğunda en yanlı kestirilen parametre, kısa test ve 2000 örneklem büyüklüğündeki a_2 parametresi iken, en yansız kestirilen parametre kısa test ve 4000 örnekleme de kestirilen d parametresi olmuştur.

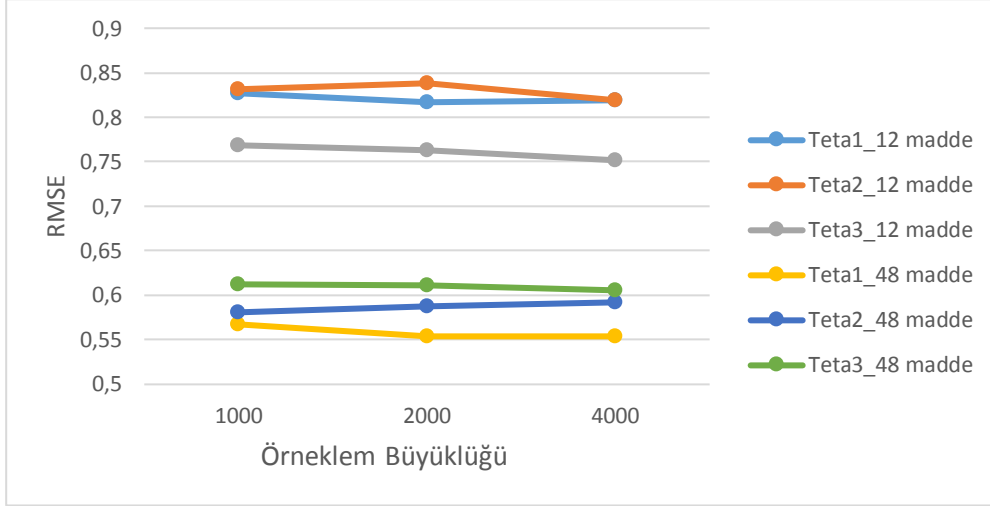


Şekil 4.27. Üç Boyutlu Basit Yapılı Testler İçin Madde Parametrelerine Ait Korelasyon Değerleri

Üç boyutlu basit yapıli testlerden kestirilen madde parametreleri ile gerçek madde parametreleri arasındaki korelasyon değerlerini içeren şekil 4.27'ye göre korelasyon

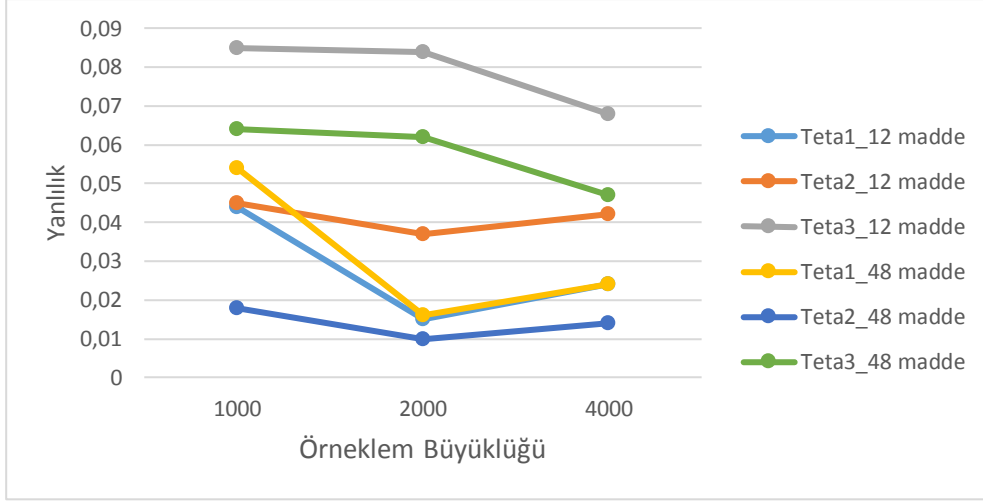
değerleri madde sayısı ve örneklem büyüklüğü arttıkça artmaktadır. Tüm parametreler 0,942 ile 0,999 arasında değişen mükemmel düzeyde korelasyon değerlerine sahiptir.

Üç boyutlu basit yapıli testlerden kestirilen birey parametrelerine ait bulgular tablo 4.10'da verilmiştir. Tablo 4.10'da yer alan değerlere ilişkin grafikler şekil 4.28-30'da yer almaktadır.



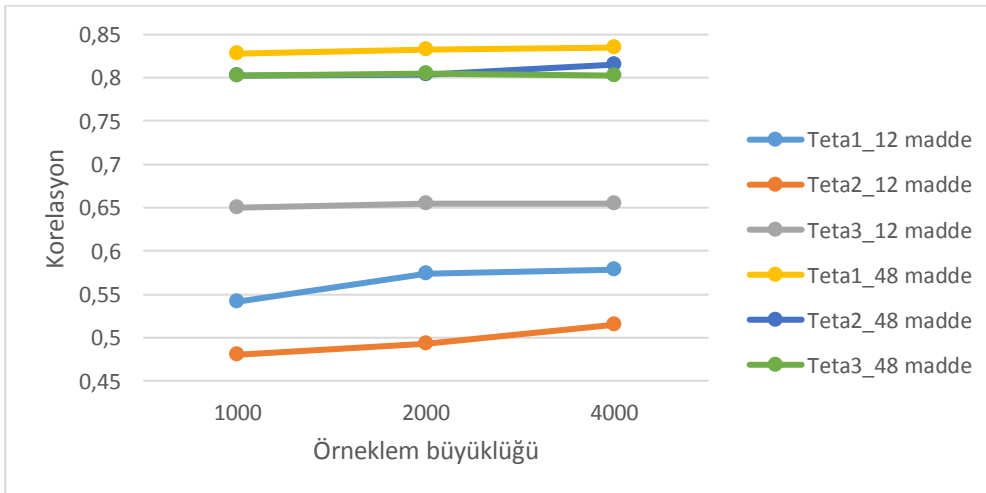
Şekil 4.28. Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait RMSE Değerleri

Şekil 4.28'e göre birey parametrelerinin RMSE değerleri 0,553 ile 0,838 arasında değerler almıştır. Test uzunluğunun artması RMSE'de fark edilir bir düşüğe sebep olmuştur. Ancak birey sayısının artması tetaların RMSE değerlerinde önemli bir etkiye sebep olmamıştır. Buna göre kısa testte tüm koşullarda diğer birey parametrelerine kıyasla en iyi kestirilen parametre teta3 iken, tüm koşullar göz önünde bulundurulduğunda en düşük hata ile kestirilen parametre uzun test koşulunda kestirilen teta1 parametresi olmuştur.



Şekil 4.29. Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Yanlılık Değerleri

Üç boyutlu basit yapıli testlerden elde edilen birey parametrelerine ilişkin yanlılık deęerleri 0,085 ile 0,010 arasında deęişen deęerler almıştır. Buna göre bu yapıdaki tüm birey parametresi kestirimleri pozitif yanlılık ile yapılmıştır. En yansız kestirilen parametre 48 madde 2000 örneklem koşulundaki teta2 parametresidir. En yanlı parametre ise 12 madde 1000 örneklem büyüklüğü koşulunda kestirilen teta3 parametresidir. Ancak yanlılık deęerlerinin çok düşük olması sebebiyle üç boyutlu basit yapıli testlerde birey parametrelerinin genel olarak benzer yanlılıkla kestirildięi sonucuna ulaşılabılır.



Şekil 4.30. Üç Boyutlu Basit Yapılı Testler İçin Birey Parametrelerine Ait Korelasyon Deęerleri

Üç boyutlu basit yapıları testlerden kestirilen birey parametrelerinin gerçek parametreler ile korelasyonlarının ortalamaları şekil 4.30'da yer almaktadır. Grafik genel itibari ile RMSE değerleri için çizilen grafik ile paralel bulgulara sahiptir. Kısa test için elde edilen korelasyonlar 0,480 ile 0,655 arasında; uzun test için elde edilen korelasyonlar 0,802 ile 0,835 arasında değerler almaktadır. Buna göre RMSE'nin yüksek olduğu kısa test koşulunda madde parametreleri için korelasyonlar düşük; RMSE'nin daha düşük değerler aldığı uzun test koşulunda korelasyonlar daha yüksek değerler almıştır.

4.4. Tartışma

Bu çalışmada tek, iki ve üç boyutlu yapıların madde ve birey parametresi kestirimlerinin farklı örneklem büyüklüğü ve test uzunluklarından nasıl etkilendiği RMSE, yanlışlık ve korelasyonlar ölçüt alınarak incelenmiştir. Örneklem büyüklüğünün ve test uzunluğunun parametre kestirimi üzerinde çok etkili olduğu bilinmektedir. Bu etkilerin niteliği ve niceliği pek çok çalışmanın konusu olmuştur. (Reise and Yu, 1990; Stone, 1992; Sinar ve Zickar, 2002; Bolt ve Lall, 2003; De Mars, 2003; de la Torre ve Patz, 2005; Sheng ve Wikle, 2007; Babcock, 2009; Finch, 2010; Finch, 2011; Sünbül, 2011; Lee, 2012; Kieftenbeld and Natesan, 2012; Zhang, 2012). Bu çalışmadan elde edilen bulgular madde ve birey parametrelerine ilişkin olmak üzere iki başlık altında tartışılmıştır.

4.4.1. Madde parametrelerine ilişkin tartışma

Reise ve Yu (1990) tek boyutlu testlerde MTK'nın iyi sonuçlar verebilmesi için en az 500 kişilik örneklem büyüklüğü ve 10 maddelik testlere ihtiyaç duyulduğunu belirtmektedir. Lee (2012) boyut sayısı arttıkça ele alınan örneklem büyüklüğünün artması gerektiğini belirtmiş ve 6 boyut için en az 2000; 3 boyut için en az 1000 kişilik örneklem büyüklüğü önermiştir. Hulin vd. (1982) 30 maddelik testte 2PLM'de yapılacak kestirimler için örneklem büyüklüğünün en az 500; 3PLM'de yapılacak kestirimler için örneklem büyüklüğünün en az 1000 kişi olması gerektiğini belirtmiştir. Bu çalışmada ele alınan koşullarda en az 1000 kişilik örneklem büyüklüğü ve 12 maddelik testler yer almaktadır. Tek boyutlu MTK ile yapılan analiz sonuçlarına göre 12 madde ve 1000 örneklem büyüklüğü koşulunda kestirilen ayırt edicilik ve güçlük parametrelerine ilişkin RMSE değerleri sırasıyla 0,125 ve 0,083 iken; 48 madde ve 1000 kişilik örneklem büyüklüğü koşulunda her iki RMSE değeri sırası ile 0,095 ve 0,079 olarak hesaplanmıştır. Drasgow ve Parsons (1983) çalışmalarında ayırt edicilik için 0,136 ve güçlük parametresi için 0,220 olan RMSE değerlerini kararlı kestirim sonuçları olarak

yorumlamıştır. Bu çalışmadan elde edilen bulguların Drasgow ve Parsons (1983)'ten daha kararlı yapılmış kestirim sonuçlarına işaret ettiği söylenebilir. Bu sebeple 12 madde ve 1000 kişilik örneklem büyüklüğünün tek boyutlu MTK'da parametrelerin doğru bir biçimde kestirilmesi için yeterli olduğu ifade edilebilir.

Çalışmadan elde edilen diğer bulgulara göre örneklem büyüklüğünün artması tek boyutlu testlerin yanı sıra hem iki hem de üç boyutlu testlerde madde parametrelerine ilişkin RMSE değerlerinin düşmesine sebep olmuştur. Bu bulgu Bolt ve Lall (2003), Finch (2010) ve Zhang (2012)'nin elde ettiği sonuçlar ile paralellik göstermektedir. Zhang (2012) boyutlar arası korelasyon ve madde sayısı sabit iken birey sayısı arttıkça parametrelerin RMSE değerlerinin azaldığını ifade etmiş ve birey sayısının artmasının hem tek hem çok boyutlu yapıların kestirimleri için iyi sonuçlar oluşturduğunu ifade etmiştir. Finch (2010) ise madde ve örneklem büyüklüğünün yüksek olduğu durumda MTK ile düşük hatalı ve kararlı kestirimler yapılabileceğini belirtmiş ve çalışmasında madde ayıricılığına ait RMSE değerlerinin ve güçlük parametresine ilişkin standart hataların birey ve madde sayısı arttıkça düştüğünü, ancak yanlılığın büyük örneklerde daha yüksek değerlere sahip olduğunu rapor etmiştir. Bu sebeple araştırmacı, çalışmasında, çok sayıda birey ve fazla sayıda madde kullanmanın her zaman doğru bir ÇBMTK parametre kestirimini sağlamayacağını ifade etmiştir. Ancak Finch (2010)'a ait bu bulgular Drasgow (1989)'un bulguları ile çelişmektedir. Drasgow çalışmasında yanlılığın örneklem büyüklüğü arttıkça azaldığını, ayrıca parametre kestirimlerinin MML ve JML ile yapıldığı durumda 500 ve 1000 örneklem büyüklüğü için sonuçların neredeyse yansız olduğunu belirtmiştir. Benzer biçimde Seong (1990) a ve b parametrelerinin yanlılık ve RMSE değerlerinin birey sayısının 100'den 1000'e çıkması ile neredeyse yarıya düştüğünü ifade etmiştir. Bu çalışmada da Finch (2010)'un çalışmasında elde edilen sonuçlara benzer durum söz konusudur ve bulgular Seong (1990) ve Drasgow (1989) ile çelişmektedir. Bu çalışmada yapılan analizlerin sonuçlarına göre madde parametrelerine ilişkin yanlılık değerleri örneklem büyüklüğünün ve/veya madde sayısının artmasına bağlı olarak düzenli bir artış veya azalış sergilememiştir. Örneğin 3 boyutlu basit yapıya kısa testten kestirilen a_1 parametresinin yanlılık değeri örneklem büyüklüğü 1000 iken 0,008; 2000 iken -0,014 ve 4000 iken 0,004 olarak hesaplanmıştır. Buna göre bu parametrenin kestirim yanlılığı örneklem büyüklüğünün ilk kez artışı (1000'den 2000'e) ile artmış, ikinci kez artışı (2000'den 4000'e) ile azalmıştır. Benzer durum farklı koşullarda diğer parametreler için de geçerlidir. Bu çalışmada ele alınan yapılara göre madde

parametrelerinin hesaplanan mutlak yanlılık değerlerinin alt ve üst sınırları incelenirse, boyut sayısı arttıkça ve yapı basitten karmaşığa geçtikçe yanlılık değerlerinin arttığı görülmektedir. Yapılara göre mutlak yanlılık değerlerinin alt ve üst sınırları şu şekilde hesaplanmıştır: Tek boyutlu yapılar için (0,001-0,013); iki boyutlu basit yapılar için (0,001- 0,023); iki boyutlu karmaşık yapılar için (0-0,197); üç boyutlu basit yapılar için (0,001-0,043) ve üç boyutlu karmaşık yapılar için (0,02-0,247)'dir. Bu değerlere göre yanlılık değerlerinin artışı boyut sayısı ve yapının karmaşık olup olmaması ile ilişkilidir. Bu bulgu Lee (2012)'nin ifadeleriyle paralellik göstermektedir. Lee (2012) çalışmasında ÇBMTK modelinin parametrelerinin kestirimini etkileyen faktörlerin artmasının kestirim yanlılığının artışına sebep olduğunu belirtmiştir.

Çalışmadan elde edilen bir diğer bulgu da elde edilen yanlılık ve RMSE değerlerinin aynı boyut sayısı, test uzunluğu ve örneklem büyüklüğü koşullarındaki basit yapı ve karmaşık yapı testlerde farklılık göstermesidir. Buna göre iki boyutlu basit yapılarda a_1 ve a_2 parametreleri için elde edilen yanlılık ve RMSE değerleri tüm test uzunluğu ve örneklem büyüklüğü koşullarında karmaşık yapıdan daha düşük değerdedir. İki boyutlu karmaşık yapı testlerde d parametresi daha kararlı kestirilmiştir ve bu parametreye ilişkin RMSE değerleri tüm koşullarda, yanlılık değerleri ise iki koşul (uzun testteki 1000 ve 4000 örneklem koşulları) dışındaki tüm koşullarda basit yapıdan daha düşük değerlere sahiptir. Bu bulguları doğrular biçimde madde parametrelerine ilişkin elde edilen korelasyon ortalamaları da a_1 ve a_2 parametreleri için basit yapı testte, d parametresi için karmaşık yapı testte daha yüksek değerde olma eğilimindedir.

Üç boyutlu yapılarda tüm örneklem büyüklüğü ve test uzunluğu koşullarında basit yapıdan elde edilen RMSE ve yanlılık değerleri, karmaşık yapıdan elde edilen RMSE ve yanlılık değerlerinden daha düşüktür. Bu durum sadece kısa test ve 1000 örneklem büyüklüğü koşulunda a_2 parametresinin RMSE değeri için istisna teşkil etmektedir. (Bu koşulda karmaşık yapıdan elde edilen a_2 RMSE değeri basit yapıdan elde edilen değerden 0,01 düşüktür.) Üç boyutlu yapılardan elde edilen madde parametrelerinin korelasyonlarının ortalaması ise yukarıdaki bulguyu destekler biçimde basit yapılarda daha yüksek değerlere sahiptir. Ancak uzun test koşulunda kestirilen d parametrelerinin gerçek d parametreleri ile korelasyonlarının ortalaması hem basit hem karmaşık yapıda bire bir aynı değerlere sahiptir. Tek boyutlu yapının tüm koşullarından elde edilen RMSE ve yanlılık değerleri ise diğer çok boyutlu kestirimlerden elde edilenlerden daha düşük değerdedir. Basit ve karmaşık yapıların

parametre kestirimlerinin RMSE, yanlılık ve korelasyonların bu şekilde farklılık göstermesi Finch (2011) ile tutarlık göstermektedir. Finch (2011) yaptığı çalışmada basit yapıda olmayan maddelerin basit yapıya göre daha yüksek yanlılık ve standart hata ile kestirildiğini, basit yapıdaki maddelerin daha kararlı kestirildiğini ifade etmiştir.

Bu çalışmadan madde parametrelerine ilişkin olarak elde edilen diğer bir bulgu ele alınan koşulların büyük bir kısmında d parametresine ait RMSE değerlerinin aynı koşullarda kestirilen a parametrelerinden daha düşük, d parametresine ait korelasyon ortalamalarının a parametrelerinininkinden daha yüksek olmasıdır. Bu bulgu Zhang (2005) ve Sheng ve Wikle (2007) tarafından yapılan çalışmalar ile örtüşmektedir. Zhang (2005) d parametrelerinin daha kararlı kestirildiğini ve Sheng ve Wikle (2007) ise ele aldıkları koşullarda d parametrelerine ait RMSE'lerin a parametrelerinininkinden daha düşük olduğunu, dolayısıyla d parametrelerinin a parametrelerinden daha kararlı kestirildiğini ifade etmişlerdir.

4.4.2. Birey parametrelerine ilişkin tartışma

Bu çalışmada ele alınan koşullarda madde parametrelerinin yanı sıra birey parametreleri de kestirilmiştir. Buna göre; ele alınan tüm boyutluluk koşullarında madde sayısı sabit tutulduğunda birey sayısının artması birey parametrelerinin (tetaların) kestirim iyiliğini etkilememiş, teta parametrelerinin RMSE değerlerinde ciddi bir fark meydana gelmemiştir. Buna paralel olarak gerçek teta parametreleri ile kestirilen teta parametreleri arasındaki korelasyonlarının ortalaması da örneklem büyüklüğü artışından etkilenmemiştir. Ancak birey parametrelerinin kestirim iyiliği test uzunluğunun artması ile değişim göstermiştir. Buna göre 12 maddeli testten 48 maddeli teste geçildiğinde birey parametresi kestirimine ilişkin RMSE değerlerinde ciddi bir azalma meydana gelmiştir. Bu bulgu literatürde yer alan birey parametrelerinin örneklem büyüklüğünden etkilenmediği bulgusu ile paralellik göstermektedir. Reise ve Yu (1990), Seong (1990), Kirisci, Hsu ve Yu (2001), de la Torre ve Patz (2005), Köse (2010), Kieftenbeld ve Natesan (2012) çalışmalarında örneklemin büyümesinin birey parametrelerinin daha iyi kestirilmesine bir etkisinin olmadığını, ancak test uzunluğunun artmasıyla teta kestirim iyiliğinin geliştirilebileceğini belirtmişlerdir.

Çalışmadan elde edilen sonuçlara göre madde sayısının artmasıyla gerçek birey parametreleri ile kestirilen birey parametreleri arasındaki korelasyonlar yükselmiştir. Bu bulgu de la Torre ve Patz (2005)'in bulguları ile tutarlıdır. de la Torre ve Patz (2005) yaptıkları çalışmada madde sayısı, test uzunluğu ve boyut sayısı koşulları tek başına

arttığında elde edilen korelasyonların da arttığını göstermişlerdir. Sheng ve Wikle (2007) ise yaptıkları çalışmada boyutlar arası korelasyonu değişen iki boyutlu basit yapıların teta parametrelerini incelemişlerdir. Buna göre bu çalışmaya en yakın koşulda teta1 ve teta2 parametrelerinin gerçek ve kestirilen değerleri arasındaki korelasyonu 0,80 ve 0,86 olarak bulmuşlardır. Eldeki çalışmada da Sheng ve Wikle (2009)'un koşullarına benzeyen iki boyutlu basit yapıda uzun test koşulunda elde edilen korelasyonlar 0,86 civarındadır. Bu açıdan elde edilen korelasyon değerleri literatür ile tutarlık göstermektedir.

Madde parametrelerine benzer biçimde birey parametrelerinde de boyut sayısı arttıkça ve yapı basitten karmaşığa geçtikçe aynı koşullar için elde edilen RMSE değerleri artmış ve korelasyon değerleri azalmıştır. Birey parametresi için tek boyutlu MTK ve basit yapılı ÇBMTK kestirimleri sonunda pozitif yanlılıklar elde edilmiş diğer bir ifade ile birey parametreleri için üst kestirimde bulunulmuştur. Boyut sayısı arttıkça hesaplanan yanlılık değerlerinin ranjı genişlemiştir. Ele alınan yapılara göre mutlak yanlılık değerlerinin alt ve üst sınırları şu şekilde hesaplanmıştır: Tek boyutlu yapılar için (0,012-0,048); iki boyutlu basit yapılar için (0,001-0,023); iki boyutlu karmaşık yapılar için (0,05-0,21); üç boyutlu basit yapılar için (0,001-0,085) ve üç boyutlu karmaşık yapılar için (0,001-0,007)'dir. Bu sebeple birey parametresi için kestirimi etkileyen faktör sayısının artmasının kestirim iyiliğini düşürdüğü söylenebilir. Bu bulgu Lee (2012)'nin ÇBMTK'da madde parametrelerinin kestirim yanlılığı ile ilgili bulgularıyla paralellik göstermektedir.

5. SONUÇ ve ÖNERİLER

Bu bölümde araştırmanın bulgu ve yorumlarına dayalı olarak ulaşılan sonuçların özetine ve bu sonuçlardan yola çıkarak geliştirilen önerilere yer verilmiştir.

5.1. Sonuçlar

Araştırmadan elde edilen sonuçlar alt problemlere ilişkin başlıklar halinde verilmiştir.

5.1.1. Birinci Alt Probleme İlişkin Sonuçlar

Araştırmanın birinci alt probleminde tek boyutlu yapıların madde ve birey parametreleri farklı örneklem büyüklüğü ve test uzunluğu koşullarında kestirilmişti. Elde edilen sonuçlara göre örneklem büyüklüğünün ve/veya test uzunluğunun artması madde parametrelerinin RMSE değerlerinde düşüş ve gerçek parametrelerle olan korelasyonlarında artış sağlamıştır. Tek boyutlu yapılarda değişen madde sayısı ve test uzunluğu koşulları göz önünde bulundurulduğunda madde parametrelerine ait yanlılıklar sıfıra yakın değerlerde hesaplanmıştır. Tek boyutlu yapılara ilişkin olarak elde edilen madde parametresi kestirimleri literatür ile birlikte değerlendirildiğinde, ele alınan tüm koşullardaki kestirimlerin kararlı ve yansız olduğu sonucu çıkarılabilir.

Çalışmada ele alınan tek boyutlu yapılarda, birey parametreleri madde parametrelerinden daha yüksek RMSE ve yanlılık değerleri ile kestirilmiştir. Birey parametreleri kestiriminde örneklemin artması kestirimleri çok etkilememiş, ancak 12 maddelik testten 48 maddelik teste geçildiğinde RMSE ve yanlılık değerlerinde düşme yaşanmıştır. Ayrıca kısa testte birey parametreleri için elde edilen korelasyonlar 0,76-0,77 iken, uzun test koşulunda elde edilen korelasyonlar 0,92'dir. Buna göre test uzunluğunun artması ile birey parametresi kestirimleri daha kararlı yapılmıştır.

5.1.2. İkinci Alt Probleme İlişkin Sonuçlar

Bu çalışmada iki boyutlu testler basit ve karmaşık yapıları olmak üzere iki başlık altında ele alınmıştır. Karmaşık yapıları testlerde, testte yer alan tüm maddeler her iki boyutla ilişkiliyken; basit yapıda testin ilk yarısı birinci, ikinci yarısı ikinci boyutla ilişkilidir. Analiz sonuçlarına göre her iki yapıda da hem birey hem de madde sayısının artmasıyla madde parametrelerine ait RMSE'ler düşmüş, korelasyonlar artmıştır. Basit yapıda elde edilen madde parametresi yanlılıkları karmaşık yapıdan daha yüksek değerlerde çıkmıştır. a_1 ve a_2 parametrelerine ait RMSE değerleri tüm koşullarda karmaşık yapıda basit yapıdan daha yüksek çıkmıştır. d parametresi ise her iki yapıda

da a_1 ve a_2 parametrelerinden daha düşük RMSE ve sifıra yakın yanlılık ile kestirilmiştir. Karmaşık yapıda a parametrelerine ait korelasyon değerleri basit yapıdan daha düşük değerler almıştır. Buna karşılık d parametreleri iki yapıda da gerçek parametrelerle pozitif yönde mükemmel korelasyon vermiştir.

İki boyutlu yapıda birey parametreleri, örneklem büyüklüğünün değişmesinden etkilenmemiş ancak test uzunluğunun artması ile daha düşük hatalı kestirilmiştir. Basit yapıdaki birey parametreleri, kısa test koşulunda, karmaşık yapıdan daha yüksek RMSE ile kestirilmiştir. Her iki yapıda da uzun test koşuluna geçildiğinde tetalar benzer hatalar ile kestirilmiştir. Birey parametrelerine ilişkin yanlılıklar da genel olarak karmaşık yapıda daha düşük değerlerde çıkmıştır. Neredeyse tüm koşullarda teta parametrelerinin yanlılık değerleri 0,1'in altındadır. Bu sebeple her iki yapıda da birey parametresi kestirimlerinin çok düşük yanlılıkla yapıldığı söylenebilir. Birey parametrelerine ilişkin korelasyonlar da RMSE gibi örneklem büyüklüğü artışından pek etkilenmemiş, test uzunluğunun artmasıyla korelasyonlarda artış gerçekleşmiştir. Sonuçlara göre her iki yapının da kısa testinde orta, uzun testinde ise iyi derecede korelasyon ortaya konmuştur. Dolayısıyla her iki yapıda da gerçek parametreler ile kestirilen parametrelerin benzer uyuma sahip olduğu söylenebilir.

5.1.3. Üçüncü Alt Probleme İlişkin Sonuçlar

Bu çalışmada üç boyutlu testler de iki boyutlu testlere benzer biçimde basit ve karmaşık yapıları olmak üzere iki başlık altında ele alınmıştır. Karmaşık yapıları testlerde, testte yer alan tüm maddeler üç boyutla da ilişkililyken; basit yapıda testin ilk 1/3'ü birinci, ikinci 1/3'ü ikinci, son 1/3'ü ise üçüncü boyutla ilişkiliydi. Üç boyutlu yapıları madde parametreleri örneklem büyüklüğünün ve madde sayısının değişiminden etkilenmiş, madde ve birey sayısındaki artış ile daha düşük RMSE'li ve daha yüksek korelasyon veren kestirimler elde edilmiştir. Karmaşık yapıda özellikle 12 maddeli test koşulunda çok yüksek RMSE değerleri ($>1,0$) elde edilmiştir. Buna göre karmaşık yapıdaki 3 boyutlu ve 12 maddeli testin parametre kestirimi için çok kısa olduğu sonucuna ulaşılabilir. Üç boyutlu yapıları da d parametreleri genel olarak a parametrelerinden daha düşük hata ile kestirilmiştir. Basit ve karmaşık yapının aynı koşullarında hesaplanan yanlılıklar genel olarak karmaşık yapıda daha yüksek değerlerde olup, basit yapıda daha az yanlı kestirimler yapılmıştır. Kestirilen madde parametrelerinin gerçek parametreler ile verdiği korelasyonlar, RMSE ve yanlılık sonuçlarına paralellik göstermiştir. Buna göre a parametrelerine ait korelasyonlar basit yapıda karmaşık

yapıdan daha yüksek değerler almış, d parametresi ise hem basit hem karmaşık yapıda yüksek korelasyon değerlerine sahip çıkmıştır.

Üç boyutlu yapılarda birey parametrelerinin RMSE değerleri her ne kadar basit ve karmaşık yapıda benzer aralıkta yer alsalar da, genel olarak karmaşık yapılarda basit yapıdan daha düşük RMSE içeren kestirimler yapılmıştır. Karmaşık yapı ile basit yapı arasındaki bu RMSE farkları 0,1 ve altındadır. Birey parametreleri basit yapıda tüm koşullarda, karmaşık yapıda ise üç koşul dışındaki tüm koşullarda pozitif yanlılık ile kestirilmiştir. Aynı koşullarda basit yapıda yapılan kestirimler çok düşük farklarla ($\leq 0,045$) karmaşık yapıdan daha yanlıdır. Gerçek birey parametreleri ile kestirilen birey parametreleri arasındaki korelasyonlar madde sayısının artması ile artmış, örneklem büyüklüğünün artmasından ise pek etkilenmemiştir. Korelasyonlar her iki yapıda da orta ve iyi korelasyona karşılık gelen değişen değerler almıştır. Çalışmada ele alınan iki boyutlu yapılarda korelasyon değerlerinin genel olarak RMSE ile paralel sonuçlar ürettiği gözlenmişti. Üç boyutlu yapılarda da benzer bir durum gözlenmiş olup, RMSE'nin düşük olduğu koşullarda korelasyonlar yüksek çıkmıştır.

5.2. Öneriler

Araştırma sonuçlarından yola çıkarak geliştirilen öneriler iki başlık halinde aşağıda verilmiştir.

5.2.1. Araştırmacılara Yönelik Öneriler

1. Madde ve birey sayısı, boyut sayısı fazla olan testlerde parametre kestirim kararlığı üzerine çok etkilidir. Boyut sayısı arttıkça madde parametresi kestirimlerinin iyileştirilmesi için madde ve birey; birey parametresi kestirimlerinin iyileştirilmesi için madde sayısı arttırılmalıdır.
2. Karmaşık yapıli testlerde kararlı kestirimler yapmak için, aynı boyut sayısına sahip basit yapıdaki testlerde kararlı kestirim yapılmasını sağlayan madde ve birey sayısından daha fazla madde ve bireye gerek vardır.
3. BA-EM ile yapılacak kestirimlerde madde parametrelerinin RMSE değerlerini makul düzeyde tutabilmek açısından bulgularından yola çıkarak şu öneriler geliştirilmiştir:
 - o Tek boyutlu madde tepki kuramında 12 madde ve 1000 kişilik örneklem kullanmak kararlı kestirim sonuçları elde etmek için yeterlidir.

- İki boyutlu basit yapılar için en az 12 maddelik testler kullanıldığı durumda 2000'den büyük örnekleme ihtiyaç vardır. İki boyutlu karmaşık yapılar için 12 maddelik testlerde tatmin edici sonuçlara ulaşılamamaktadır. Bu yapılarda 48 maddenin yer aldığı testler ve 1000 kişilik örneklem büyüklüğü kullanıldığında kabul edilebilir düzeyde RMSE sonuçlarına ulaşılabilir.
 - Üç boyutlu basit yapılarda 4000 örneklem ve üstünde 12 maddelik kısa testlerden faydalanılabilir. Üç boyutlu karmaşık yapılarda ise 48 madde ve 4000 örneklemin üstüne çıkılmasına ihtiyaç vardır.
4. Birey parametresi kestirimlerini incelemek amacıyla hesaplanan RMSE değerleri genel olarak madde parametrelerinin RMSE değerlerinden yüksektir. Bu çalışmada hesaplanan korelasyon değerleri göz önünde bulundurularak tek boyutlu yapılar, iki boyutlu (basit ve karmaşık) yapılar ve üç boyutlu basit yapılar için 12 maddelik, üç boyutlu karmaşık yapılarda ise 48 maddelik testlerin kullanılması önerilmektedir.
5. Bu çalışmada ele alınan istatistiklerden olan RMSE, parametre kestirim iyiliğini test etmek için başlangıç parametreleri ile kestirilen parametre değerlerini hesaba katması nedeniyle duyarlı bir istatistiktir. Elde edilen bulgulara göre RMSE ve gerçek parametrelerle kestirilen parametreler arasındaki basit korelasyonlar genel olarak birbirine paralel sonuçlar üretmiştir. Bu sebeplerle kestirim kararlılığını belirlemede genel bir bilgi edinmek için korelasyonların, daha hassas bilgi sağlamak ve hatanın büyüklüğünü anlamak için RMSE'nin hesaplanması önerilmektedir.

5.2.2. İleri Araştırmalara Yönelik Öneriler

1. Bu çalışmada madde sayısının etkisi kısa ve uzun testi temsil etmek üzere 12 ve 48 maddelik testlerle ortaya konmaya çalışılmıştır. Ancak yüksek boyut sayılarında farklı uzunluktaki testlerin parametre kestirimi üzerinde etkisini görebilmek amacıyla farklı sayıda madde içeren testler ile çalışılabilir.
2. Çalışmada 3 örneklem büyüklüğü ele alınmıştır. Örneklem büyüklüğünü etkisini daha detaylı incelemek için farklı örneklem büyüklükleri ile çalışılabilir.
3. Bu çalışmada, çok boyutlu yapılarda boyutlar arası düşük korelasyon ile çalışılmış, bu nedenle boyutlar arası korelasyonlar 0,3 olarak sabitlenmiştir. Gelecekte

boyutların ikili kombinasyonları arasındaki korelasyonları farklılaşan yapılar ile çalışılabilir.

4. Bu çalışmada 2PL modeller ile çalışılmıştır. İleri çalışmalarda 3PLM kullanılarak sonuçlar incelenebilir.
5. Çalışmada yetenek parametreleri normal dağılımdan üretilmiştir. Kestirim üzerine yetenek dağılımının etkisini görebilmek için farklı dağılımlar kullanılarak çalışmalar yapılabilir.
6. Araştırmada parametre kestirimleri için IRTPRO 2.1 kullanılmıştır. Program farklı model ve yapıların birey ve madde parametrelerini kestirmeye yönelik araştırma problemlerine tek başına cevap verebilecek niteliktedir. Ancak boyut sayısı arttıkça kestirim süresi logaritmik olarak artmaktadır. Araştırmacıların kestirim süresini göz önünde bulundurarak planlama yapmaları önerilmektedir.

KAYNAKÇA

- Ackerman, T.A. (1989). Unidimensional IRT calibration of compensatory and noncompensatory multidimensional items. *Applied Psychological Measurement*, 13(2), 113-127.
- Ackerman, T.A. (1994). Using multidimensional item response theory to understand what items and tests are measuring. *Applied Measurement In Education*, 7(4), 255-278.
- Ackerman, T.A. (1996). Graphical representation of multidimensional item response theory analyses. *Applied Psychological Measurement*, 20(4), 311-329.
- 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.
- Ansley, T. N., & Forsyth, R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two-dimensional data. *Applied Psychological Measurement*, 9(1), 37-48.
- Babcock, B. G. E. (2009). *Estimating a noncompensatory IRT model using a modified Metropolis Algorithm*. Unpublished Doctoral Dissertation. University Of Minnesota Faculty Of The Graduate School.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: Pegem Yayınları.
- Bolt, D. M., & Lall, V. F. (2003). Estimation of compensatory and noncompensatory multidimensional item response models using Markov Chain Monte Carlo. *Applied Psychological Measurement*, 27(6), 395-414.
- Bulut, O. (2013). *Between-person and within-person subscore reliability: Comparison of unidimensional and multidimensional IRT models*. Unpublished Doctoral Dissertation. University Of Minnesota Faculty Of The Graduate School.
- Choi, S.-W. (1996) *A response dichotomization technique for item parameter estimation of the multidimensional graded response model*. Unpublished Doctoral Dissertation. Faculty of the Graduate School of The University of Texas at Austin.
- de la Torre, J., & Patz, R. L. (2005). Making the most of what we have: A practical application of multidimensional item response theory in test scoring. *Journal of Educational and Behavioral Statistics*, 30(3), 295-311.
- DeMars, C.E. (2003). Sample size and the recovery of nominal response model item parameters. *Applied Psychological Measurement*, 27(4), 275-288.
- Drasgow, F. (1989). An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model. *Applied Psychological Measurement*, 13(1), 77-90.
- Drasgow, F., & Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, 7(2), 189-199.
- Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associate, Inc.

- Finch, H. (2010) Item parameter estimation for MIRT model: Bias and precision of confirmatory factor analysis based models. *Applied Psychological Measurement*, 34(1), 10-26.
- Finch, H. (2011) Multidimensional item response theory parameter estimation with nonsimple structure items. *Applied Psychological Measurement*, 35(1), 67-82.
- Gilbert, N. (1999). Simulation: A new way of doing social sciences. *American Behavioral Scientist*, 42(10), 1485-1487.
- Hambleton, R. K., & Jones, R. W. (1993). An NCME module on comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, 12(3), 38–47.
- Hambleton, R.K., & Swaminathan, H. (1985). *Item response theory principles and applications*. Kluwer-Nijhoff Publishing. Boston-USA.
- Hartig, J., & Höhler, J. (2009). Multidimensional IRT models for the assessment of competencies. *Studies in Educational Evaluation*, 35(2009), 57–63
- Harvey, R.J., & Hammer, A. L. (1999). Item response theory. *The Counselling Psychologist*, 27(3), 353-383.
- Harwell, M., Stone, C.A., Hsu, T.C., & Kirisci L. (1996). Monte Carlo studies in item response theory. *Applied Psychological Measurement*, 20(2), 101-125.
- Hulin, C. L., Lissak, R. I., & Drasgow, F. (1982). Recovery of two- and three-parameter logistic item characteristic curves: A Monte Carlo study. *Applied Psychological Measurement*, 6(3), 249-260.
- IRTPRO: Users Guide (2011). *Scientific Software International*.
- IRTPRO 2.1 for Windows (Item Response Theory for Patient-Reported Outcomes). *Scientific Software International*.
- Kieftenbeld, V., & Natesan, P. (2012). Recovery of graded response model parameters: A comparison of marginal maximum likelihood and Markov Chain Monte Carlo Estimation. *Applied Psychological Measurement*, 36(5), 399-419.
- Kirisci, L., Hsu, T., & Yu, L. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied Psychological Measurement*, 25(2), 146-162.
- Koğar, H. (2014). *Madde tepki kuramının farklı uygulamalarından elde edilen parametrelerin ve model uyumlarının örneklem büyüklüğü ve test uzunluğu açısından karşılaştırılması*. Yayımlanmamış Doktora Tezi. Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü.
- Köse, İ.A. (2010). *Madde tepki kuramına dayalı tek boyutlu ve çok boyutlu modellerin test uzunluğu ve örneklem büyüklüğü açısından karşılaştırılması*. Yayımlanmamış Doktora Tezi. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Lee, J. (2012). *Multidimensional item response theory: An investigation of interaction effects between factors on item parameter recovery using Markov Chain Monte Carlo*. Unpublished Doctoral Dissertation. Michigan State University Measurement and Quantitative Methods.

- Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, 64(2), 187-212.
- McDonald, P., & Paunone, S. V. (2002). A Monte Carlo comparison of item and person statistics based on item response theory versus classical test theory. *Educational and Psychological Measurement*, 62(6), 921-943.
- Md Desa, Z. N. D. (2012) *Bi-factor multidimensional item response theory modeling for subscores estimation, reliability, and classification*. Unpublished Doctoral Dissertation. Department of Psychology and Research in Education and the Faculty of the Graduate School of the University of Kansas.
- Özkan, Y. Ö. (2012). *Öğrenci başarılarının belirlenmesi sınavından (ÖBBS) klasik test kuramı, tek boyutlu ve çok boyutlu madde tepki kuramı modelleri ile kestirilen başarı puanlarının karşılaştırılması*. Yayınlanmamış Doktora Tezi. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Reckase, M. D. (2009). *Multidimensional item response theory (Statistics for social and behavioral sciences)*. New York: Springer.
- Reckase, M.D., & McKinley, R. L. (1999). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement*, 15(4), 361-373.
- Reise, S. P., & Yu, J. (1990). Parameter recovery in graded response model using MULTILOG. *Journal of Educational Measurement*, 27(2), 133-144.
- Seong, T.-J. (1990). Sensitivity of marginal maximum likelihood estimation of item and ability parameters to the characteristics of the prior ability distributions. *Applied Psychological Measurement*, 14(3), 299-311.
- Sheng Y. (2005). *Bayesian analysis of hierarchical IRT models: Comparing and combining the unidimensional & multi-unidimensional IRT models*. Unpublished Doctoral Dissertation. Faculty of the Graduate School University of Missouri-Columbia.
- Sheng Y. and Wikle C. K. (2007). Comparing multiunidimensional and unidimensional item response theory models. *Educational and Psychological Measurement*, 68(3), 413-430.
- SimuMIRT (2003). Software. *Lihua Yao*.
- Sinar, E.F., & Zickar, M. J. (2002). Evaluating the robustness of graded response model and classical test theory parameter estimates to deviant items. *Applied Psychological Measurement*, 26(2), 181-191.
- Stone, C. A. (1992). Recovery of marginal maximum likelihood estimates in the two-parameter logistic response model: An evaluation of MULTILOG. *Applied Psychological Measurement*, 16(1), 1-16.
- Sünbül, Ö. (2011). *Çeşitli boyutluluk özelliklerine sahip yapılarda, madde parametrelerinin değişmezliğinin klasik test teorisi, tek boyutlu madde tepki kuramı ve çok boyutlu madde tepki kuramı çerçevesinde incelenmesi*. Yayınlanmamış Doktora Tezi. Mersin Üniversitesi Eğitim Bilimleri Enstitüsü.
- Turhan, A. (2006) *Multilevel 2PL item response model vertical equating with the presence of differential item functioning*. Unpublished Doctoral Dissertation. Florida State University.

- Way, W. D., Ansley, T. N., & Forsyth, R. A. (1988). Unidimensional IRT estimates the comparative effects of compensatory and noncompensatory two-dimensional data on unidimensional IRT estimates. *Applied Psychological Measurement*, 12(3), 239-252.
- Yao, L., & Boughton, K. A. (2007). A multidimensional item response modeling approach for improving subscale proficiency estimation and classification. *Applied Psychological Measurement*, 31(2), 83-105.
- Zhang, J. (2012). Calibration of response data using MIRT models with simple and mixed structures. *Applied Psychological Measurement*, 36(5), 375-398.
- Zhang, B., & Stone, C. (2004). *Direct and indirect estimation of three-parameter compensatory multidimensional item response models*. Paper presented in annual meeting of the American Educational Research Association, San Diego, CA.

EKLER DİZİNİ

EK 1. SIMUMIRT İLE ÖRNEK VERİ ÜRETME KODU

Basit yapıda iki boyutlu 12 maddeli testin madde parametrelerini oluşturmak için kullanılan SimuMIRT ItemPool1.ctl dosyası kodu:

```
12 2 6 6 59494 0.5 0.4 0 1 6 16 0.4 1.8 -2 2
```

Basit yapıda iki boyutlu 12 maddeli testin madde parametrelerini oluşturmak için kullanılan SimuMIRT simudriver.bat dosyası kodu:

```
REM simulate item of simple structure , all MC items  
for %%c in (1) do call SimulateSimpleItemParam ItemPool%%c.ctl  
out\SimpleItemPool%%c  
pause;
```

Basit yapıda iki boyutlu 12 maddeli teste ilişkin iki kategorili veri seti ve yetenek parametresi üretmek için kullanılan .par dosyası kodu:

```
12      1000 2 0 0 1 0.3 0.3 1 2 9999 999001  
  
1      1      0.3938      0      1.2007      0  
2      1      1.0069      0      -0.9809      0  
3      1      0.3535      0      1.6941      0  
4      1      1.4552      0      -0.7435      0  
5      1      1.0139      0      1.3928      0  
6      1      0.6928      0      -1.6081      0  
7      1      0      1.1903      0.0726      0  
8      1      0      0.9725      1.4293      0  
9      1      0      0.5442      0.4173      0  
10     1      0      0.3303      0.0001      0
```


25 tekrar ile basit yapıda iki boyutlu 12 maddeli teste ilişkin iki kategorili veri seti ve yetenek parametresi üretmek için kullanılan driver.bat dosyası kodu:

```
REM MIRT compensatory model
for %%f in (1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25) do
call SimulateRwo %%f.par out\%%f
pause;
```

EK 2. VERİ ÜRETMEDE KULLANILAN MADDE PARAMETRELERİ

12 maddelik testler için kullanılan madde parametreleri

Madde No	Bir Boyutlu Test		İki Boyutlu Test						Üç Boyutlu Test							
	a	d	Karmaşık Yapı			Basit Yapı			Karmaşık Yapı				Basit Yapı			
	a	d	a1	a2	d	a1	a2	d	a1	a2	a3	d	a1	a2	a3	d
1	0,4039	-1,2007	0,3416	0,4456	-1,2007	0,4039	0	-1,2007	0,4892	1,3629	0,5618	-1,6641	0,4039	0	0	-1,2007
2	0,9352	0,9809	1,0069	0,3535	0,9809	0,9352	0	0,9809	0,7755	0,52	0,2194	-0,1685	0,9352	0	0	0,9809
3	1,3	-1,6941	0,1831	1,4163	-1,6941	1,3	0	-1,6941	0,1537	0,2599	1,4209	-0,643	1,3	0	0	-1,6941
4	0,4256	0,7435	0,9513	0,5962	0,7435	0,4256	0	0,7435	0,2076	0,9811	0,3948	1,2644	0,4256	0	0	0,7435
5	0,8853	-0,9998	0,717	0,8228	-1,6127	0,8853	0	-0,9998	2,0636	0,8316	0,7107	-1,6236	0	0,8853	0	-0,9998
6	0,5593	-0,4612	0,5264	1,3737	1,1348	0,5593	0	-0,4612	0,9967	0,2663	0,5006	-2,1814	0	0,5593	0	-0,4612
7	0,5587	-1,2266	1,1903	0,9725	-0,1771	0	0,5587	-1,2266	1,5009	0,9752	0,164	0,6891	0	0,5587	0	-1,2266
8	0,5394	-0,9409	0,5006	2,7779	-0,1756	0	0,5394	-0,9409	0,1592	0,7713	0,2267	-0,3066	0	0,5394	0	-0,9409
9	1,0456	-1,147	1,1252	0,7491	-0,1199	0	1,0456	-1,147	0,2318	0,5206	0,8222	1,8655	0	0	1,0456	-1,147
10	0,763	-0,5716	0,8709	0,8414	0,5864	0	0,763	-0,5716	1,2121	0,2101	0,36	0,9626	0	0	0,763	-0,5716
11	0,5702	-1,3187	0,1774	0,4626	1,2624	0	0,5702	-1,3187	0,32	1,3127	0,7814	-0,605	0	0	0,5702	-1,3187
12	1,5468	-0,5683	0,812	0,7342	0,9234	0	1,5468	-0,5683	0,5562	0,4994	0,4789	-1,2012	0	0	1,5468	-0,5683

48 maddelik testler için kullanılan madde parametreleri

Madde No	Bir Boyutlu Test		İki Boyutlu Test						Üç Boyutlu Test							
	a	d	Karmaşık Yapı			Basit Yapı			Karmaşık Yapı				Basit Yapı			
			a1	a2	d	a1	a2	d	a1	a2	a3	d	a1	a2	a3	d
1	0,4039	-1,2007	0,3416	0,4456	-1,2007	0,4039	0	-1,2007	0,3416	0,4456	1,0069	-1,2007	0,4039	0	0	-1,2007
2	0,9352	0,9809	1,0069	0,3535	0,9809	0,9352	0	0,9809	0,3535	1,6566	0,2956	0,9809	0,9352	0	0	0,9809
3	1,3	-1,6941	0,1831	1,4163	-1,6941	1,3	0	-1,6941	0,9513	0,5962	1,564	1,4203	1,3	0	0	-1,6941
4	0,4256	0,7435	0,9513	0,5962	0,7435	0,4256	0	0,7435	0,2241	0,717	0,8228	-1,4725	0,4256	0	0	0,7435
5	0,8853	-0,9998	0,717	0,8228	-1,6127	0,8853	0	-0,9998	0,5667	0,5661	1,1903	-0,0726	0,8853	0	0	-0,9998
6	0,5593	-0,4612	0,5264	1,3737	1,1348	0,5593	0	-0,4612	0,9725	0,5006	2,7779	-1,4293	0,5593	0	0	-0,4612
7	0,5587	-1,2266	1,1903	0,9725	-0,1771	0,5587	0	-1,2266	1,1252	0,7491	0,8709	-0,1199	0,5587	0	0	-1,2266
8	0,5394	-0,9409	0,5006	2,7779	-0,1756	0,5394	0	-0,9409	0,8414	0,2048	0,2603	0,5864	0,5394	0	0	-0,9409
9	1,0456	-1,147	1,1252	0,7491	-0,1199	1,0456	0	-1,147	0,812	0,7342	0,5832	1,4658	1,0456	0	0	-1,147
10	0,763	-0,5716	0,8709	0,8414	0,5864	0,763	0	-0,5716	0,2926	0,7161	0,6581	0,1098	0,763	0	0	-0,5716
11	0,5702	-1,3187	0,1774	0,4626	1,2624	0,5702	0	-1,3187	1,6822	0,5781	0,2245	1,2745	0,5702	0	0	-1,3187
12	1,5468	-0,5683	0,812	0,7342	0,9234	1,5468	0	-0,5683	1,7673	0,2139	0,7445	-0,837	1,5468	0	0	-0,5683
13	0,4034	1,2009	0,7161	0,6581	-0,2177	0,4034	0	1,2009	0,3879	0,3024	0,5901	0,7768	0,4034	0	0	1,2009
14	1,281	-0,5631	0,203	0,9037	0,7577	1,281	0	-0,5631	0,2648	0,3794	0,5258	-1,2098	1,281	0	0	-0,5631
15	1,5156	-0,5179	0,2245	1,7673	-1,7158	1,5156	0	-0,5179	0,2347	0,9026	0,0914	0,5208	1,5156	0	0	-0,5179
16	0,4298	0,885	0,2139	0,7445	-0,2054	0,4298	0	0,885	0,1143	0,2494	0,4222	-0,4608	0,4298	0	0	0,885
17	0,7432	-0,7044	0,3879	0,3024	0,7768	0,7432	0	-0,7044	0,8683	0,3927	0,1808	-0,7044	0	0,7432	0	-0,7044
18	1,197	0,3408	0,5901	0,2648	-1,2098	1,197	0	0,3408	0,2059	0,1681	0,7788	0,3408	0	1,197	0	0,3408
19	0,4768	-0,1423	0,346	0,6926	0,3903	0,4768	0	-0,1423	0,2121	0,3241	0,6376	-0,1255	0	0,4768	0	-0,1423
20	0,779	0,0979	0,2347	0,9026	-0,0712	0,779	0	0,0979	0,5586	0,6807	0,4849	0,2896	0	0,779	0	0,0979
21	0,516	-1,0511	0,2494	0,4222	2,4033	0,516	0	-1,0511	0,1386	0,3247	1,8708	-1,3323	0	0,516	0	-1,0511
22	0,4445	1,874	0,8228	0,3929	2,0872	0,4445	0	1,874	2,7608	0,544	0,3071	2,6318	0	0,4445	0	1,874
23	1,0282	-1,435	0,1808	0,2059	-0,7806	1,0282	0	-1,435	0,4384	1,1195	0,1238	0,5952	0	1,0282	0	-1,435
24	0,4046	1,7339	0,1681	0,7788	0,3415	0,4046	0	1,7339	0,1493	0,9937	0,3024	0,1319	0	0,4046	0	1,7339
25	0,7069	0,5798	0,2121	0,3241	-0,1255	0	0,7069	0,5798	0,7644	0,3825	0,5405	1,1552	0	0,7069	0	0,5798
26	0,9028	0,8369	0,6376	0,5586	0,2896	0	0,9028	0,8369	1,0076	0,9681	0,3237	0,326	0	0,9028	0	0,8369
27	1,1902	-0,0864	1,2826	0,0778	-0,4364	0	1,1902	-0,0864	0,2378	0,91	0,8443	1,5054	0	1,1902	0	-0,0864
28	0,5896	-0,8428	0,1386	0,3247	0,0434	0	0,5896	-0,8428	1,2166	0,6012	0,2672	-1,2833	0	0,5896	0	-0,8428
29	0,6943	0,8228	0,544	0,3071	-1,8661	0	0,6943	0,8228	0,7217	3,3971	0,2373	-0,359	0	0,6943	0	0,8228
30	0,6175	1,0542	0,3282	0,4555	-2,4164	0	0,6175	1,0542	0,7989	0,4299	0,2509	-0,4579	0	0,6175	0	1,0542
31	1,0045	-0,6628	0,1238	0,1493	0,1859	0	1,0045	-0,6628	1,145	0,3857	1,2411	1,5319	0	1,0045	0	-0,6628
32	0,904	-0,2726	0,9937	0,3024	-1,1399	0	0,904	-0,2726	0,6873	1,3442	0,4101	0,9918	0	0,904	0	-0,2726
33	0,4306	-1,4017	0,7644	0,3825	1,1552	0	0,4306	-1,4017	1,311	0,3864	1,0208	-1,0809	0	0	0,4306	-1,4017
34	0,4365	0,3358	0,5405	1,0076	0,326	0	0,4365	0,3358	0,2757	0,1925	0,2583	-0,1692	0	0	0,4365	0,3358

35	0,4061	0,0054	0,1725	1,239	-0,9343	0	0,4061	0,0054	0,3943	0,4981	0,2041	0,2361	0	0	0,4061	0,0054
36	1,1232	-0,0671	0,2378	0,91	0,6149	0	1,1232	-0,0671	0,5641	0,3962	1,2357	0,2149	0	0	1,1232	-0,0671
37	0,6085	-1,7463	0,6012	0,2672	-0,7409	0	0,6085	-1,7463	0,1835	2,4504	0,9367	-0,0671	0	0	0,6085	-1,7463
38	0,675	0,3604	0,6445	0,6912	-1,2575	0	0,675	0,3604	0,8034	1,3726	0,9319	2,2195	0	0	0,675	0,3604
39	0,6349	-0,6871	0,2373	0,7989	-0,5191	0	0,6349	-0,6871	0,8128	0,4936	0,2977	-0,4745	0	0	0,6349	-0,6871
40	0,7954	0,0182	0,4299	0,2509	-2,7097	0	0,7954	0,0182	0,3761	0,6531	0,8401	1,4451	0	0	0,7954	0,0182
41	1,4347	-1,8325	1,145	0,3857	1,5319	0	1,4347	-1,8325	0,9405	1,3179	1,6248	-0,9603	0	0	1,4347	-1,8325
42	0,7831	-0,131	1,2411	0,6873	0,9918	0	0,7831	-0,131	0,1636	0,5703	0,4493	-2,8969	0	0	0,7831	-0,131
43	0,4358	-1,9674	1,0737	0,5636	-1,3986	0	0,4358	-1,9674	0,2766	0,1822	0,5238	0,3136	0	0	0,4358	-1,9674
44	0,8919	-0,5792	1,311	0,3864	0,2802	0	0,8919	-0,5792	0,7664	0,911	0,6583	0,4515	0	0	0,8919	-0,5792
45	1,1497	-0,1177	0,1925	0,2583	-1,0094	0	1,1497	-0,1177	0,4615	0,5876	0,6576	0,924	0	0	1,1497	-0,1177
46	0,563	-1,0439	0,4231	0,4295	0,8421	0	0,563	-1,0439	0,6142	0,1211	0,7189	-1,5529	0	0	0,563	-1,0439
47	0,6696	-0,8142	0,2041	0,5641	0,3358	0	0,6696	-0,8142	1,578	1,7612	0,5709	-1,6598	0	0	0,6696	-0,8142
48	0,7891	-0,7606	0,3962	1,2357	0,0054	0	0,7891	-0,7606	0,6931	0,8892	0,8561	2,0861	0	0	0,7891	-0,7606

EK 3. IRTPRO 2.1 İLE ÖRNEK KESTİRİM KODLARI

Basit yapılu iki boyutlu 12 maddeli testin parametrelerini kestirmek için kullanılan IRTPRO 2.1 kodu:

```
Project:
  Name = 1;
Data:
  File = .\1.ssig;
Analysis:
  Name = Test1;
  Mode = Calibration;
Title:
12 madde analizleri
Comments:
Estimation:
  Method = BAEM;
  E-Step = 5000, 1e-005;
  SE = S-EM;
  M-Step = 500, 1e-006;
  Quadrature = 49, 6;
  SEM = 0.001;
  SS = 1e-005;
Save:
  PRM
Scoring:
  Pattern = MAP;
  Score Persons;
  Mean = 0;
  SD = 1;
Miscellaneous:
  Decimal = 2;
  Processor = 1;
  Print CTLD, P-Nums, Diagnostic;
  Min Exp = 1;
Groups:
Group :
Dimension = 2;
  Items = VAR1, VAR2, VAR3, VAR4, VAR5, VAR6, VAR7, VAR8, VAR9,
  VAR10,
  VAR11, VAR12;
```

```
Codes(VAR1) = 0(0), 1(1);
Codes(VAR2) = 0(0), 1(1);
Codes(VAR3) = 0(0), 1(1);
Codes(VAR4) = 0(0), 1(1);
Codes(VAR5) = 0(0), 1(1);
Codes(VAR6) = 0(0), 1(1);
Codes(VAR7) = 0(0), 1(1);
Codes(VAR8) = 0(0), 1(1);
Codes(VAR9) = 0(0), 1(1);
Codes(VAR10) = 0(0), 1(1);
Codes(VAR11) = 0(0), 1(1);
Codes(VAR12) = 0(0), 1(1);
Model(VAR1) = 2PL;
Model(VAR2) = 2PL;
Model(VAR3) = 2PL;
Model(VAR4) = 2PL;
Model(VAR5) = 2PL;
Model(VAR6) = 2PL;
Model(VAR7) = 2PL;
Model(VAR8) = 2PL;
Model(VAR9) = 2PL;
Model(VAR10) = 2PL;
Model(VAR11) = 2PL;
Model(VAR12) = 2PL;
Means = 0.0, 0.0;
Covariances = 1.0,
              Free, 1.0;
Constraints:
(VAR1, Slope[1]) = 0.0;
(VAR2, Slope[1]) = 0.0;
(VAR3, Slope[1]) = 0.0;
(VAR4, Slope[1]) = 0.0;
(VAR5, Slope[1]) = 0.0;
(VAR6, Slope[1]) = 0.0;
(VAR7, Slope[0]) = 0.0;
(VAR8, Slope[0]) = 0.0;
(VAR9, Slope[0]) = 0.0;
(VAR10, Slope[0]) = 0.0;
(VAR11, Slope[0]) = 0.0;
(VAR12, Slope[0]) = 0.0;
```

IRTPRO 2.1'i batch modunda çalıştırmak için kod:

```
"C:\Program Files (x86)\IRTPRO21\IRTPRO" -RunList "C:\Users\Derya\Desktop\
dosyalar\2 boyutlu basit\41000 kişi 12 madde \IRTPRO\batch.txt"
```

EK 4. MADDE PARAMETRESİ KESTİRİMLERİNE İLİŞKİN ORTALAMANIN GÜVEN ARALIĞI TESTİ SONUÇLARI

Gerçek madde parametresi ortalamalarının, kestirim sonuçlarının güven aralığına düşüp düşmediği aşağıdaki formül kullanılarak incelenmiştir.

90

$$\bar{X} - t_{\alpha/2} \frac{\sigma_x}{\sqrt{n}} \leq \mu_x \leq \bar{X} + t_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}$$

sd=25, $\alpha=0,95$ olarak alınmıştır.

Gerçek parametre değerlerinin ortalamasının kestirilen parametrelerin güven aralığına düşüp düşmediği görülmüştür. Sonuçlar tablo verilmiştir

Sabitler	Değişkenler		Kestirim sonuçlarına ait alt ve üst değerler				Gerçek parametrelerin ortalaması			
	Yapı	Madde Sayısı	Örneklem Büyüklüğü	a ₁	a ₂	a ₃	d	a ₁	a ₂	a ₃
1 Boyutlu	12	1000	0,64 0,94	-	-	-1,04 0,35	0,79	-	-	-0,70
		2000	0,63 0,93	-	-	-1,05 -0,37				
		4000	0,63 0,93	-	-	-1,04 -0,36				
	48	1000	0,65 0,92	-	-	-0,72 0,07	0,77	-	-	-0,33
		2000	0,64 0,90	-	-	-0,73 0,06				
		4000	0,63 0,90	-	-	-0,72 0,06				
2 Boyutlu Karmaşık Yapılı	12	1000	0,65 1,00	0,57 1,05	-	-0,41 0,51	0,70	0,96	-	0,05
		2000	0,66 0,99	0,55 1,04	-	-0,40 0,52				
		4000	0,67 1,01	0,52 1,02	-	-0,41 0,49				
	48	1000	0,40 0,71	0,44 0,84	-	-0,59 0,36	0,55	0,67	-	-0,10
		2000	0,39 0,70	0,45 0,85	-	-0,58 0,37				
		4000	0,39 0,70	0,45 0,85	-	-0,59 0,35				
2 Boyutlu Basit Yapılı	12	1000	0,60 0,88	0,66 1,00	-	-1,03 -0,34	0,75	0,84	-	-0,70
		2000	0,60 0,88	0,68 0,99	-	-1,02 -0,34				
		4000	0,60 0,88	0,66 0,97	-	-1,03 -0,35				
	48	1000	0,62 0,91	0,66 0,89	-	-0,71 0,05	0,78	0,77	-	-0,33
		2000	0,61 0,90	0,65 0,88	-	-0,70 0,05				
		4000	0,62 0,92	0,65 0,88	-	-0,71 0,04				
3 Boyutlu Karmaşık Yapılı	12	1000	0,49 0,96	0,51 1,33	0,27 1,33	-1,08 0,13	0,72	0,71	0,55	-0,30
		2000	0,63 1,15	0,42 1,39	0,13 1,13	-1,05 0,14				
		4000	0,64 1,15	0,56 0,96	0,16 0,62	-0,89 0,20				
	48	1000	0,46 0,89	0,35 0,82	0,56 1,01	-0,44 0,55	0,69	0,73	0,69	0,10
		2000	0,43 0,86	0,34 0,81	0,57 1,03	-0,41 0,57				
		4000	0,46 0,90	0,35 0,82	0,58 1,03	-0,41 0,57				

3 Boyutlu Basit Yapılı	12	1000	0,60 0,95	0,53 0,67	0,84 1,17	-1,06 -0,37	0,77	0,64	0,98	-0,70
		2000	0,58 0,93	0,52 0,66	0,82 1,17	-1,05 -0,37				
		4000	0,59 0,95	0,54 0,68	0,79 1,13	-1,04 -0,36				
	48	1000	0,65 1,00	0,62 0,83	0,61 0,85	-0,74 0,06	0,82	0,76	0,74	-0,33
		2000	0,65 0,99	0,62 0,84	0,61 0,86	-0,72 0,06				
		4000	0,66 1,00	0,65 0,86	0,60 0,84	-0,73 0,06				

EK 4. ETİK KURUL İZİN MUAFİYET FORMU



Hacettepe Üniversitesi
Eğitim Bilimleri Enstitüsü

Form: 40

Tez Çalışması Etik Kurul İzin Muafiyeti Formu

18 / 11 / 2014

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

Tez Başlığı / Konusu: Çok Boyutlu Madde Tepki Kuramının Farklı Modellerinden Çeşitli Koşullar Altında Kestirilen Parametrelerin İncelenmesi

Yukarıda başlığı/konusu gösterilen tez çalışmam:

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

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

Gereğini saygılarımla arz ederim.


Derya ÇAKICI ESER
(Öğrencinin Adı Soyadı, İmzası)

Öğrenci Bilgileri

Adı Soyadı	Derya ÇAKICI ESER
Öğrenci No	N11140655
Anabilim Dalı	Eğitim Bilimleri
Programı	Eğitimde Ölçme ve Değerlendirme
Statüsü	<input type="checkbox"/> Yüksek Lisans <input checked="" type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr.

Danışman Görüşü ve Onayı

Yapılan bu tez çalışması için etik kurul izni almaya gerek duyulmamıştır.


Prof. Dr. Selahattin GELBAL
(İmza)
(Danışmanın Ünvanı, Adı ve Soyadı)

EK 5. ORJİNALLİK RAPORU

The screenshot displays the iThenticate Professional Plagiarism Prevention web interface. The browser address bar shows the URL: https://app.ithenticate.com/en_us/folder/415430. The page header includes "Welcome Derya ÇAKICI ESER | Logout" and "Help". The main navigation bar contains "Folders", "Settings", and "Account Info".

The interface is divided into several sections:

- Left Sidebar:** Contains "My Folders", "My Folders", "My Documents", and "Trash".
- Top Bar:** Includes "Search" and "Trash" buttons.
- Main Content Area:**
 - My Documents:** Shows a document titled "ÇOK BOYUTLU MADDE TEPKİ KURAMININ FARKLI MODELLERİNDEN ÇEŞİTLİ KOŞULLAR ALTINDA KESTİRİLEN PARAMETRELERİN İNCELENMESİ.docx" with a progress bar at 3%. The document is 1 part, 241,167 words, and is on page 1 of 1.
 - Report:** Shows "Report Author Processed Actions" with a progress bar at 3%. The report is dated January 19, 2015, at 2:02:48 PM EET.
 - Submit a document:** Shows "128,090 Pages remaining" and options to "Upload a File", "Zip File Upload", "Multiple File Upload", and "Cut & Paste". A "View: Recent Uploads" link is also present.
 - Folder Info:** Shows "Folder Info" with "Name: My Documents".

The footer of the page contains the URL: https://app.ithenticate.com/en_us/folder/415430.

ÖZGEÇMİŞ

Kişisel Bilgiler

<i>Adı Soyadı</i>	Derya ÇAKICI ESER
<i>Doğum Yeri</i>	Ankara
<i>Doğum Tarihi</i>	29.08.1985

Eğitim Durumu

<i>Lise</i>	Yavuz Selim Anadolu Lisesi	2003
<i>Lisans</i>	Hacettepe Üniversitesi/Kimya Eğitimi	2009
<i>Yüksek Lisans</i>	Hacettepe Üniversitesi/Eğitimde Ölçme ve Değerlendirme	2011
<i>Yabancı Dil</i>	İngilizce: Okuma (İyi), Yazma (İyi), Konuşma (İyi)	

İş Deneyimi

<i>Stajlar</i>		
<i>Projeler</i>		
<i>Çalıştığı Kurumlar</i>	Kırıkkale Üniversitesi/Eğitim Bilimleri Bölümü	2010-halen

Akademik Çalışmalar

Yayınlar (Ulusal, uluslararası makale, bildiri, poster vb gibi.)

Seminer ve Çalıştaylar

Sertifikalar

İletişim

<i>e-Posta Adresi</i>	deryacakicieser@gmail.com

<i>Jüri Tarihi</i>	
--------------------	--