

**BİLİŞSEL TANI VE ÇOK BOYUTLU MADDE TEPKİ  
KURAMI MODELLERİNİN KARŞILIKLI UYUMLARININ  
İNCELENMESİ**

**RETROFITTING OF COGNITIVE DIAGNOSIS AND  
MULTIDIMENSIONAL ITEM RESPONSE THEORY  
MODELS**

**Levent YAKAR**

Hacettepe Üniversitesi

Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

Doktora Tezi

olarak hazırlanmıştır.

2017

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'n¼,

Levent YAKAR'ın hazırladıđı "Bilişsel Tanı ve Çok Boyutlu Madde Tepki Kuramı Modellerinin Karşılıklı Uyumlarının İncelenmesi" başlıklı bu çalıřma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalı'nda Doktora Tezi** olarak kabul edilmiřtir.

Başkan Prof. Dr. Selahattin GELBAL



¼ye (Danıřman) Prof. Dr. Nuri DOĐAN



¼ye Doç. Dr. Tahsin Ođuz BAŐOKÇU



¼ye Doç. Dr. Őeref TAN



¼ye Doç Dr. Burcu ATAR



#### ONAY

Bu tez Hacettepe ¼niversitesi Lisans¼st¼ Eđitim-¼đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından 03 / 02 / 2017 tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca ..... / ..... / ..... tarihinde kabul edilmiřtir.

Prof. Dr. Ali Ekber ŐAHİN  
Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

## İkinci Tez Danışmanı Onay Bildirimi

### 1- Öğrenci Bilgileri (Student Info)


<b>Adı Soyadı</b> (Name/Lastname)	<b>Levent YAKAR</b>
<b>Anabilim Dalı</b> (Department)	Eğitim Bilimleri
<b>Bilim Dalı</b> (Division)	Eğitimde Ölçme ve Değerlendirme
<b>Statüsü</b> (Status)	<input type="checkbox"/> Y. Lisans (Masters - MA, MS) <input checked="" type="checkbox"/> Doktora (Doctorate - Ph.D.) <input type="checkbox"/> Bütünleşik Doktora (Integrated PhD)

### 2- Tez Bilgileri (Thesis/Dissertation Info)

<b>Tezin Başlığı-1</b> (Title of thesis/ dissertation-1)	Bilişsel Tanı ve Çok Boyutlu Madde Tepki Kuramı Modellerinin Karşılıklı Uyumlarının İncelenmesi
<b>Tezin Başlığı-2</b> (Title of thesis/ dissertation-2)	Retrofitting Of Cognitive Diagnosis And Multidimensional Item Response Theory Models

Hacettepe Üniversitesi  
Eğitim Bilimleri Enstitüsü Müdürlüğüne  
Yukarıda özdeş ve tez bilgileri bulunan öğrencinin ikinci tez danışmanı olarak, tez çalışmasının bilgim dâhilinde tamamlandığını ve sonuçlandırılan çalışmayı onayladığımı belirtirim.

Hacettepe University  
Graduate School of Educational Sciences  
I, as the co-adviser of the student whose identification and thesis/dissertation information provided above, approve his work.

<b>Danışmanın Ünvanı, Adı ve Soyadı</b> (Title, First and Last Name of Co-advisor)	<b>Prof. Dr. Jimmy de la TORRE</b>
<b>Kurumu (Institution)</b>	The University of Hong Kong
<b>İmza (Signature)</b>	

## YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

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

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

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

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

Tezimin/Raporumun 01.09.2017 tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.

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

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

Serbest Seçenek/Yazarın Seçimi: .....

.....

03 /02 /2017


Levent YAKAR

## 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.

  
İmza  
Levent YAKAR

## TEŞEKKÜR

Kendisiyle çalışmaktan büyük zevk aldığım, danışmanım Prof. Dr. Nuri DOĞAN'a,

Bilgeliğiyle yol gösteren ve sıcak davranışlarıyla bizi bölümde aile gibi hissettiren, hocam Prof. Dr. Selahattin GELBAL'a,

Derslerinde öğrettikleriyle, alandaki gelişimime katkı sağlayan, hocam Prof. Dr. Hülya KELECİOĞLU'na,

Doktora zamanının değerli, verimli ve huzurlu geçmesinde etken olan ARKADAŞLARIMA,

Desteklerini esirgemeyen bölümdeki HOCALARIMA VE ÇALIŞANLARA,

Eğitimim süresince yurtiçi ve yurtdışı destekler sunan DEVLETİME,

Ve tabi ki, tüm eğitim hayatı boyunca olduğu gibi beni yine yalnız bırakmayan sevgili AİLEME sonsuz teşekkürlerimi sunmayı borç bilirim.

And special thanks to my good friend, great hospitable and wise co-advisor

Prof. Jimmy de la TORRE.

# BİLİŞSEL TANI VE ÇOK BOYUTLU MADDE TEPKİ KURAMI MODELLERİNİN KARŞILIKLI UYUMLARININ İNCELENMESİ

Levent YAKAR

## ÖZ

Bu araştırmada, ortak yönere sahip olan bilişsel tanı modelleri (BTM) ve çok boyutlu madde kuramı (ÇBMTK) modellerinin birbiri yerine kullanım durumları incelenmiştir. Uyarılma olarak ifade edilen bu işlemde birey parametre kestiriminin gerçek düzeye ne derecede yakın olduğu üzerinde durulmuştur. Bu amaçla ÇBMTK'ya yakın özellikler barındıran çok kategorili bir bilişsel tanı modeli olan tam eklenir model (fully-additive model; fA-M) önerilmiştir.

Araştırma kapsamında BTM'lerden çok düzeyli genelleştirilmiş belirleyici girdi ve kapısı (polytomous generalized deterministic input noisy and gate; pG-DINA; de la Torre, 2011) ve fA-M modelleri ile 2PL ÇBMTK modeli kullanılmıştır. Her modelde madde ayırt edicilik indeksi, test maddelerinin yapılarına göre oranı, test uzunluğu ve yetenekler arası korelasyon değerleri farklılaştırılarak toplamda 54 koşul için telafi edici ve telafi edici olmayan yaklaşımlarda 108 veri seti üretilmiştir. ÇBMTK için iki boyut olarak tasarlanan yapı, BTM'ler için dörder düzeye sahip iki nitelik olarak düzenlenmiştir. ÇBMTK'de sürekli olan yetenek parametresi kesme puanları ile kategorik hale getirilerek BTM'ye benzer hale getirilmiştir. Verilerin her üç model ile analizi sonucu birey parametrelerinin doğru sınıflama oranları elde edilmiştir.

Araştırma iki alt problemde tanımlanmıştır. İlki ÇBMTK verisinin analizini içeren BTM'nin ÇBMTK'ye uyarlanması, diğeri ise BTM verilerinin analizi ile ÇBMTK'nin BTM'ye uyarlanmasıdır. Her iki alt problemde de faktörlerin sonuçlara etkisi benzer olmuştur. Sonuçlarda, kullanılan koşullardan test uzunluğu artışının birey parametre doğruluğuna artışını getiren en kuvvetli faktör olduğu görülmüştür. Test uzunluğunu sırasıyla madde ayırt edicilik ve madde yapısı oranı izlemektedir. Yetenekler arası korelasyonun varlığı ise birey parametre doğruluğuna etkisi pozitif ancak diğer faktörlere göre sınırlı olduğu görülmüştür. BTM'nin uyarlanmasını içeren ilk alt problemde telafi edici ile telafi edici olmayan verilerin analiz sonuçları arasındaki farkın az olduğu, ÇBMTK'nin uyarlanmasını içeren ikinci alt problemde ise bu farkın telafi edici veri analiz sonuçları lehine daha fazla olduğu görülmüştür.

Sonuçlar modeller açısından ele alındığında beklendiği üzere tüm veri türleri, ait oldukları modeller tarafından en yüksek doğrulukta kestirilmiştir. Uyarlama çalışmalarında ise ÇBMTK ve pG-DINA verisi fA-M tarafından, fA-M verisi ise ÇBMTK tarafından verinin ait olduğu modele yakın bir oranda doğru kestirilmiştir. ÇBMTK verisinden bilişsel tanı bilgisi alınmak istendiğinde, yeteneklerin çok kategorili niteliğe dönüştürülebilmesi durumunda fA-M uygulaması doğru bir seçenek olacaktır. Benzer şekilde fA-M verisinden sürekli yetenek puanı elde edilmek istendiğinde, nitelikler boyutlara dönüştürülebiliyorsa ÇBMTK uygulamasının kullanılması önerilmektedir. PG-DINA ve ÇBMTK'nin karşılıklı olarak uyarlamasında ise elde edilen doğru sınıflama oranları diğerlerinden düşüktür. FA-M verisinin pG-DINA ile analizi yine düşük doğrulama oranı ile sonuçlanmıştır.

Çalışma kapsamında önerilen fA-M'ın, ÇBMTK ve kendisi gibi BTM olan pG-DINA verilerini üretilen modele yakın bir doğrulukta kestirmesi modelin başarısını göstermektedir. FA-M'ın diğer çok kategorili BTM'lerden farklı olarak geliştirdiği madde yanıtama olasılığı hesaplamasına uygun yeni modeller önerilebilir. Bu sayede fA-M'ın daha farklı verilerde de başarılı kestirimler yapabilecek ve model- veri uyumu yüksek yeni bir model ailesine öncülük edeceği düşünülmektedir.

**Anahtar sözcükler:** Bilişsel tanı modelleri, çok boyutlu madde tepki kuramı, uyarlama, çok kategorili nitelik, fully additive model (fA-M).

**Danışman:** Prof. Dr. Nuri DOĞAN, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

**İkinci Danışman:** Prof. Dr. Jimmy de la TORRE, Hong Kong Üniversitesi, Eğitim Fakültesi, Öğrenme, Gelişme ve Çeşitlilik Bölümü



# RETROFITTING OF COGNITIVE DIAGNOSIS AND MULTIDIMENSIONAL ITEM RESPONSE THEORY MODELS

Levent YAKAR

## ABSTRACT

In this study, interchangeable use of cognitive diagnosis models (CDM) and multidimensional item response theory (MIRT) models, which have common features, were investigated. This study can be considered as a retrofitting study, which concentrates on the accuracy of estimated person parameters to their true levels. To this end, a polytomous attribute CDM, which is referred to as fully-additive model (fA-M) with features close to MIRT, has been proposed.

In the study, polytomous generalized deterministic input noisy and gate (pG-DINA) and fA-M models from CDMs and 2PL MIRT models were used. By manipulating the item discrimination index, the ratio of item structure of test, the test length, and the correlation between the abilities; 54 conditions formed for compensatory and non-compensatory cases. Therefore, for each model, a total of 108 data sets were generated under compensatory and non-compensatory approaches. Two dimensional MIRT structure was organized as two attributes with four levels for CDM cases. Continuous person parameters in MIRT was categorized using cut-off scores to make it compatible with CDM. Correct classification rates of person parameters by the models were obtained for each dataset.

The study was characterized with two sub-questions. The first was concerned with retrofitting CDM to MIRT, which includes the analysis of MIRT data; and the other was concerned with the retrofitting MIRT to CDM by analysis of CDM data. In both sub-questions, the effect of the conditions to results was similar. In conclusion, it was seen that the increase in test length was the strongest factor that led to an increase in person parameter accuracy. The item discrimination index and the ratio of item structure of test followed the test length, respectively. The presence of a positive correlation between the abilities had limited positive influence to the accuracy of person parameters when compared to other factors. In the first sub-question involving retrofitting CDM, the difference between the results of the analysis of the compensatory and the non-compensatory data was very small.

However, in the second sub-question involving the adaptation of the MIRT, it was seen that this difference was larger and favoring the compensatory data analysis.

When the results are considered in terms of models, as expected, all data types were estimated with highest accuracy by the true models. In the retrofitting cases, accuracy rates were close to true model estimation accuracy levels when MIRT and pG-DINA data analyzed by fA-M and fA-M data analyzed by MIRT. To obtain cognitive diagnostic information from the MIRT data, fA-M may be a convenient tool if abilities can be transformed into polytomous attributes. Likewise, if continuous ability score is desired from fA-M data, MIRT application is recommended if attributes can be converted into dimensions. In addition to these results, lower correct classification rates were obtained when pG-DINA data were retrofitted to MIRT, and vice versa. Furthermore, analysis of fA-M data with pG-DINA resulted in lower accuracy rate.

High classification accuracy rates of fA-M when MIRT and pG-DINA data fitted may be considered as success indicators of the proposed model. Future studies may propose new models using the success probability computation developed for fA-M. In this respect, fA-M may lead to a new model family yielding successful estimations for various datasets with high model-data fit.

**Keywords:** Cognitive diagnosis models, multidimensional item response theory, retrofitting, polytomous attribute, fully additive model (fA-M).

**Advisor:** Prof. Dr. Nuri DOĞAN, Hacettepe University, Department of Educational Science, Division of Educational Measurement and Assessment

**Co-Advisor:** Prof. Dr. Jimmy de la TORRE, The University of Hong Kong, Faculty of Education, Division of Learning, Development and Diversity

# İÇİNDEKİLER

KABUL ve ONAY.....	ii
ETİK BEYANNAMESİ.....	v
TEŞEKKÜR.....	vi
ÖZ.....	vii
ABSTRACT.....	ix
İÇİNDEKİLER.....	xi
ÇİZELGELER DİZİNİ.....	xiii
ŞEKİLLER DİZİNİ.....	xiv
SİMGELER VE KISALTMALAR DİZİNİ.....	xv
1. GİRİŞ.....	1
1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı ve Önemi:.....	4
1.3. Problem Cümlesi:.....	5
1.3.1. Alt Problemler:.....	5
1.4. Sınırlılıklar:.....	6
1.5. Araştırmanın Kuramsal Temeli.....	6
1.5.1. Madde Tepki Kuramı (MTK).....	6
Varsayımlar.....	7
Modeller.....	7
Model parametreleri.....	8
1.5.2. Çok Boyutlu Madde Tepki Kuramı (ÇBMTK).....	9
Model Parametreleri.....	12
Telafi edici Modeller.....	13
Telafi edici Olmayan Modeller.....	14
1.5.3. Bilişsel Tanı Modelleri.....	17
Bilişsel Tanı Modellerinin Temelleri.....	18
Q Matrisi.....	21
Başlıca Bilişsel Tanı Modelleri.....	24
G-DINA (Generalized Deterministic-Input, Noisy-And Gate) Model.....	25
DINA (Deterministic-Input, Noisy-And Gate) Model.....	26
DINO (Deterministic-Input, Noisy-Or Gate) Model.....	27
A-CDM (Additive Cognitive Diagnosis Model).....	28
Polytomous G-DINA (pG-DINA).....	29
fA-M (Fully Additive Models).....	31
2. İLGİLİ ARAŞTIRMALAR.....	35
2.1. ÇBMTK ile İlgili Çalışmalar.....	35
2.2. BTM ile İlgili Çalışmalar.....	37
2.3. ÇBMTK ve BTM Uyarlaması ile İlgili Çalışmalar.....	39
3. YÖNTEM.....	45
3.1. Araştırmanın Türü.....	45
3.2. Verilerin Elde Edilmesi.....	45

3.3. Verilerin İşlenmesi ve Çözümlemesi .....	52
4. BULGULAR VE TARTIŞMA .....	56
4.1. ÇBMTK Verisi Analizlerine İlişkin Bulgular ve Tartışma .....	56
4.1.1. Madde Ayırt Edicilik Koşullarına Göre Doğru Sınıflama Oranları ....	56
4.1.2. Madde Yapısı Oranı Koşullarına Göre Doğru Sınıflama Oranları ...	60
4.1.3. Test Uzunluğu Koşullarına Göre Doğru Sınıflama Oranları .....	64
4.1.4. Yetenekler Arası Korelasyon Koşullarına Göre Doğru Sınıflama Oranları.....	67
4.1.5. Yetenek Düzeylerinde Doğru Sınıflama Oranları .....	71
4.2. BTM Verileri Analizlerine İlişkin Bulgular ve Tartışma.....	72
4.2.1. Madde Ayırt Edicilik Koşullarına Göre Doğru Sınıflama Oranları ....	72
4.2.2. Madde Yapısı Oranı Koşullarına Göre Doğru Sınıflama Oranları ...	81
4.2.3. Test Uzunluğu Koşullarına Göre Doğru Sınıflama Oranları .....	88
4.2.4. Yetenekler Arası Korelasyon Koşullarına Göre Doğru Sınıflama Oranları.....	96
4.2.5. Yetenek Düzeylerinde Doğru Sınıflama Oranları .....	102
5. SONUÇ ve ÖNERİLER .....	106
5.1. Sonuçlar.....	106
5.2. Öneriler.....	110
5.2.1. Araştırmaya Dönük Öneriler .....	110
5.2.2. Uygulamaya Dönük Öneriler.....	111
KAYNAKÇA.....	112
EKLER DİZİNİ .....	117
EK 1. ETİK KOMİSYONU ONAY BİLDİRİMİ .....	118
EK 2. ORJİNALLİK RAPORU.....	119
ÖZGEÇMİŞ .....	120

## ÇİZELGELER DİZİNİ

Tablo 1.1: Örnek Q Matrisi .....	21
Tablo 1.2: Orijinal, Düşmüş ve Çökmüş Örtük Gruplar.....	30
Tablo 3.3: Örnek Örtük Sınıflar İçin ÇBMTK Grup Büyüklükleri ve Doğru Yanıtlama Olasılıkları .....	47
Tablo 3.4: pG-DINA İçin Örnek Grupların Doğru Yanıtlama Olasılıkları .....	48
Tablo 3.5: fA-M Örnek Madde Parametreleri.....	50
Tablo 3.6: Örnek Örtük Sınıflar İçin ÇBMTK Grup Büyüklükleri ve Doğru Yanıtlama Olasılıkları .....	50
Tablo 4.7: Madde Ayırt Ediciliğe Göre ÇBMTK Verisi DSO'ları .....	56
Tablo 4.8: Madde Yapısı Oranına Göre ÇBMTK Verisi DSO'ları.....	60
Tablo 4.9: Test Uzunluğuna Göre ÇBMTK Verisi DSO'ları .....	64
Tablo 4.10: Yetenekler Arası Korelasyona Göre ÇBMTK Verisi DSO'ları .....	68
Tablo 4.11: Yetenek Düzeylerinde ÇBMTK Verisi DSO'ları .....	71
Tablo 4.12: Madde Ayırt Ediciliğe Göre pG-DINA Verisi DSO'ları .....	73
Tablo 4.13: Madde Ayırt Ediciliğe Göre fA-M Verileri DSO'ları.....	77
Tablo 4.14: Madde Yapısı Oranına Göre pG-DINA Verisi DSO'ları.....	81
Tablo 4.15: Madde Yapısı Oranına Göre fA-M Verileri DSO'ları .....	85
Tablo 4.16: Test Uzunluğuna Göre pG-DINA Verisi DSO'ları.....	89
Tablo 4.17: Test Uzunluğuna Göre fA-M Verileri DSO'ları .....	93
Tablo 4.18: Yetenekler Arası Korelasyona Göre pG-DINA Verisi DSO'lar .....	97
Tablo 4.19: Yetenekler Arası Korelasyona Göre fA-M Verileri DSO'ları .....	100
Tablo 4.20: Yetenek Düzeylerinde pG-DINA DSO'ları .....	102
Tablo 4.21: Yetenek Düzeylerinde fA-M Verileri DSO'ları .....	104

## ŞEKİLLER DİZİNİ

Şekil 1.1. Madde Karakteristik Eğrisi .....	9
Şekil 1.2. Basit ve Karmaşık Yapıdaki Maddeler .....	11
Şekil 1.3. Telafi Edici ve Telafi Edici Olmayan Madde Karakteristik Eğrileri .....	15
Şekil 1.4. Telafi edici ve Telafi edici olmayan Madde Tepki Yüzeyleri .....	16
Şekil 1.5. Değerlendirme Üçgeni .....	19
Şekil 1.6. Farklı Hiyerarşi Yapıları .....	22
Şekil 1.7. Modeller İçin Örnek Olasılıklar .....	28
Şekil 1.8. Çok Kategorili Nitelik İçin Model Örnek Olasılıkları .....	33
Şekil 3.9. Araştırma İşlem Basamakları .....	52
Şekil 4.10. Düşük Düzey Madde Ayırt Edicilik Koşulu ÇBMTK Verisi DSO'ları ....	57
Şekil 4.11. Orta Düzey Madde Ayırt Edicilik Koşulu ÇBMTK Verisi DSO'ları .....	58
Şekil 4.12. Yüksek Düzey Madde Ayırt Edicilik Koşulu ÇBMTK Verisi DSO'ları...	59
Şekil 4.13. Çoğunlukla Karmaşık Yapı Koşulu ÇBMTK Verisi DSO'ları.....	61
Şekil 4.14. Eşit Orandaki Yapı Koşulu ÇBMTK Verisi DSO'ları .....	62
Şekil 4.15. Çoğunlukla Basit Yapı Koşulu ÇBMTK Verisi DSO'ları.....	63
Şekil 4.16. 15 Maddelik Test Uzunluğu Koşulu ÇBMTK Verisi DSO'ları.....	65
Şekil 4.17. 30 Maddelik Test Uzunluğu Koşulu ÇBMTK Verisi DSO'ları.....	66
Şekil 4.18. 60 Maddelik Test Uzunluğu Koşulu ÇBMTK Verisi DSO'ları.....	67
Şekil 4.19. Yetenekler Arası 0 Korelasyon Koşulu ÇBMTK Verisi DSO'ları.....	69
Şekil 4.20. Yetenekler Arası 0,6 Korelasyon Koşulu ÇBMTK Verisi DSO'ları.....	70
Şekil 4.21. Yetenek Düzeylerine Göre ÇBMTK Verisi DSO'ları .....	71
Şekil 4.22. Düşük Düzey Madde Ayırt Edicilik Koşulu pG-DINA Verisi DSO'ları ..	74
Şekil 4.23. Orta Düzey Madde Ayırt Edicilik Koşulu pG-DINA Verisi DSO'ları ....	75
Şekil 4.24. Yüksek Düzey Madde Ayırt Edicilik Koşulu pG-DINA Verisi DSO'ları.	76
Şekil 4.25. Düşük Düzey Madde Ayırt Edicilik Koşulu fA-M Verisi DSO'ları .....	78
Şekil 4.26. Orta Düzey Madde Ayırt Edicilik Koşulu fA-M Verisi DSO'ları .....	79
Şekil 4.27. Yüksek Düzey Madde Ayırt Edicilik Koşulu fA-M Verisi DSO'ları.....	80
Şekil 4.28. Çoğunlukla Karmaşık Yapı Koşulu pG-DINA Verisi DSO'ları.....	82
Şekil 4.29. Eşit Oran Yapı Koşulu pG-DINA Verisi DSO'ları.....	83
Şekil 4.30. Çoğunlukla Basit Yapı Koşulu pG-DINA Verisi DSO'ları.....	84
Şekil 4.31. Çoğunlukla Karmaşık Yapı Koşulu fA-M Verisi DSO'ları.....	86
Şekil 4.32. Eşit Oran Yapı Koşulu fA-M Verisi DSO'ları.....	87
Şekil 4.33. Çoğunlukla Basit Yapı Koşulu fA-M Verisi DSO'ları.....	88
Şekil 4.34. 15 Maddelik Test Uzunluğu Koşulu pG-DINA Verisi DSO'ları.....	90
Şekil 4.35. 30 Maddelik Test Uzunluğu Koşulu pG-DINA Verisi DSO'ları.....	91
Şekil 4.36. 60 Maddelik Test Uzunluğu Koşulu pG-DINA Verisi DSO'ları.....	92
Şekil 4.37. 15 Maddelik Test Uzunluğu Koşulu fA-M Verisi DSO'ları .....	94
Şekil 4.38. 30 Maddelik Test Uzunluğu Koşulu fA-M Verisi DSO'ları .....	95
Şekil 4.39. 60 Maddelik Test Uzunluğu Koşulu fA-M Verisi DSO'ları .....	96
Şekil 4.40. Yetenekler Arası 0 Korelasyon Koşulu pG-DINA Verisi DSO'ları.....	98
Şekil 4.41. Yetenekler Arası 0,6 Korelasyon Koşulu pG-DINA Verisi DSO'ları.....	99
Şekil 4.42. Yetenekler Arası 0 Korelasyon Koşulu fA-M Verisi DSO'ları.....	101
Şekil 4.43. Yetenekler Arası 0,6 Korelasyon Koşulu fA-M Verisi DSO'ları.....	102
Şekil 4.44. Yetenek Düzeylerinde pG-DINA Verisi DSO'ları .....	103
Şekil 4.45. Yetenek Düzeylerinde fA-M Verileri DSO'ları.....	104

## SİMGELER VE KISALTMALAR DİZİNİ

**ÇBMTK:** Çok Boyutlu Madde Tepki Kuramı

**BTM:** Bilişsel Tanı Modeli

**PG-DINA:** Polytomous Generalized Deterministic Input Noisy And gate

**FA-M:** Fully Additive Model

## 1. GİRİŞ

Bu bölümde problem durumu, Çok Boyutlu Madde Tepki Kuramı (ÇBMTK) ve Bilişsel Tanı Modellerine (BTM) değinilmiştir.

### 1.1. Problem Durumu

Eğitim toplum olarak daha yüksek hayat standardına sahip olmak adına kaliteli bireyler yetiştirme faaliyetlerini gerektirir. Eğitimin kalite düzeyini belirlemek için ölçme ve değerlendirmeye başvurulur. Bu sayede yapılan işin istenilen düzeyde gerçekleştirilip gerçekleştirilmediği öğrenilebilir. Ölçme ve değerlendirme ile öğretim programının veya bireylerin hedef alandaki yeterlik düzeyleri gözlemlenebilir.

Ölçme ve değerlendirme işlemi amacın belirlenmesi ile başlar. Bu amaca uygun olarak ölçme aracının seçimi ve uygulanması gerekmektedir. Ölçme işleminin kaliteli ve anlamlı olması geçerli sonuçlara bağlıdır. Ölçmenin geçerliğini sağlamak için ölçülmek istenilen özelliğin/özelliklerin tam olarak saptanması ve bu amaca hizmet edecek testlerin hazırlanması gerekmektedir. Bu gerekliliklerin sağlanmaması, en güvenilir sonuçları dahi geçersiz kılacaktır. Buradaki kritik nokta neyi, hangi düzeyde ölçüyoruz sorusuna verilecek cevaptır. Birbirine yakın olan beş beceri, kullanılan ölçme modeline göre tek puanla temsil edilip değerlendirilebileceği gibi, her biri için farklı puanlar elde edilip bireyler bu puanlarla değerlendirilebilir.

Ölçülen bir özelliğin birden fazla puanla gösterilmesi kadar ölçülen birden fazla özelliğin bir puanla temsil edilmesi hatalı sonuçlara neden olabilir. Testte yoklanan birden fazla özelliğin tek puanla temsili halinde bu puanın hangi özelliğin ne kadarının temsil ettiğini bilmek mümkün olmayacaktır. Bu durum testin geçerliği üzerinde ciddi şüpheler oluşturacaktır. Bu türden hatalar yanlışlıkla yapılabileceği gibi tek puanın yorum ve kullanım kolaylığı nedeniyle ihmale bağlı olarak da gerçekleşebilir.

Sadece bir özellik ölçtüğü düşünülen testlerin dahi gerçekte birden fazla özelliği ölçtüğü; ancak ölçülen özellikler arasındaki yüksek korelasyon nedeniyle testler diğer özelliklerin ölçülmesine duyarsız olabilir (Reckase, 2007). Bu durumda testin sadece başat faktör olarak yer alan özelliği ölçtüğü söylenebilir. Ancak özellikle psikolojik temelli ölçümlerde ölçülen özellikler arası korelasyon ana faktörün



diğerlerini baskılamaya yetecek kadar çok yüksek olmayabilir. Bir fizik sorusunda matematik işlem becerisi istem dışı bir şekilde yoklanabilir, aynı şekilde sözel ifadelerle anlatılmış bir matematik sorusu okuduğunu anlama becerisinden arındırılmaz.

Eğitimde ölçme çoğunlukla testler aracılığıyla yapılmaktadır. Testlerin geliştirilmesinde kullanım amacı kadar teste ilişkin altyapı da önemli bir yere sahiptir. Bu altyapı ölçme kuramları olarak yerlerini almıştır. Kuramlardan biri Klasik Test Teorisi (KTK)'dir. Yirminci yüzyılın başında temelleri atılan kuram hala yaygın olarak kullanılmaktadır. Yüzyılın ortalarından sonra Madde Tepki Kuramı (MTK) görünmeye başlamıştır (Embretson ve Reise, 2000).

Gerçek test maddelerinin karmaşık yapısı ve bireyin maddeyi çözmeye kullandığı çoklu yetenek ve bilgi varlığı çok boyutlu modelin ihtiyacını oluşturmaktadır. Bu ihtiyacın görmezden gelinerek çok boyutlu verilere tek boyutlu MTK uygulanması sorunlu sonuçlara neden olabilir. Ayrıca ortak metne sahip maddeler gibi madde grubu içindeki maddeler için verilen yanıtlar yüksek ihtimalle birbirine bağımlı olacaktır. Sonrasında bu maddeler MTK'nin yerel bağımsızlığı göz ardı edilip tek boyutlu uygulamaya tabi tutulursa madde parametreleri yanlı ve test güvenilirliği olduğundan yüksek tahmin edilecektir (Wang & Wilson, 2005). Bu sebeple öncelikle maddelerin doğru yanıtlanması için gerekli yetenek veya yeteneklerin sayısının tespit edilmesi şarttır. Sonrasında uygulanması gereken model karar verilmesi sonuçlar açısından isabetli karar olacaktır.

Ölçülen özelliğin tek boyutlu olmasını gerektiren MTK'de bu sınırlılığın giderilmesine yönelik olarak çok boyutlu testler için çok boyutlu madde tepki kuramı (ÇBMTK) geliştirilmiştir. Böylece kimi zaman ihlallere neden olan sınırlılık giderilmiştir. ÇBMTK ile her bir yeteneğe ilişkin yetenek parametreleri elde edilebilmektedir (Reckease, 2009). ÇBMTK uygulamaları çok sayıda özelliğin ölçülebilmesi nedeniyle fark yaratmış ve giderek yaygın hale gelmiştir. MTK gibi değer biçmeye yönelik değerlendirme yaklaşımına sahiptir.

MTK, ÇBMTK'den başka modellere de öncülük etmiştir. Bu modellerden biri de bilişsel tanı modelleridir (BTM). BTM'nin ana amacı bireylerin öğrenme eksikliklerini tespit etmek adına testte yoklanan nitelikler için bireyin niteliklere sahip olup olmadığını göstermektir. Bu amaçla daha çok eğitim süreci içerisinde yapılan ve

bireyleri geliřtirmeyi ve yetiřtirmeyi amalayan biimlendirici deęerlendirme zellięi tařıtmaktadır. Bu gstergelerin tek tek incelenebilmesi BTM'lerin avantajı olarak grlebilir. BTM, MTK'ye nazaran daha kk birimleri ayrı ayrı lmesi sayesinde son yıllarda giderek yaygınlařmaktadır.

llen zellięin ifade edilmesi modeller arasında grlen farklılıklardan birisidir. KTK, MTK ve BMTK bireylerin llen zelliklerini srekli yapıdaki parametre ile ifade ederken yaygın olarak kullanılan BTM zellikleri bireylerde grlp grlmemelerine gre kesikli olarak sınıflandırılmaktadır. Elde edilen yetenek parametrelerine bakıldığında BTM'nin 0-1 ikili sınıflama trnde veri retmesi BMTK'ya gre daha alt dzey lek tr olması nedeniyle dezavantaj olarak grlebilir. retilen sonuların farklı olması bir yandan da yeteneęin srekli mi kesikli mi olduęu tartiřmalarına dayandırılabilir (Wang, 2009).

llen birimin byklę bir bařka tartiřma konusudur. MTK ve BMTK uygulamalarında llen zellik geniř sınırlara sahiptir ve genellikle boyut, yetenek, faktr gibi isimlerle anılır. BTM'lerin lme birimleri nitelik, zellik, grev gibi isimlerle anılmaktadır ve boyut/yetenek/faktre gre ok daha kk sınırlara sahiptir. BTM'lerde llen zellięin kklę elde edilen sonuların detaylı olmasını saęlamaktadır. MTK tarafından aritmetik iřlem yeteneęini lmeyi amalayan bir boyutlu test BTM'ler iin toplama ıkarma, arpma ve blme niteliklerini lmeyi amalayan 4 nitelikli bir test olarak ele alınabilir.

BTM ve BMTK'nin en nemli ortak ynleri ok boyutlu leklerde uygulanabilmeleridir. Bunun yanında her iki kuram da telafi edici ve telafi edici olmayan modeller barındırmaktadır. Bu benzerlięin dayanak noktası yine birden fazla zellik veya yetenekten hareketle sonuca ulařmaları olarak grlebilir. Modellerde kullanılacak madde yapıları da benzemektedir. Her iki kuram iin basit yapı ve karmařık yapıdaki maddeler bulunmaktadır. Bu zellik kuramlara ait modellerin analizlerinde nemli bir konuma sahiptir.

BTM ve BMTK'nin farklılıkları modellerin birbiri yerine uygulanmaları durumunda ortaya ıkabilecek olumsuzlukları meydana getirmektedir. Model farklılıklarına gz atıldığında, modellerin kullanım amaları, madde parametreleri, birey parametrelerinin srekli veya kesikli olma durumları ve lme birimlerinin byklkleri n plana ıkmaktadır.

Psikometrik modellerin birbirine benzerlikleri hangi modelin ne zaman kullanılacağına ilişkin soru işaretlerine yol açmaktadır. Bu benzerlikler aynı zamanda testin uygulandığı modelin belli olması durumunda ise teste farklı bir bakış açısı kazandırmak adına uygulanabilecek yakın model seçeneği sunmaktadır. Benzer altyapıya sahip modellerden elde edilen sonuçların yorumlanması, model çıktılarının karşılaştırması yapıldıktan sonra daha anlamlı hale gelir. Bu yüzden gittikçe yaygınlaşan bu iki psikometrik modele ait çıktılar karşılıklı olarak değerlendirilmelidir.

Uyarılama olarak ifade edilebilecek (retrofit) işlem sayesinde veriden, mevcut bilgiler dışında başka bilgiler edinilmesi amaçlanmaktadır. Bir modele aitliği bilinen veri setinin başka modele uygulaması durumunda elde edilecek yeni bilgilerle mevcut çıktıların çeşitlendirilmesi, desteklenmesi ve zenginleştirilmesi uyarılama işleminin temelini oluşturmaktadır. Uyarılama işleminde mevcut model ile uyarılama ile katkıda bulunacak modelin sonuçlar açısından birbirini destekler durumda olması işlem sonucunu değerli kılacaktır.

Birbirinden uzak modelleri birbiri yerine kullanarak uyumu yüksek sonuçlar beklemek gerçeklikten uzaktır. Daha iyi kıyaslanabilir sonuçlar için farklı alandaki modellerin yakın olanlarının uygulamaya sokulması doğru bir yaklaşımdır. Bu amaçla ÇBMTK modellerine yakın BTM modelleri kullanılarak geniş perspektifte uyarılama, uygulama ve değerlendirme yoluna gidilmiştir. Amacın gerçekleşmesi için yeni bir BTM önerilerek alana katkıda bulunulmuştur.

## **1.2. Araştırmanın Amacı ve Önemi:**

Bu araştırmanın amacı BTM ile ÇBMTK'yi karşılıklı uygulayarak modellerin birbirlerine uyarlanabilme düzeylerini incelemektir. ÇBMTK'ya yakın bir BTM ortaya koymak da bir başka amaç olarak değerlendirilebilir. Bu amacı gerçekleştirme çabası ile ortaya konan yeni BTM ailesi hem amacı zenginleştirmiş hem de çalışma sonunda ürün olarak ortaya çıkmıştır.

MTK'den doğan BTM modelleri ve ÇBMTK'nin arasındaki ilişkiyi görebilmek için parametreleri karşılaştırmak gerekmektedir. Böylece birbirlerinin yerine kullanılma imkânları görülecektir. Birbirlerinin yerine kullanılması durumunda ortaya çıkacak sonuçlar irdelenmiş olacaktır.

ÇBMTK ile BTM arasındaki temel benzerlikler her ikisinin de MTK'ya dayalı olması ve birden fazla özellik içeren durumlarda kullanılabilmesidir. Bu benzerlikler nedeniyle modeller karşılaştırmaya değerdir. Ayrıca modellerin birbiri yerine kullanılması modeller arasındaki farklılıklara bağlı olarak ortaya çıkabilecek olumsuz sonuçları görmek açısından önem arz etmektedir. Özellikle kullanılan çok sayıda koşul birçok olası durumda karşılaşılabilecek sorun ve sonuçlara ışık tutmaktadır.

Araştırmanın benzerleri literatürde olmasına rağmen birden fazla modeli karşılaştıran bir çalışma olmaması ve ilgili modeller için ilk çalışma olması da araştırmanın önemini artırmaktadır. Önerilen model benzer formülasyon kullanarak önerilebilecek birçok modelden biridir ve alanda yeni bir bakış açısı olarak değerlendirilebilir.

### **1.3. Problem Cümlesi:**

Aynı verinin Bilişsel Tanı Modelleri (BTM) ve Çok Boyutlu Madde Tepki Kuramına (ÇBMTK) uygulanması ile elde edilen birey parametrelerinin doğrulanma düzeyi nasıldır?

#### **1.3.1. Alt Problemler:**

1- ÇBMTK verisinin ÇBMTK ve BTM ile analizinde;

- a) madde ayırt ediciliğe göre elde edilen verilerin kestiriminde, modellerin birey parametrelerini doğru sınıflama oranı nasıldır?
- b) madde yapısına göre elde edilen verilerin kestiriminde, modellerin birey parametrelerini doğru sınıflama oranı nasıldır?
- c) test uzunluğuna göre elde edilen verilerin kestiriminde, modellerin birey parametrelerini doğru sınıflama oranı nasıldır?
- d) yetenekler arası korelasyona göre elde edilen verilerin kestiriminde, modellerin birey parametrelerini doğru sınıflama oranı nasıldır?
- e) düzeylerine göre yeteneğin doğru sınıflama oranları nasıldır?

2- BTM verilerinin ÇBMTK ve BTM ile analizinde;

- a) madde ayırt ediciliğe göre elde edilen verilerin kestiriminde, modellerin birey parametrelerinin doğru sınıflama oranı nasıldır?

- b) madde yapısına göre elde verilerin kestiriminde, modellerin birey parametrelerini doğru sınıflama oranı nasıldır?
- c) test uzunluğuna göre elde edilen verilerin kestiriminde, modellerin birey parametrelerini doğru sınıflama oranı nasıldır?
- d) yetenekler arası korelasyona göre elde edilen verilerin kestiriminde, modellerin birey parametrelerini doğru sınıflama oranı nasıldır?
- e) düzeylerine göre yeteneğin doğru sınıflanma oranları nasıldır?

#### **1.4. Sınırlılıklar:**

Araştırma pG-DINA ve fA-M BTM'leri ile sınırlıdır.

Telafi edici ve telafi edici olmayan 2PL ve 3PL ÇBMTK modelleri ile sınırlıdır.

Farklılaştırılan faktörler olan madde ayırt edicilik indeksi, yapılarına göre maddelerin testteki oranı, test uzunluğu ve yetenekler arası korelasyona uygun üretilen veri ile sınırlıdır.

Kestirimler sadece expected a posteriori (EAP) metodu ile gerçekleştirilmiştir.

#### **1.5. Araştırmanın Kuramsal Temeli**

Bu bölümde tezin teorik iskeletini oluşturan MTK, ÇBMTK ve BTM üzerinde durulacaktır. Analizde kullanılacak ÇBMTK ve BTM'ler üzerinde detaylı açıklamalar yapılacaktır.

##### **1.5.1. Madde Tepki Kuramı (MTK)**

Yirminci yüzyılın başlarında ortaya çıkan KTK psikometri alanının ilk kuramıdır. Uygulama kolaylığı sayesinde günümüzde halen yaygın olarak kullanılmaktadır. MTK'nin gün yüzüne çıkışı en fazla yirminci yüzyılın ortalarına kadar dayandırılabilir. Örtük özellik teorisi ve modern test teorisi (Crocker & Algina, 1986) olarak da adlandırılan teori ile KTK arasındaki temel farkların isimlerinde ortaya çıktığı görülmektedir. Her iki teoride de madde parametreleri bulunmasına rağmen KTK testin bütünü üzerine yoğunlaşırken, MTK testin bütünü değil maddeleri üzerinde işlem yapmayı temele alır. İsimden anlaşılabilir diğer fark ise MTK'nin modern (yeni) KTK'nin klasik (eski) olarak isimlendirilmesi çıkış zamanlarındaki farkı ortaya koymaktadır.

MTK temelleri Lord ve Novick'in 1968'de yayınlanan kitabına Birnbaum tarafından yazılan bölüme dayandırılabilir (Embretson ve Reise, 2000). Bu tarihten sonra sağlam matematiksel temellere dayanan MTK'de önemli gelişmeler görülmeye ve kuramın temelini oluşturan modeller ortaya sürülmeye başlanmıştır. MTK'de soruyu doğru yanıtlayabilmek için yoklanan yeteneğin bulunması gerekir. Sorunun doğru yanıtlanması sorunun altında yatan gizli yeteneğin bireyde göstergesi olarak kabul edilir. Diğer bir deyişle gözlenen madde performanslarından örtük yetenek parametresine ulaşmaya çalışılmaktadır. Yeteneğe ilişkin soruların zorluk seviyelerine göre doğru ya da yanlış olarak yanıtlanması arka plandaki yeteneğin doğru şekilde ölçülmesiyle sonuçlanır. MTK'nin genel özelliklerini sıralayacak olursak, madde parametrelerinin gruptan bağımsız olması, birey parametrelerinin ilgili popülasyon için kalibre edilmiş maddelerden bağımsız olması ve yetenek tahminin kesinliğinin bilinmesidir (Hambleton & Swaminathan, 1985).

### **Varsayımlar**

MTK için iki temel varsayımdan bahsedilebilir, madde karakteristik eğrisinin belli bir şeklinin olması ve yerel bağımsızlığın sağlanması. Madde karakteristik eğrisi yetenek düzeyi değişiminde maddenin doğru yanıtlanma olasılığını gösteren eğridir. Bireyin yeteneği arttıkça maddeyi doğru yanıtlanma olasılığında artış olmasını gerektirir. Bu ilişki monotonik artışı tarif eder. Uç yeteneklerde olasılık artışı küçük iken özellikle ayırt ediciliğin yüksek olduğu orta düzeylerde daha belirgindir. Yerel bağımsızlık madde ve yetenek parametrelerinin kontrollü halinde bir maddeyi doğru yanıtlanma olasılığının diğer maddelerden bağımsız olması durumudur. Kimi kaynaklarda geçen tek boyutluluk varsayımı tek boyutlu MTK için geçerlidir. Bu modellerde tek boyutluluk yerel bağımsızlık için kanıt teşkil eder (Embretson ve Reise, 2000).

### **Modeller**

MTK modelleri genellikle verinin 2 kategorili (0-1) veya çok kategorili olmasına göre sınıflandırılmaktadır. 2 kategorili modellerde çok kullanılan 1, 2 ve 3 parametrelili lojistik (PL) modellere değinilmiştir.

*1 PL Model:* En basit MTK modelidir. Madde parametrelerinden sadece zorluk parametresi maddeler için değişkendir. Bu da modelin  $b$  parametresine bağlı olduğunun gösterir. Madde ayırt edicilik ( $a$ ) parametresi tüm maddeler için 1 olarak

ele alındığında Rasch model olarak adlandırılır (de Ayala, 2009).  $D=1,7$  ölçekleme sabiti kullanılarak ogive modele eşdeğer sonuç üreten modelin madde doğru yanıt olma olasılığı şu şekilde hesaplanır;

$$p(x_j = 1|\theta, a, b) = \frac{1}{1 + e^{-Da(\theta-b)}}$$

*2 PL Model:* Yukarı verilen madde yanıt olma olasılığı 2 PL için de geçerlidir. 1 PL model ile arasındaki fark  $a$  parametresinin maddeler için farklı değer alabilmesidir. Böylece  $a$  üzerindeki kısıtlama kaldırılarak formül maddeler arası farkı gösterir hale gelmiştir.

*3 PL Model:* Modelin 2 PL'den farkı şans parametresinin etkin hale gelmesidir. 2 PL modelde 0 olarak değerlendirilen şans parametresi, 3 PL modelin formülünde yerini almaktadır.

$$p(x_j = 1|\theta, a, b) = c + \frac{1 - c}{1 + e^{-Da(\theta-b)}}$$

### **Model parametreleri**

*Ayırt edicilik parametresi (a):* Maddenin yoklanan yeteneğe sahip olan ve olmayan bireyleri ne derecede ayırt edebildiğini gösterir. Madde karakteristik eğrisinin şeklini yatay/dikey olması yönünden belirler. Daha dik madde karakteristik eğrisine sahip maddenin daha ayırt edici olduğu anlamını verir. Genellikle 0-6 arasında ölçeklenir. Sıkça kullanılan lojistik modellerde 1 orta düzey ayırt edicilik olarak kabul edilebilir. Tüm modellerde bulunur ancak tek parametrelilikte tüm maddeler için sabittir.

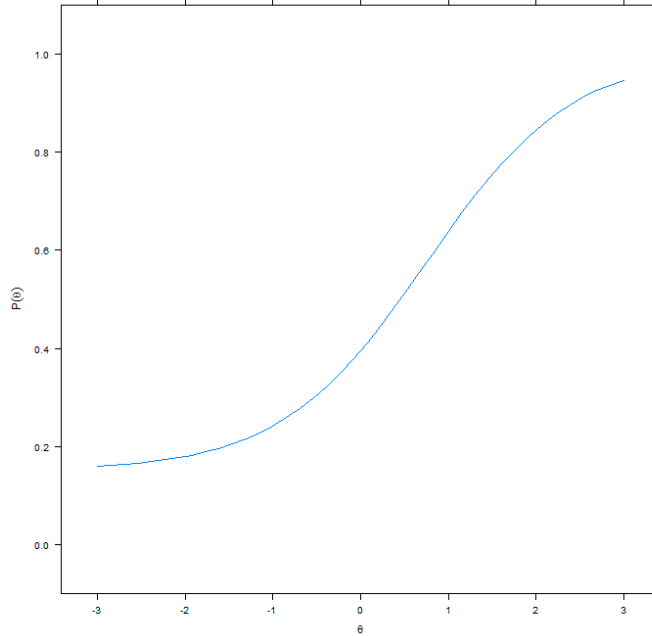
*Zorluk parametresi (b):* Madde karakteristik eğrisinin konumunu belirler. Güçlük veya konum parametresi olarak da isimlendirilir. Yatay eksene göre daha sağda olan maddenin daha zor olduğunu gösterir. Bireyin yetenek ve maddenin zorluk parametresinin aynı ölçekte olması yorumlamayı kolaylaştırır ve bireyin maddeye hangi olasılıkla doğru yanıt vereceğini de ortaya koyar. Şans parametresinin olmadığı modellerde maddenin doğru yanıtlanma olasılığının %50 olduğu yetenek seviyesini işaret eder. Genellikle -3 ile +3 arasında ölçeklenir. Tüm modellerde bulunur.

*Şans parametresi (c):* En düşük yetenek seviyesindeki bireyin maddeyi doğru yanıt olma olasılığını verir. Çoktan seçmeli testlerde bilgisi olmayan bireyin şansla doğru seçeneği bulması örnek olarak verilebilir. Madde karakteristik eğrisinin y

ekseni üzerinde başlayacağı noktayı belirler. 0-1 arasında değer alır. Üç ve daha üzeri parametreye sahip modellerde bulunur.

*Üst limit parametresi (d)*: Şans parametresinin tersi olarak düşünülebilir. En yüksek yetenek seviyesindeki bireyin maddeyi doğru yanıtlama olasılığını verir. Madde karakteristik eğrisinin y ekseninde yatay hale geleceği olasılığı gösterir. 0-1 arasında değer alır. Dört ve daha üzeri parametreye sahip modellerde bulunur.

*Yetenek parametresi ( $\theta$ )*: Maddeler aracılığıyla yoklanan gizil değişken için modelce kestirilen bireyin yeteneğini temsil eden değerdir. Tanımlanan yeteneğin bireyde bulunan sayısal anlamını taşır. Birey için sabit olarak görülmemeli, zamanla değişebileceği göz önünde bulundurulmalıdır (Hambleton & Swaminathan, 1985). Madde zorluk parametresi ile aynı ölçek üzerinde bulunur.



**Şekil 1.1. Madde Karakteristik Eğrisi**

Şekil 1.1'de  $a=1,2$ ,  $b=0,75$  ve  $c=0,15$  parametrelerine sahip madde için ilgili yetenek düzeylerinde maddeye doğru yanıt verme olasılığını gösteren madde karakteristik eğrisi bulunmaktadır. Görüldüğü üzere yetenek artışı maddeyi doğru yanıtlama olasılığındaki artışı da beraberinde getirmektedir.

### **1.5.2. Çok Boyutlu Madde Tepki Kuramı (ÇBMTK)**

ÇBMTK konu uzmanının bakış açısına göre tek boyutlu MTK'nin, faktör analizinin veya yapısal eşitlik modelinin farklı bir formu olarak görülebilir. ÇBMTK ve faktör



analizinde benzer matematiksel altyapı kullanılmaktadır. Ayrıca doğrulayıcı ve açıklayıcı çeşitlilik yönünden de yakın görünmektedir. Ancak faktör analizinde veriyi açıklamak adına korelasyon matrisini kullanarak en az sayıda boyut üzerinde yoğunlaşırken veriye ait değişkenlerin zorluk ve ayırt edicilik gibi birçok özelliği göz ardı edilir. Bu özellikler ise ÇBMTK'nin ayrılmaz bir parçasıdır (Reckase, 1997a). MTK çıkışı itibariyle tek boyutluluk üzerinedir. Hatta tek boyutluluk karşılanması gereken varsayımlar arasında bulunmaktadır (Hambleton ve Swaminathan, 1985). Geleneksel MTK veya MTK denildiğinde tek boyutlu MTK kastedilmektedir. Ancak psikolojik ve eğitsel testlerde tek boyutluluğu sağlamada görülen güçlükler ve tek boyutluluk varsayımının ihlali ÇBMTK'yi gündeme getirmiştir. Tüm testler birden fazla boyuta ilişkin yüklemeye sahiptir ancak bazı boyutlar arasında bulunan çok yüksek korelasyon nedeniyle testler boyutlar arası farka duyarlı olmayabilir. Bu durumlarda tek boyutlu MTK modellerine başvurulmalıdır (Reckase, 2007). ÇBMTK iki ya da daha fazla yapı veya boyut olarak bir arada ele alınan gözlenemeyen değişkenler ve bireyin herhangi bir test maddesine doğru yanıt verme olasılığı arasındaki ilişki için kullanılır (Ackerman, Geirl & Walker, 2003).

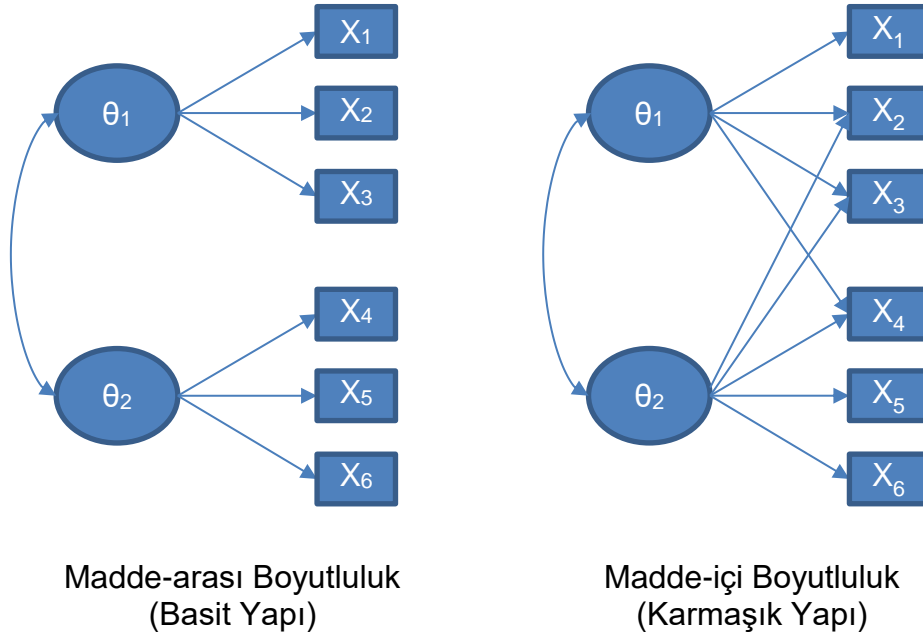
ÇBMTK henüz tam olarak tüm yönleriyle ortaya konulmuş bir model değildir. Tek boyutlu MTK kıyaslandığında bu eksiği ön plana çıkmaktadır. Bunda modelin sonradan ortaya çıkış zamanı ve karmaşık yapısı önemli rol oynamaktadır. ÇBMTK'ye ilişkin verilebilecek ilk model örnekleri 1970'li yıllara dayanmaktadır (Reckase, 2007). Sonraki yıllarda temel modellerin birçok formu ortaya konmuş bazıları pratik yetersizliğinden çok fazla uygulama alanı bulamamıştır (Embretson ve Reise, 2000).

Modelin karmaşıklığı parametre tahminlerine yansımaktadır. Bireyin tüm boyutları kapsayan uzay konumu bilinseydi, bireyin maddeye vereceği cevap yüksek bir doğruluk payıyla bilinebilirdi. Fakat bireyin uzay konumu pratikte bilinemez (Reckase, 2007). Birey ve madde parametrelerinin birden fazla boyut için konumlandırma çabası parametre tahminlerine MTK'ye nazaran daha fazla standart hata olarak yansıyabilir.

MTK'de yer alan birey yeteneği tek bir değer iken ÇBMTK'de her bir boyut için ayrı ayrı yer almaktadır. Bireye ait yetenekler vektör halinde tanımlanır. Her boyuta ilişkin yetenek ilgili olduğu maddelerin arka planında örtük değişken olarak bulunur. ÇBMTK'de maddeyi çözmek için gerekli yetenek bir veya birden fazla olabilir.

Maddeyi doğru yanıtlamak için gerekli yeteneğin bir tane olması testin değil maddenin tek boyuta ait olduğu anlamına gelmektedir.

Testte yoklanan yeteneğin birden fazla olması ÇBMTK uygulamasını gündeme getirir. ÇBMTK ile birlikte psikometri alanına madde-arası ve madde-içi boyutluluk gündeme gelmiştir. Madde-arası boyutluluk testteki tüm maddelerin testte yoklanan herhangi tek yetenek için gözlenen değişken olması durumudur. Basit yapı olarak da adlandırılmaktadır. Madde-içi boyutluluk ise testteki en az bir maddenin doğru yanıtlanması için birden fazla yeteneğin varlığı gerekmektedir. Karmaşık yapı ismi de bu tür için kullanılmaktadır. Çoğu çok boyutlu analiz yorumlamasının kolaylığı nedeniyle basit yapıdaki testlere odaklanılmaktadır (Ackerman, Geirl ve Walker, 2003).



**Şekil 1.2. Basit ve Karmaşık Yapıdaki Maddeler**

Şekil 1.2'de basit ve karmaşık yapıdaki modellere örnek verilmiştir. Altı maddeden oluşan bir test için iki farklı yetenek gerektirdiği düşünüldüğünde basit yapıdaki ilk üç madde sadece  $\theta_1$  olarak gösterilen boyutla ilişkilidir.  $X_{4,5,6}$  olarak gösterilen diğer üç madde ise sadece diğer boyutla ilişkilidir. Şeklin sağ tarafında görülen karmaşık yapıdaki teste baktığımızda ise 2,3 ve 4. maddelerin her iki boyut için de gözlenen

değişken olduğu görülmektedir. Bir maddenin birden fazla boyuta ait olması nedeniyle madde-içi boyutluluk yapısında olduğu görülmektedir.

ÇBMTK'de en çok dikkat çeken sınıflama telafi edici ve telafi edici olmayan modellerdir. Telafi edici modellerde bireylerin yetenekleri birbirinin eksikliğini tamamlayabilirken, telafi edici olmayan modellerde böyle bir durum söz konusu değildir. Model seçiminde boyutların birbiri açısından telafi durumu incelenip karar verilmesi yerinde olacaktır. Araştırma için geçerli madde birey etkileşimi varsayımının uyumuna göre model seçimi yapılabilir. Telafi edici model madde birey etkileşimine bütüncül gözle bakarak bireyin tüm bilgi, yeteneğini maddelerin hepsine birden yansıtır. Telafi edici olmayan modellerde ise maddelerin farklı parçaları, farklı bilgi, yeteneklerle ilişkilidir ve maddede başarılı olmak için her parçada başarılı olmak gereklidir (Reckase, 2009).

### **Model Parametreleri**

Daha önce değinilen geleneksel MTK'de olmayıp ÇBMTK'de yer alan parametreler şunlardır.

*d (Kesişim) Parametresi:* ÇBMTK'deki  $d$  parametresi MTK'de nadiren olsa rastlanan üst limit parametresiyle ilişkili değildir. Buradaki  $d$  parametresi MTK'deki güçlük ( $b$ ) parametresinin yerini almıştır. Ancak ÇBMTK'de  $d$  parametresi  $a$  parametresi gibi vektör olarak değil her bir madde için

$$d_i = - \sum_{k=1}^m (a_{ik} b_{ik})$$

olarak hesaplanır. Formüldeki  $d$  parametresi güçlük (konum) ve ayırt edicilik (eğim) parametrelerinden oluşması nedeniyle kesişim (intercept) parametresi olarak adlandırılmaktadır (de Ayala, 2009).

*Çok Boyutlu Madde Ayırt Edicilik (MDISC) Parametresi:* Maddenin genel olarak bireyler için ayırt ediciliğini ortaya koyar. Bu parametre hesaplanmadan ifade edilecek her ayırt edicilik parametresi maddede bulunan boyutların madde için ayırt edicilik indeksi olarak görülür. MDISC, madde bir boyuta yük vermediği sürece, boyutlar için hesaplanan ayırt edicilik indekslerinden büyük olarak hesaplanacaktır. Maddenin ayırt edicilik gücü boyutların en ayırt ediciliğe sahip kombinasyonundan oluşur (Reckase & McKinley, 1991; Reckase, 1997b).

$$MDISC_i = \sqrt{\sum_{k=1}^m a_{ik}^2}$$

*Çok Boyutlu Madde Güçlük (MDIFF) Parametresi:* Kesişim parametresi boyutların güçlüğü'nün yanı sıra ayırt ediciliğinin de etkisinde olduğundan madde için toplam olarak hesaplanırsa dahi MTK'deki  $b$ , güçlük parametresi gibi yorumlanamaz. Maddenin tamamının için verilen toplam  $d$  parametresinden ayırt ediciliğin etkisi arındırıldığında elde edilecek parametre MTK'deki  $b$  parametresine denk gelecektir. ÇBMTK'de madde için hesaplanan kesişim parametresi, ayırt edicilik parametresine bölüldüğünde elde edilecek parametre güçlük parametresi olarak ele alınabilir. En yüksek ayırt ediciliğe sahip noktanın eksenden uzaklığını belirtir (Reckase, 1985; 1997b). ÇBMTK'de de konum parametresi olarak da adlandırılır.

$$MDIFF_i = \frac{-d_i}{MDISC_i}$$

### **Telafi edici Modeller**

Telafi edici modellerde bireyin bir maddeye doğru yanıt verme olasılığı hesaplanırken bir boyuta ait yüksek düzeyde yeteneğin bulunması başka bir boyutta bulunan düşük düzeydeki yeteneğin negatif etkisini ortadan kaldırır. Yoklanan yetenek parametreleri logit formunda toplama işlevine sahiptir. Bu model de MTK'de olduğu gibi parametre girdisine göre farklı modeller içermektedir.

*1 ve 2 PL ÇBMTK Modelleri:* Farklı formlarda başka yazarlar tarafından ifade edilse de McKinley ve Reckase (1982) tarafından ortaya konulan model halen yaygın olarak görülmektedir.

$$p(x_{ij} = 1 | \theta_j, a_i, d_i) = \frac{1}{1 + \exp(-D \sum_{k=1}^m (a_{ik} \theta_{jk}) + d_i)}$$

Formül MTK için verilen 2PL formülü ile kıyaslandığında farklılıklar göze çarpmaktadır. MTK'de bireyin sahip olduğu tek yetenek parametresi burada  $(\theta_1, \dots, \theta_m)$  olarak  $m$  tane boyuta ait birey parametresi olarak vektör şeklinde karşımıza çıkmaktadır. Aynı şekilde MTK'de madde için bir tane bulunan ayırt edicilik,  $a$  parametresi yine  $m$  tane boyut için  $m$  tane hesaplanmaktadır.

Verilen formülde  $a$  parametresi MTK'de olduğu gibi tüm maddeler ve boyutlar için eşit ele alınırsa 1 PL ÇBMTK modeli için bireyin maddeyi doğru yanıtlama olasılığı

elde edilecektir. Formül bu haliyle kısıtlama olmaksızın kullanıldığında 2PL ÇBMTK modeli olarak kullanılır. Formülde  $j$  indisi bireyi,  $i$  indisi ise maddeyi temsil etmektedir.

**3 PL ÇBMTK Modeli:** MTK'deki 2PL ve 3PL modelleri arasında bulunan fark ÇBMTK modellerine de yansımıştır.

$$p(x_{ij} = 1 | \theta_j, a_i, d_i, c_i) = c_i + \frac{1 - c_i}{1 + \exp(-D \sum_{k=1}^m (a_{ik} \theta_{jk}) + d_i)}$$

Formüldeki  $c$  parametresi tek boyutlu MTK modellerinde olduğu gibi en düşük yetenek düzeyindeki bireyin maddeyi doğru yanıtlama olasılığı, şans parametresi olarak ele alınır. Her madde için sadece bir değer alır.

### **Telafi edici Olmayan Modeller**

Telafi edici olmayan modellerde boyutlardan gelen yetenek parametreleri çarpım halinde son terimi oluştururlar. Çarpımda telafi etkisi tamamen kaldırılmadığından kısmi telafi edici modeller olarak da adlandırılır (Reckase, 2009). Model logit formunda çarpım halindedir. Maddeye doğru yanıt verme olasılığı en fazla, boyutların ayrı ayrı değerlendirildiğinde elde edilen olasılıklardan yani logit bölümlerin küçük olanına eşit olabilir. Çünkü bir boyuttan gelebilecek maddeye doğru yanıt verme olasılığı çok nadir bir şekilde en fazla 1 olabilir. Bu olasılık daha küçük olanla çarpıldığında elde edilen sonuç daha küçük çıkacaktır. 1978 yılında Sympson tarafından ortaya konulan telafi edici olmayan ÇBMTK formülü 3PL için şu şekildedir (Akt:Reckase, 2007);

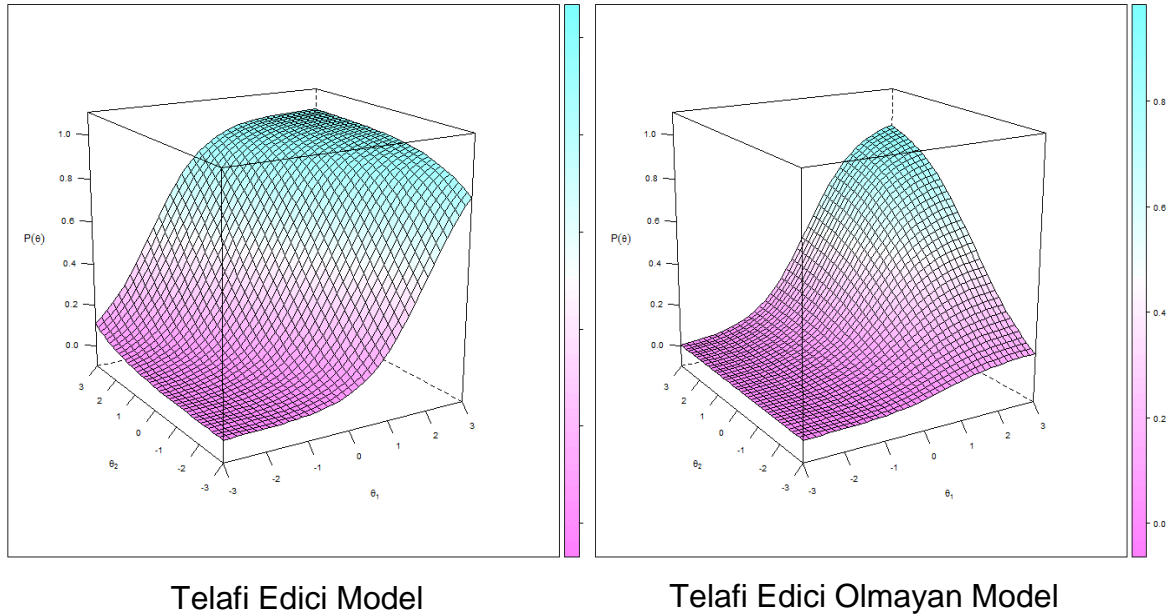
$$p(x_{ij} = 1 | \theta_j, a_i, d_i, c_i) = c_i + 1 - c_i \prod_{k=1}^m \frac{1}{1 + \exp(-D a_{ik} (\theta_{jk} - b_{ik}))}$$

Model direk 3PL olarak tasarlanmıştır. MTK'ye benzer şekilde  $c$  parametresi 0'a eşitlenirse 2PL'ye indirgenir. Madde ayırt edicilik,  $a$  parametresi de tüm madde ve boyutlar için eşitlenirse 1PL modele indirgenmiş olur. Model formülü telafi edici model formülü ile kıyaslandığında toplam işlevi yerine çarpım işlevi olduğu görülmektedir. Birey parametresi güçlük parametresine eşit olduğunda maddeye doğru yanıt verme olasılığının 0,50 olduğunu hatırlayarak formüle bakacak olursak maddeye bu özelliklere sahip bir boyut eklenmesi maddeye doğru yanıt verme olasılığını yarıya indirecektir (Reckase, 2007). Bu durum telafi edici modeller için söylenemez. Ayrıca madde kesişim parametresi yerine her madde ve her boyut için hesaplanan güçlük parametresi vektör olarak ifade edilir.

Telafi edici olmayan model özellikle açılmayıcı çalışmanın gerektiđi durumlarda tahmin hesaplama sıkıntısı nedeniyle arařtırmalarda ikinci planda kaldıđı grlmektedir (Chalmers & Flora, 2014). Bu sebeple alanda telafi edici modellerin hâkimiyeti sz konusudur.

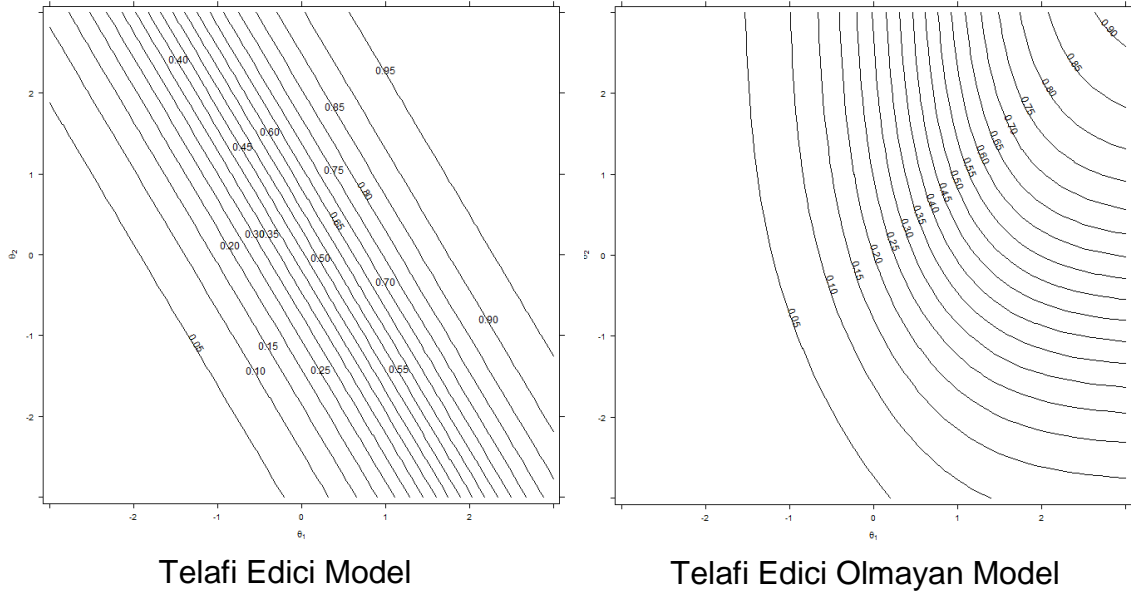
Tek boyutlu MTK'de eđri olarak gsterilen madde tepki fonksiyonu, ÇBMTK'de yzey olarak temsil edilir. Ancak grafik aısından 3'den fazla boyutun birlikte madde tepki fonksiyonunda gsterimi pek mmkn grnmemektedir. Bu yzden kolay anlatım aısından 2 boyutlu madde tepki fonksiyonu gsterimi yaygındır.

İki boyutlu  $a_1=1,4$  ,  $a_2=0,8$  ,  $b_1=0,55$  ,  $b_2=-0,61$  ,  $-a_1*b_1=d_1=-0,75$  ,  $-a_2*b_2=d_2=0,5$  ve  $d_1+d_2=d_f=-0,25$  parametrelerine sahip madde telafi edici ve telafi edici olmayan modellerine uygulayıp madde tepki fonksiyonları zerinden rnek olarak Őekil 1.3'de verilmiřtir.



**Őekil 1.3. Telafi Edici ve Telafi Edici Olmayan Madde Karakteristik Eđrileri**

Modeller arasındaki farkın grafiđe son derece net bir Őekilde yansıdađı grlmektedir. Grafikler arasında gze batan ilk fark telafi edici modelin forml geređi daha yksek dođru yanıtlanma olasılıđına sahip olduđunun grlmesidir. Sađda grlen telafi edici olmayan model ise iki farklı boyuttan gelen olasılıđın arpımı zerine kurulduđundan maddeye dođru yanıt vermek ok daha zor hale geldiđi grlmektedir.



**Şekil 1.4. Telafi edici ve Telafi edici olmayan Madde Tepki Yüzeyleri**

Şekil 1.4 ise madde tepki fonksiyonu yüzeylerinin başka bir gösterimi olarak kabul edilebilir. Şekil 1.3’de verilen yüzeyin izlerini barındırmaktadır. Madde tepki yüzeyinin daha dik olması, çizgilerin birbirine yaklaşmasını ve o bölgenin daha ayırt edici olduğu gösterir (Ackerman, Geirl & Walker, 2003). Özellikle madde parametrelerinin etkisi bakımından daha kolay yorumlanabilir. Maddeye ait ilk boyut orta üzeri ayırt ediciliğe sahipken, ikinci boyut orta-düşük ayırt ediciliğe sahiptir. Madde konumları açısından ilk boyut güçlük parametresi kısmen yüksek iken, ikinci boyutun güçlük parametresi kısmen düşüktür. Buna göre ilk boyut daha seçici ve zor, ikinci boyut daha zayıf ve kolaydır.

Bu parametrelerin telafi edici modele yansımalarına baktığımızda ilk boyutun daha yüksek ayırt ediciliğe sahip olması nedeniyle ilk boyuttaki yetenek değişimi maddeyi doğru yanıtlama olasılığına daha hızlı yansımaktadır. Bu durum ikinci boyut için daha yavaştır. Eksenlerin biri sabit tutularak bakıldığında ve olasılığın değişimi incelendiğinde bu sonuç ortaya çıkmaktadır. Maddenin kesişim parametresi telafi edici modelde tüm madde için geçerli olduğundan boyutlar için farklılık göstermeyecektir. Daha zor yani daha büyük b değerleri için  $d_i$ 'nin -0,25’den daha küçük bir değer alması x eksenini açısından baktığımızda çizgileri sağa doğru kaydıracaktır.

Şeklin sağ tarafındaki telafi edici olmayan modele ait madde tepki izlerine bakıldığında telafi edici modelde görülmeyen çizginin eğriliği göze çarpmaktadır. Bu durum olasılık hesaplamada kullanılan çarpım etkisinden kaynaklanmaktadır. Maddede bulunan boyutların ayırt edicilikleri bu modelde çok daha fazla ön plana çıkmaktadır. Yatay eksene dik bir çizgi çizdiğimizde x ekseninde sabitlenen grafiğimiz için  $\theta_1$  değeri için  $\theta_2$ 'deki değişimin olasılığa yansımalarının çok az olduğu görülmektedir. Bu durum  $a_2$  parametresinin düşüklüğünden kaynaklanmaktadır. İşlem dikey eksen için tekrarlandığında  $a_1$  parametresinin daha iyi olduğu yorumuna erişebiliriz. Telafi edici modelde her bir boyut için hesaplanan madde güclüğü etkisine baktığımızda düşük  $\theta_1$  değerleri için olasılık çizgilerine rastlanılmamaktadır. Bu durumda ilk boyutun daha zor olduğu sonucuna götürür. Ayrıca her boyut için gerçekleşecek zorluk değişimi ilgili boyut eksenindeki çizgilerin başlangıç konumlarını etkileyecek ve çizgi diğer boyut eksenine ulaştığında etki kaybolacaktır.

### 1.5.3. Bilişsel Tanı Modelleri

Son 20 yılda psikometrik alanda kendini gösteren BTM'lere olan ilgi artarak devam etmektedir. Literatürdeki kaynaklarda BTM'ler Bilişsel Tanı Değerlendirme (Cognitive Diagnosis Assessment) amacını taşıyan testlere aracılık eden modeller olarak tanımlanabilir (de la Torre & Minchen, 2014). BTM'ler genel olarak kişileri biliş göstergelerine göre sınıflama işlemi yapmaktadır.

Yurtdışı kaynaklarda BTM'ler pek çok farklı isimlerle anılmaktadır (Rupp, Templin & Henson, 2010). Bunlardan en çok kullanılanları "Cognitive Diagnosis Models" (Bilişsel Tanı Modelleri) ve "Diagnostic Classification Models" (Tanısal Sınıflandırma Modelleri)'dir. Yapılan işlemler göz önüne alındığında farklılığın isimlendirme boyutunda kaldığı söylenebilir. Türkiye'de ise modeller ilk kez Başokçu (2011) tarafından kullanılmış ve Bilişsel Tanı Modelleri olarak ifade edilmiştir. Bu araştırmada da aynı isimle anılması uygun görülmüştür. BTM'lerde ölçme birimi olarak görülebilecek nitelikler ise yine yurtdışı kaynaklarda "skill", "attribute" gibi ifadelerle anılmaktadır. Yurtiçi kaynaklarda geçen "özellik" ifadesi (Başokçu, 2011; 2012, Sünbül, 2013 & Demir, 2013) yerine BTM doğasına daha uygun olduğu düşünülen "nitelik" ifadesi kullanımı daha uygun görülmüştür.

BTM'nin amacı testte yoklanmak istenen niteliklerin öğrencilerde olup olmamasına göre öğrencileri örtük kategorilere sınıflamaktır. BTM testi alan bireylerin yeteneklerini, maddeleri doğru cevaplamak için gerekli olan niteliklere sahip olup



olmamalarına göre 0-1 olarak belirleyen modellerdir. Araştırılan örtük değişkenler beceri, yetenek, nitelik, özellik, görev olarak isimlendirilebilir. Diğer psikometrik yöntemlerden farklı parametreler elde ederek hem testi alan bireylerin maddeleri doğru cevaplamak için gerekli özelliklere sahip olup olmadığını belirlemekte hem de maddeler hakkında farklı altyapıda parametreler üretmektedir. Bireyin soruyu doğru yanıtlaması ilgili soruyu doğru yanıtlamak için gerekli nitelik veya niteliklerin öğrencide bulunduğuna işaret etmektedir. Belirli bir nitelik için düşünülecek olursa bu niteliği gerektiren soruların yüksek oranda doğru yanıtlanması öğrencinin niteliğe sahip olduğuna dair sonucun geçerliliğini artıracaktır.

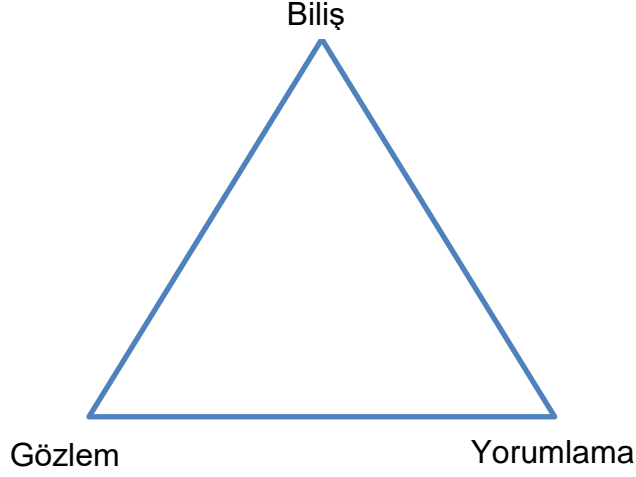
Sinharay, Puhan ve Haberman (2009) BTM'lerin ortak noktalarını şu şekilde göstermektedir:

- i. Her test maddesinin doğru çözümü bir veya birden fazla nitelik gerektirir.
- ii. Her birey her bir beceriye karşılık gelen örtük yetenek parametresine sahiptir.
- iii. Bireyin maddeye doğru yanıt verme olasılığı maddenin gerektirdiği niteliğin ve bireyin örtük yetenek parametresinin matematiksel fonksiyonudur.

### **Bilişsel Tanı Modellerinin Temelleri**

Bilişsel değerlendirme sürecinde karşılaşılan temel zorlukları Junker (1999) "değerlendirmenin amacı nedir, öğrenci hakkında ne tür çıkarımlarda bulunulabilir, çıkarımlar için gereken bilgi nedir ve nasıl elde edilir, nasıl çıkarımda bulunur ve raporlarız" sorularıyla sıralamıştır. Buradan ölçme-değerlendirme sürecine ilişkin adımlar da görülebilir. Benzer şekilde işlemler zinciri başka değerlendirme yaklaşımlarında da bulunmaktadır.

Pellegrino, Chudowsky ve Glaser (2001) tarafından Ulusal Araştırma Kurulu (National Research Council) için hazırlanan "Öğrencilerin Ne Bildiğini Bilmek" (Knowing What Students Know) adlı raporda üzerinde durulan değerlendirme üçgeni Bilişsel Tanı Değerlendirmenin teorik yapı taşlarındandır. Şekil 5'te gösterilen değerlendirme üçgeni biliş, gözlem ve yorumlama köşelerini içermektedir. Biliş köşesi öğrencinin alan hakkındaki bilgisini ve bunu yansıtmalarını gösterir, gözlem köşesi öğrenci seviyesi hakkında kanıt sağlama işlevini görür, yorumlama köşesi ise elde edilen kanıtları kullanılabilir hale getirmeyi temsil etmektedir. Sağlıklı bir değerlendirme için bu üç eleman birbiriyle etkileşim halinde olmalıdır.



**Şekil 1.5. Değerlendirme Üçgeni**

Değerlendirme üçgenine benzer bir yaklaşım da kanıt merkezli yaklaşımdır. Bu yaklaşımda öğrenci, kanıt, görev, birleştirme, sunum ve teslim sistemi modellerini barındırmaktadır. Öğrenci modeli ölçülmek istenen yetenek, bilgi, nitelik vb. ile ilgili bir veya birden fazla değişken tanımlar. Kanıt modeli, öğrencilerin değişkenlere verdikleri cevaba göre öğrenciler hakkındaki bilgileri nasıl güncelleneceği üzerinde durur. Görev modeli, kanıt modeline türlü destek sağlamak için değişik durumların nasıl yapılandırılması gerektiğini tanımlar. Birleştirme modeli ise öğrenci kanıt ve görev modellerinin birlikte çalışarak değerlendirmenin temelini oluşturmasını sağlar. Bu modelle ne kadar ölçüm yapılacağı tanımlanmış olur. Sunum modeli görevlerin farklı formatlarda nasıl ortaya konulabileceğini araştırır. Teslim sistemi modeli diğer tüm modelleri kapsayan birlikte çalışmalarını düzenleyen sistemdir (Mislevy, Almond & Lukas, 2004).

Hem değerlendirme üçgeni hem de kanıt merkezli yaklaşım, bilişsel tanı değerlendirme teorik temellerini oluşturmaktadır. Bu sistemler takip edilerek nitelik tanımlama, nitelik madde arası ilişki kurma, niteliklerin testte temsili, istatistiksel model seçimi, analizlerin yorumlanması gibi birçok işlem düzenli bir şekilde sonlandırılabilir. İşlem döngüsü takip edilerek ölçüm öncesindeki amaç ile değerlendirme sonrasındaki yorumlar tutarlı hale getirilir.

BTM'nin uygulamada yaygınlaşmasında 2001 yılında başlayan No Child Left Behind eğitim hareketinin büyük rol oynadığı söylenebilir (DiBello & Stout, 2007). Günümüzde hâlâ yaygın olarak kullanılan modeller bu akım sonrası ortaya çıkmıştır.

Bu eğitim hareketinin felsefesi ile BTM'nin amacının yakınlığı, BTM'lerin daha ön planda çıkmasına yardımcı olmuştur. BTM'ler konu alanlarında verdikleri geri bildirimler ile eğitim faaliyetinin bireydeki etkisini, zayıf ve güçlü yönlerini ortaya koyarak eğitimin bireyler için gerektiğinde yeniden düzenlenmesine ilişkin önemli ipuçları sunmaktadır. Böylece bireysel dönütler ile eğitimin yeniden düzenlenip öğrencinin üst düzeye çıkmasına yardımcı olunabilir.

BTM'ler MTK uzantısı olarak görülebileceğinden ilk BTM uygulamaları olarak MTK modelleri gösterilebilir. Ancak burada Linear Logistic Test Model (LLTM) ve Rule Space Model (RSM) üzerinde durulacaktır. LLTM Fischer tarafından 1973 yılında çıkarılmış ve 1983'de güncellenmiştir. Model günümüzdeki BTM uygulamalarında bulunan nitelik vektörünü tahmin etme işlemine sahip değildir. Modeli BTM için ayrıcalıklı kılan unsur model tek boyutlu MTK yapısında olmasına rağmen her bir madde için bilişsel işlemi içeren Q matrisini barındırmasıdır. Bu bilişsel işlemler modelde maddelerin özelliği olarak ele alınmakta, bireyler için sınıflama işlemine gidilmemektedir. Q matrisi aracılığıyla 1 parametrelili lojistik modeldeki madde zorluğu parametresi elde edilmektedir.

$$b_i = \sum_{k=1}^K q_{ik}\eta_k + c$$

$b_i$ : madde zorluğu

$q_{ik}$ :  $k$  niteliğinin  $i$  maddesinde bulunma durumu (0-1)

$\eta_k$ :  $k$  niteliğinin güçlük düzeyi

$c$ : ölçek orjininde sabitleme sayısı

Model tek boyutlu yapıları ölçmeyi amaçlamasına rağmen elde edilen yetenek düzeyi maddelerde bulunan niteliklerle doğrudan ilişkili olduğundan, yetenek parametresi niteliklere sahip olma durumlarının ağırlıklandırılmış ortalaması olarak görülebilir (DiBello, Roussos & Stout, 2007).

Rule Space Modeli (RSM) Tatsuoka (1983) tarafından tanıtılmıştır. Model LLTM de olduğu gibi tek boyutlu MTK üzerine inşa edilmiştir. Model temel olarak bireylerin verdikleri yanıt örüntüsüyle, model tarafından hesaplanan ideal cevap örüntüsü arasında kurallar oluşturularak bireyleri boş bir düzleme yerleştirme işlemi yapmaktadır. Düzlemin yatay boyutunu MTK'den elde edilen yetenek parametresi,

dikey boyutunu ise ideal cevap örüntüsünden uzaklaşma birimi oluşturur.  $\zeta$  olarak verilen bu parametre MTK modeline uymayan cevap örüntüsünü belirtir ve uyarı indeksi olarak da anılır (Junker, 1999). Bu düzlemde birbirine yakın bireyler kümelenerek aynı nitelik örüntüsüne sahip olduklarına ilişkin veri elde edilmektedir. Öğrencilere nitelik profili vermesi bakımından RSM, BTM için kilometre taşlarından biridir.

### Q Matrisi

BTM'lerin ortak özellikleri Q matrisi girdisidir. Q matrisinde soru ile soruları yanıtlamak için bulunması gereken nitelik ilişkilendirilir. 2 boyutlu olan matriste günümüzdeki modeller satırları maddeler, sütunları ise nitelikler olarak ele alırlar. Yapı olarak basit görünmesine rağmen Q matrisinin BTM'ler için oynadığı rol kritiktir. Q matrisinde yapılabilecek küçük bir hata bile analiz sonuçlarını olumsuz bir şekilde etkileyebilir. Bu sebeple Q matrisinin geçerliği hayati bir önem taşımaktadır. Q matrisi sadece madde nitelik ilişkisinde değil aynı zamanda bireylerin yeteneklerin belirlenmesinde önemli rol oynar. Q matrisinde temsil edilen nitelikler için bireylere profil ataması yapılır.

Q matrisinin geliştirilmesi BTM'lerde en fazla çabanın gösterildiği süreç olarak görülebilir. Süreç istenilen sonuca erişilinceye kadar madde ve niteliklerin birbiriyle etkileşimli bir şekilde döngüsel olarak devam eder.

**Tablo 1.1: Örnek Q Matrisi**

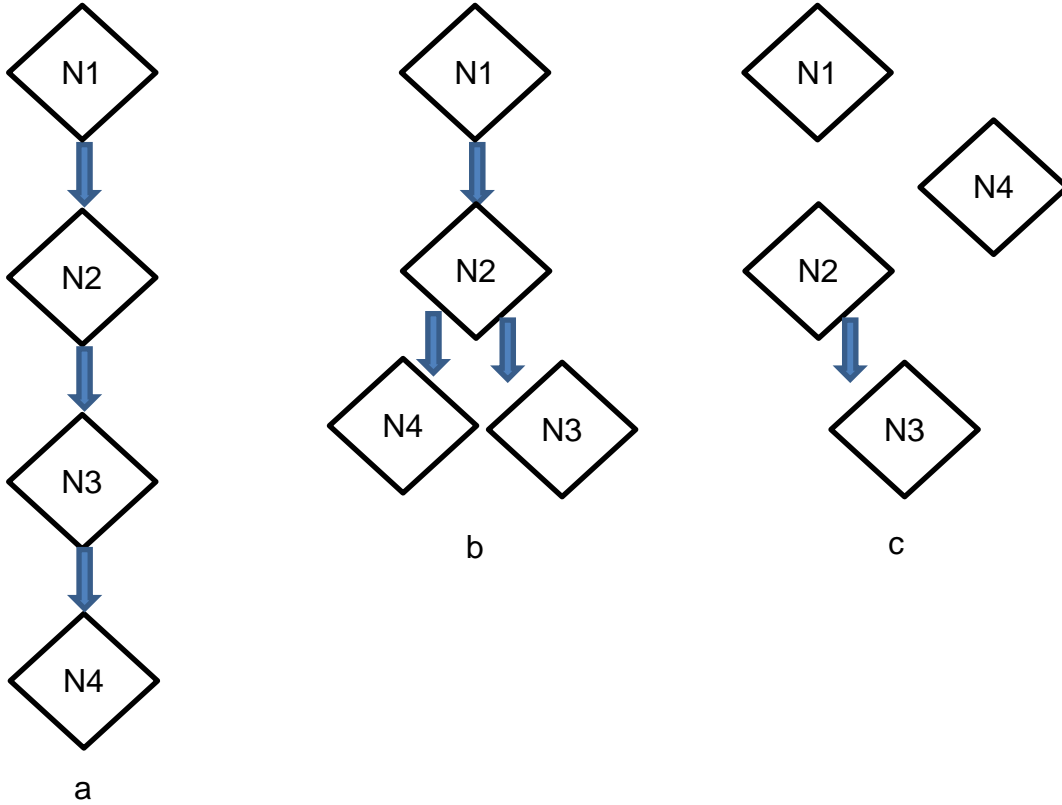
	<i>Nitelik 1</i>	<i>Nitelik 2</i>	<i>Nitelik 3</i>
<b>Madde 1</b>	1	0	0
<b>Madde 2</b>	1	1	0
<b>Madde 3</b>	0	1	1
<b>Madde 4</b>	1	1	1

Tablo 1.1'de 3 nitelik ve 4 madde için hazırlanmış örnek bir Q matrisi görülmektedir. Tabloya göre birinci maddede birinci nitelik, 2. maddede 1. ve 2. nitelik, 3. maddede 2. ve 3 nitelik, son madde ise her üç nitelik ile ilişkilendirilmiştir.

Q matrisinin BTM'ler için merkezi bir yere sahip olması araştırma konusu olarak modelde yer edinmiştir. Q matrisi geçerliği ve yapısı araştırılan konularda ön plana çıkmaktadır. Q matrisi geçerliğinde maddeler ile niteliklerin ilişkilendirilmesinin kalitesi ve doğruluğu araştırılmaktadır. Q matrisi yapısında ise niteliklerin

maddelerle ve kendi aralarındaki ilişkiler göz önünde tutulmakta, değişik yapılardaki Q matrislerden elde edilen sonuçlar karşılaştırılmaktadır.

En temel Q matris ayrımı basit ve karmaşık yapıdaki matrislerdir. Basit yapıdaki matrislerde her bir madde yalnızca bir niteliği ölçecek şekilde yapılandırılmıştır. Q matriste satır toplamı birdir, ancak sütun toplamı birden büyük olabilir. Karmaşık yapıdaki matriste ise en az bir madde birden fazla nitelikle ilişkilendirilmiştir. Diğer bir matris yapısı ayrımı da hiyerarşik olan ve hiyerarşik olmayan yapıdaki matrislerdir. Hiyerarşik olmayan Q matrisinde nitelikler birbirinden bağımsız olarak maddelerde temsil edilebilir, nitelikler arasında ilişki yoktur. Hiyerarşik yapıdaki Q matrisinde ise niteliklerin hepsi veya bazıları birbiri ile ilişkilidir. Bu ilişki birbirinin önkoşulu olma yönündedir. Hiyerarşik matrisler kendi içlerinde de farklı türler barındırmaktadır.



**Şekil 1.6. Farklı Hiyerarşi Yapıları**

Şekil 1.6'da verilen hiyerarşik örnekler incelendiğinde, a örneğinde lineer bir ilişki olduğu görülmektedir. Burada N1 niteliği N2 niteliğine, N2 niteliği N3'e, N3 ise N4 niteliğine ön koşuldür. Bireylerden N1 niteliğine sahip olmaksızın N2'ye doğrudan, N3 ve N4 ise dolaylı olarak sahip olmaları beklenmez. Aynı durum N2'ye sahip

olmadan N3 doğrudan ve N4'e dolaylı olarak, N3'e sahip olmadan da doğrudan N4'e sahip olmaları beklenmez. Şekilde ortada yer alan b örneğine bakıldığında a örneğinden farklı olarak N3 ile N4 arasındaki bağın koptuğunu, N4'ün direk N2 ile bağlantılı hale geldiğini görmekteyiz. Buradan N4'ün önkoşulunun N2 olduğunu ve N2'nin artık N3 ve N4 ile doğrudan ilişkili hale geldiği görülmektedir. Sağdaki c örneğine bakıldığında ise sadece N2nin N3 niteliği için önkoşul olduğunu görmekteyiz. Bu durumda N1, N2 ve N4 nitelikleri birbirinden bağımsız bir halde Q matrisinde temsil edilebilir.

Hiyerarşik yapıdaki Q matrislerinin yazılabilecek soru sayısına ve analiz sonrası ortaya çıkabilecek öğrenci profillerine doğrudan etkisi bulunmaktadır. Hiyerarşik olmayan 4 nitelik barındıran bir testte  $2^k-1=16-1=15$  farklı nitelik örüntüsü barındıran soru oluşturulabilir. Soldaki a örneği ele alınacak olursa sadece 4 farklı nitelik örüntüsü barındıran soru oluşturulabilir. Bu sayılara hiçbir niteliğe sahip olmayan bireylerin oluşturduğu 0 grubu da eklenirse hiyerarşik olmayan 4 nitelikli bir testte 16 farklı öğrenci profili oluşturulabilirken, yine a örneği için sadece 5 farklı öğrenci profili oluşturulabilir.

Q matrisindeki diğer bir farklı tür yaklaşımı ise iki kategorili ve çok kategorili nitelik bulundurmasıdır. İki kategorili yaklaşımda nitelik soruda yoktur veya vardır ve matriste 0-1 ile temsil edilir. Yine aynı şekilde birey profilinde de nitelik bireyde yoktur veya vardır ve matriste 0-1 ile temsil edilir. Çok kategorili yaklaşımda ise niteliğin soruda ve bireyde bulunma seviyesi vardır. Örneğin 5 kategorili bir niteliğin, 1 tane sahip olmama ve 4 tane de farklı seviyelerde sahip olma durumu söz konusudur. Çok kategorili yaklaşımın doğrudan lineer hiyerarşi yaklaşımı ile de bağlantılı olduğu söylenebilir. Bu iki yaklaşım birbirine dönüştürülebilir. Şekil 1.5'deki a örneğini tekrar ele alacak olursak bireyin hiçbir niteliğe sahip olamaması çok kategorili yaklaşımda 0 kategorisine atanması ile eşdeğerdir. Sadece N1 niteliğine sahip olması 1, N2 ve zorunlu olarak N1'e sahip olması 2, N3 ve zorunlu olarak N1 ve N2'ye sahip olması 3, N4'e ve zorunlu olarak N1,N2 ve N3'e sahip olması 4. seviyede sınıflanmasına neden olur. Aynı dönüşüm hiyerarşik yapıdaki Q matrisi için de geçerlidir. Buradaki 4 nitelik, 4 seviyeye sahip tek nitelik gibi ele alınmaktadır. Tersine işlem yani çok kategorili Q matrisi, lineer hiyerarşik yapıdaki Q matrisine de kolaylıkla dönüşüm sağlanabilir.

## **Başlıca Bilişsel Tanı Modelleri**

Günümüzde pek çok BTM mevcuttur. Kimi modeller üzerinde düzenlemeye gidilerek yeni bir model oluşturulduğu da sıkça görülmektedir. BTM'leri sınıflamada farklı yaklaşımlar görülebilmektedir. Ele alınan genel yaklaşımlardan biri niteliklerin birbiri ile ilişkisini gösteren bağlayıcı (conjunctive) ve bağlayıcı olmayan (disjunctive) modeller ayrımıdır. Diğer sınıflama ise yanıt ve nitelik örüntüsünün 2 veya daha fazla kategorili olmasına göre yapılmaktadır (Rupp & Templin, 2008, DiBello, Roussos & Stout, 2007). Burada tanıtılan modeller genel bir çerçeveye sahip iki ve çok kategorili nitelik örüntüsüne sahip doygun modellerin yanı sıra sınırlandırılmış modellerden bağlayıcı, bağlayıcı olmayan ve her iki gruba girmeyen modellerdir. Genel ve daha geniş sınırlara sahip kapsayıcı modeller olan doygun modellere uygun kısıtlamalar uygulandığında yaygın olarak kullanılan kimi BTM'lerin genel modellerin özel durumları olduğu görülebilir.

BTM'lerin farklılaşmasında önemli bir etmen modelin bağlayıcı olup olmadığıdır. Bağlayıcı modellerde soruya doğru yanıt verebilmek için maddenin gerektirdiği tüm niteliklerin bireyde bulunması gerekirken, bağlayıcı olmayan modellerde maddenin gerektirdiği bir niteliğin bireyde bulunması soruya doğru yanıt vermek için yeterli görülmektedir. Bağlayıcı modellerin analizde kullanılmasında niteliklerin etkileşimli ve birbirinden ayrılmaz olduğu varsayımı hâkimdir. Bağlayıcı olmayan modeller ise bireylerin soruya doğru yanıt vermek için farklı stratejiler kullanabileceği düşüncesiyle hareket etmektedir.

BTM sınıflamasında telafi edici ve telafi edici olmayan gruplamasına da gidilebilmektedir. Bu sınıflama ÇBMTK'de da kullanılmaktadır. Bağlayıcı model, telafi edici olmayan modele karşılık gelirken, bağlayıcı olmayan model, telafi edici modele karşılık gelmektedir. Ancak BTM'deki yaklaşım ÇBMTK'deki yaklaşımın uç noktasında olduğu söylenebilir. ÇBMTK'deki telafi edici modelde soruya doğru yanıt vermek için aranılan yetenek parametrelerinden zayıf olanın dezavantajı diğer boyutun iyi düzeydeki yetenek parametresi ile giderilebilir. BTM'deki bağlayıcı olmayan modele bakıldığında ise sorunun gerektirdiği herhangi bir niteliğe sahip olmak, diğer niteliklere bakılmaksızın soruya doğru yanıt verme olasılığını en üst düzeye çıkaracaktır. ÇBMTK'deki telafi edici olmayan modelde soruya doğru yanıt vermek için aranılan yetenek parametrelerinden birinin zayıf olması diğer boyutun yetenek parametresi iyi düzeyde olması durumunda dahi doğru yanıtlama olasılığını

düşürecektir. BTM'deki bağlayıcı modelde ise sorunun gerektirdiği herhangi bir niteliğe sahip olmamak, diğer niteliklere varlığına bakılmaksızın soruya doğru yanıt verme olasılığını en alt seviyeye indirecektir. Bu yüzden alan yazında BTM sınıflamaları ve özellikleri için geçen telafi edici olan ve olmayan modelleri direkt olarak ÇBMTK'de olduğu gibi birebir yorumlamak doğru olmayacaktır.

Sık kullanılan modellerden olan Generalized Deterministic-Input, Noisy-And Gate (G-DINA; de la Torre, 2011) genel bir çerçeve çizerek doygun yapıya sahip bir modeldir. Niteliklerin ikiden fazla kategorisi olması durumunda kullanılan polytomous G-DINA (p-GDINA; Chen & de la Torre, 2013) da doygun modellerdendir. Sınırlandırılmış modellere bakıldığında ise Deterministic-Input, Noisy-And Gate (DINA; Haertel, 1989; Junker & Sijtsma, 2001) bağlayıcı modellerdendir. Deterministic-Input, Noisy-Or Gate ise (DINO; Templin & Henson, 2006) bağlayıcı olmayan modellerdendir. Additive Cognitive Diagnosis Model (A-CDM; de la Torre, 2011) de bu iki sınıfın dışında yer almaktadır. Ayrıca tez kapsamında de la Torre ile birlikte geliştirilen fully Additive-Model (fA-M) tanıtılacaktır.

### **G-DINA (Generalized Deterministic-Input, Noisy-And Gate) Model**

2011 yılında de la Torre tarafından ortaya konan G-DINA, DINA model temel alınarak iki kategorili yanıtlar için geliştirilmiş doygun bir modeldir. Herhangi bir kısıtlama olmaksızın toplamda  $j \cdot 2^k$ , madde sayısı çarpı 2 üzeri nitelik sayısı (madde\*toplam nitelik örüntü sayısı) kadar parametreye sahiptir. G-DINA uygulamada maddenin gerektirdiği niteliklerden daha azına sahip bireyin daha fazlasına sahip bireyden maddeye doğru yanıtlamada yüksek olasılığa sahip olmasına izin vermektedir.

G-DINA da  $P(\alpha_{ij}^*)$  / nitelik örüntüsündeki bireyin  $j$ . maddeye doğru yanıt vermesi olarak ifade edilebilecek formül, her bir nitelik temel etkisi, bu niteliklerin birbirleriyle etkileşimi etkisi ve toplam etkilerine göre parçalara ayrılabilir. Aşağıda G-DINA model için maddeyi doğru yanıtlama olasılık formülü:

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{lk} + \sum_{k'=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'} \alpha_{lk} \alpha_{lk'} \dots + \delta_{j12} \dots \delta_{jK_j^*} \prod_{k=1}^{K_j^*} \alpha_{lk},$$



$\delta_{j0} = j$  maddesi kesişimi

$\delta_{jk} = \alpha_k$  temel etkisi

$\delta_{jkk'} \dots = \alpha_k$  ve  $\alpha_{k'}$  'nin çoklu etkileşiminin etkisi ve

$\delta_{j12\dots K_j^*} = \alpha_1$  'den  $\alpha_{K_j^*}$  'ye etkileşimin etkisi (toplu) anlamına gelmektedir.

Kısaca açıklamak gerekirse  $j$  maddesi için  $\delta_{j0}$  maddede yoklanan niteliklerin hiçbirine sahip olmayan bireyin soruyu doğru yanıtlama olasılığıdır. DINA ve DINO modeldeki  $g$  (tahmin etme) parametresi gibi yorumlanabilir.  $\delta_{jk}$   $k$  niteliğinin tek başına soruyu doğru yanıtlama olasılığına kattığı ek değer,  $\delta_{jkk'}$   $k$  ve  $k'$  niteliklerinin birlikte bulunmalarının doğru yanıtlama olasılığındaki değişime katkısı ve  $\delta_{j12\dots K_j^*}$  maddenin gerektirdiği tüm niteliklere sahip olmanın doğru yanıtlama olasılığındaki değişime katkısı olarak söylenebilir. Madde kesişimi daima pozitif, temel etkiler pozitif beklentili olmakla birlikte, her iki etkileşim etkisi de negatif değer alabilir (de la Torre, 2011).

### **DINA (Deterministic-Input, Noisy-And Gate) Model**

DINA modeli 1989'da Haertel tarafından ortaya konulan Restricted Latent Class Model temel alınarak 2001'de Junker ve Sijtsma tarafından geliştirilmiş ve isimlendirilmiştir. Analiz sonucu ortaya çıkan parametreler ve model formülüne göre en basit BTM yöntemlerindedir (de la Torre & Minchen, 2014). Formülündeki bu sadelik sayesinde üzerinde bazı değişiklikler ile Higher Order - DINA (HO-DINA; de la Torre & Douglas, 2004), Multiple Choice DINA (MC-DINA; de la Torre, 2009) ve G-DINA (de la Torre, 2011) gibi bazı modellerin oluşmasına da önderlik etmiştir.

DINA model temel olarak bireyleri her madde için 2 gruba ayırır. Birincisini madde için gerekli tüm niteliklere sahip bireyler oluşturur. Diğer grupta ise maddede yoklanan niteliklerden en az birine sahip olmayan bireyler bulunur. Buna göre bir bireyin bir maddeye vermesi beklenen yanıt  $\eta_{ij}$  şu şekilde yazılabilir;

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}}$$

Formülde bulunan  $i$  bireyi,  $j$  maddeyi,  $k=1$  Q matrisinde ilgili maddede yoklanan nitelikleri,  $\alpha_{ik}$   $i$  bireyi için  $k$ . nitelik kategorisini (0-1),  $q_{jk}$  Q matrisindeki  $j$ . Maddenin  $k$ . niteliğini temsil etmektedir. Formülden de görüleceği üzere bireyin maddenin

gerektirdiği herhangi bir niteliğe sahip olmaması ( $\alpha_{ik}=0$ ) çarpım işlemi nedeniyle beklenen cevabı 0 yapacaktır. Ancak bu beklentinin her zaman gerçekleşmesi mümkün değildir. Model beklenmeyen yanıtları  $s$  (slip, kaydırma) ve  $g$  (guess, tahmin) parametreleri ile göstermektedir. Modelde  $j$  maddesi için kaydırma parametresi  $s_j=P(X_{ij}=0|\eta_{ij}=1)$  olarak gösterilir ve beklenen cevap doğru iken maddeye yanlış yanıt verme olasılığını gösterir. Tahmin parametresi ise  $g_j=P(X_{ij}=1|\eta_{ij}=0)$  olarak gösterilir ve beklenen cevap yanlış iken maddeye doğru yanıt verme olasılığını verir. Modelin madde yanıtlama fonksiyonu,

$$P_j(\alpha_i) = P[X_{ij} = 1|\alpha] = (1 - s_j)^{\eta_{ij}} g_j^{1-\eta_{ij}}$$

$\eta_{ij}$  'nin 0 ve 1 değerleri alabileceği düşünüldüğünde  $\alpha$  nitelik örüntüsüne sahip  $i$  bireyin  $j$ . maddeye doğru yanıt verme olasılığı beklenen yanıtın 1 olması halinde  $1-s_j$  yanıtın 0 olması halinde  $g_j$  'e eşit olacaktır.

DINA model ayrıca G-DINA'nın kısıtlanmış hali olarak düşünüldüğünde  $j$  maddesi için  $g_j=\delta_{j0}$  ve  $1-s_j = \delta_{j0}+\delta_{j12...K^j}$  olarak bulunur. Yapılan işlem G-DINA formülündeki tüm temel etkileri ve çoklu etkileşim etkilerini 0'a eşitlemektir. Geriye kalan kesişim parametresi tahmin parametresine, toplu etkileşim parametresi ise 1-tahmin-kaydırma parametresine eşit olacaktır.

### **DINO (Deterministic-Input, Noisy-Or Gate) Model**

2006'da Templin ve Henson tarafından duyurulan DINO ile DINA model arasındaki temel fark DINA'daki "ve" ("and") kapısı DINO'da "veya" ("or") kapısına dönüşmektedir. DINO modelde maddenin gerektirdiği nitelikler birbirinden bağımsızdır, bireyin maddenin gerektirdiği niteliklerden herhangi birine sahip olması dahi bireyin maddeyi doğru cevaplama olasılığının en yüksek seviyede olması için yeterlidir.

$$\eta_{ij} = 1 - \prod_{k=1}^K (1 - \alpha_{ik})^{q_{jk}}$$

DINO'nun ideal yanıt formülü incelendiğinde  $i$  bireyinin  $j$ . maddede yoklanan niteliklerin hiç birine sahip olmaması durumunda beklenen yanıt 0 olacaktır. Diğer tüm durumlarda ise beklenen yanıt 1'dir. DINA modeldeki  $s$  ve  $g$  parametreleri DINO için de geçerli olup aynı şekilde hesaplanır. Ayrıca DINO modelin doğru yanıt verme olasılığı yine DINA model ile aynı şekilde hesaplanır.

DINO model de G-DINA'nın kısıtlanmış hali olarak düşünülebilir. Tahmin ve bilen bireyin doğru yanıt olma olasılığı  $j$  maddesi için  $g_j = \delta_{j0}$  ve  $1-s_j = \delta_{j0} + \delta_{jk}$  olarak bulunur. Görüldüğü üzere  $1-s_j$  parametresi madde kesişimine ek olarak DINA'da tüm niteliklere sahip olanların etkisi DINO'da ise herhangi bir niteliğe sahip olanlar etkisi toplanarak elde edilmektedir. GDINA üzerinde yapılan sadeleştirme  $k=1, \dots, K_j^*$ ,  $k'=1, \dots, K_j^*-1$ ,  $k'' > k', \dots, K_j^*$  olmak üzere  $\delta_{jk} = \delta_{jk'} \delta_{jk''} = \dots = (-1)^{K_j^*+1} \delta_{j12\dots K_j^*}$  olarak gösterilir.  $K_j^*$   $j$  maddesi için gerekli nitelikler olarak ifade edilebilir.

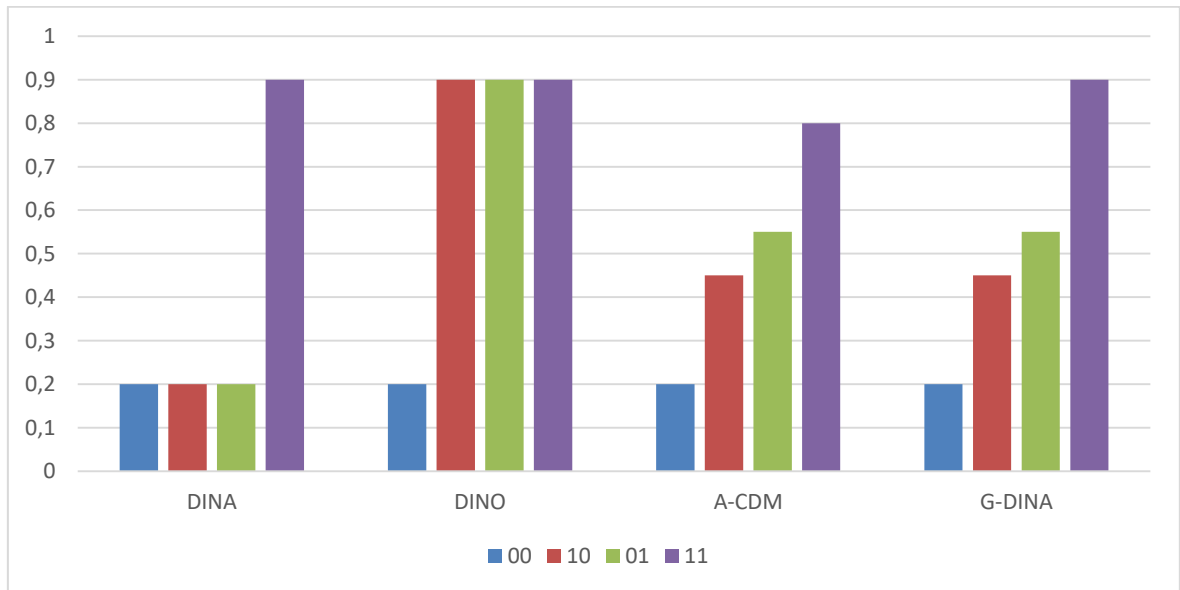
### A-CDM (Additive Cognitive Diagnosis Model)

A-CDM, G-DINA ile aynı anda tanıtılmıştır (de la Torre, 2011). Model doygun formda olmamakla beraber bağlayıcı veya bağlayıcı olmayan modeller sınıflamasına da girmemektedir. Modelin telafi edici yapıda olduğunu söylenebilir. Model temel olarak madde kesişimi üzerine eklenen her bir niteliğin temel etkisi üzerine kuruludur. G-DINA'da bulunan etkileşim etkileri A-CDM'de bulunmaz.

$$P(\alpha_{lj}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{lk}$$

Bireyde bulunan her bir nitelik maddeyi doğru yanıt olma olasılığını artıracaktır. Model her bir madde için aranan nitelik sayısından bir fazla parametre üretecektir.

İki nitelik  $q=(1,1)$  barındıran örnek bir madde için Şekil 1.7'deki nitelik örüntülerinin maddeyi doğru yanıt olma olasılığı incelenebilir.



Şekil 1.7. Modeller İçin Örnek Olasılıklar

Grafikte görüldüğü üzere tüm modeller için (0,0) grubundaki bireylerin soruyu doğru yanıtlama olasılığı aynıdır. A-CDM hariç diğer modeller için (1,1) grubundaki bireylerin soruyu doğru yanıtlama olasılığı aynıdır. A-CDM'de etkileşimlerin 0 olarak varsayılması nedeniyle bu örnek için doğru yanıtlama olasılığında küçük bir fark görülmektedir. Etkileşim etkisinin yönüne göre bu farkın yönü de değişebilir. Ayrıca bu örnek yukarıda bahsedilen etki büyüklükleri dikkate alındığında geçerlidir. Analiz sonuçları göre bu büyüklükler varsayımlar ve kısıtlamalar nedeniyle modeller arası küçük farklıklar gösterebilir.

### **Polytomous G-DINA (pG-DINA)**

BTM'lerin pek çoğunun çıkış noktası iki kategorili nitelik kodlaması (0-1) üzerinedir. Niteliklerin çok kategorili olması durumunda kullanılacak şekilde tasarlanan model General Diagnostic Modeldir (GDM; von Davier, & Yamamoto, 2004; von Davier, 2008). Ancak çoğu model çok kategorili nitelik örüntüsüne ve cevap örüntüsüne göre sonradan düzenlenmiştir. Pek çok analiz programında ve paketinde çok kategorili nitelik ve/veya yanıt örüntüsü kullanılarak model analizi yapılabilmektedir.

Lineer hiyerarşiye sahip nitelik örüntüsü çok kategorili nitelik örüntüsüne çevrilebilmektedir. Teorik açıdan aynı niteliğin farklı seviyelerinin değil de birbirine ön koşul olan farklı nitelikler için çok kategorili nitelik örüntüsüne dönüştürme işlemi doğru görünmeyebilir. Ancak model uygulamasında bu işlemin pek çok yararı vardır. Azalan nitelik sayısı parametre sayısını azaltacak yorumlamayı kolay hale getirecek ve hata oranını düşürecektir. pG-DINA da bu ihtiyaç sonucu ortaya çıkan bir modeldir. pG-DINA, fA-M ve G-DINA gibi 3 ayrı fonksiyon altında yazılabilir. Özdeş (identity), log ve logit fonksiyonları için de formül yazılabilir.

pG-DINA'nın parametre hesaplamasını  $a_1=2$ ,  $a_2=0$  ve  $a_3=1$  nitelik örüntüsündeki örnek bir madde üzerinde gösterildiği Tablo 1.2'de  $\alpha_{ij}^*$  düşmüş Q matrisi elemanını,  $\alpha_{ij}^{**}$  çökmüş Q matrisi elemanını temsil etmektedir.

**Tablo 1.2: Orijinal, Düşmüş ve Çökmüş Örtük Gruplar**

Orijinal $\alpha_{ij}$	Düşmüş $\alpha_{ij}^*$	Çökmüş $\alpha_{ij}^{**}$
(0,0,0), (0,1,0), (0,2,0) (1,0,0), (1,1,0), (1,2,0)	(0,0) (1,0)	(0,0)
(0,0,1), (0,1,1), (0,2,1) (1,0,1), (1,1,1), (1,2,1) (0,0,2), (0,1,2), (0,2,2) (1,0,2), (1,1,2), (1,2,2)	(0,1) (1,1) (0,2) (1,2)	(0,1)
(2,0,0), (2,1,0), (2,2,0)	(2,0)	(1,0)
(2,0,1), (2,1,1), (2,2,1) (2,0,2), (2,1,2), (2,2,2)	(2,1) (2,2)	(1,1)

Çok kategorili niteliğe sahip Q matrisi için beklenen olası yanıt örüntüleri sayısı  $M^K$  dir.  $M$  toplam kategori sayısı olup Tablo 1.2'deki örnek için üçtür.  $K$  toplam nitelik sayısını temsil etmekte ve örnekte üç olarak görülmektedir. Örneğimize göre beklenen olası örüntü sayısı  $3^3=27$  olup Orijinal  $\alpha_{ij}$  sütununda görülmektedir. Düşmüş  $\alpha_{ij}^*$  sütununda ise maddemiz 2. niteliği gerektirmediği için  $\alpha_{ij}$  vektöründe bulunan 2. elemanlar düşürülmüş ve düşmüş  $\alpha_{ij}^*$  elde edilmiştir. Buradaki örüntü sayımız  $M^{K^*}=3^2=9$  olarak yer almaktadır. Düşmüş  $\alpha_{ij}^*$ 'nin her bir elemanı çökmüş grup için maddenin nitelik vektöründe karşılık gelen elemanından büyük ise  $\alpha_{ij}^{**}$ 'de 1 değilse 0 olarak ele alınmıştır. Bu sayede  $2^{Kj^*}=2^2=4$  gruptan oluşan çökmüş  $\alpha_{ij}^{**}$  elde edilmiştir. Bu örtük sınıflar için G-DINA formülü kullanılarak doğru yanıt verme olasılığı hesaplanır. Formül gösterimindeki tek değişim  $\alpha_{ij}$  görülen yerler  $\alpha_{ij}^{**}$  ile temsil edilir. Ayrıca pG-DINA formülüne G-DINA'da olduğu gibi ilgili kısıtlamalar uygulanırsa kısıtlı modeller çok kategorili nitelikleri için parametreler elde edilmiş olur (Chen & de la Torre, 2013).

pG-DINA'yı teknik açıdan biraz incelemek gerekirse marginal maximum likelihood (MML) expectation-maximization (EM) uygulaması ile uyumun en üst olduğu noktada madde parametreleri tahmin edilir.  $N$  örneklem büyüklüğü,  $L$  toplam nitelik örüntüsü sayısı,  $L(X_i|\alpha_1)=\prod_{j=1}^L P(\alpha_{ij}^{**})^{X_{ij}} [1-P(\alpha_{ij}^{**})]^{(1-X_{ij})}$  olmak üzere

$$L(X)=\prod_{i=1}^N \sum_{l=1}^L L(X_i|\alpha_l)p(\alpha_l),$$

ile hesaplanır.  $\alpha_1$ 'in önsel olasılığı olan  $p(\alpha_1)$ ,  $p(\alpha_1|X)$  sonsal olasılıkla dögüsel işlemde güncellenerek hesaplanır.  $p(\alpha_l|X_i)_{t=0}=p(\alpha_l)$ 'den başlayan dögü aşağıdaki formül ile ilerler.

$$p(\alpha_l|X_i)_t = \frac{L(X_i|\alpha_l)p(\alpha_l|X_i)_{t-1}}{\sum_{l=1}^L L(X_i|\alpha_l)p(\alpha_l|X_i)_{t-1}}$$

MML işleminde  $\alpha_{ij}^{**}$  çökmüş nitelik örüntüsünün  $j$  maddesine doğru yanıt verme olasılığı,

$$\hat{P}(\alpha_{ij}^{**}) = \frac{R_{\alpha_{ij}^{**}}}{N_{\alpha_{ij}^{**}}}$$

$\alpha_{ij}^{**}$  grubu için  $R_{\alpha_{ij}^{**}}$  beklenen doğru yapan birey sayısı ve  $N_{\alpha_{ij}^{**}}$  gruptaki beklenen toplam birey sayısı olarak ifade edilir.

$$I(p_j) = \left\{ \frac{\partial^2 L(X)}{\partial P(\alpha_{ij}^{**}) \partial \hat{P}(\alpha_{ij}^{**})} \right\}$$

Tahminin standart hatası bilgi matrisinde elde edilebilir. Bireylerin her bir nitelik için marjinal sonsal olasılığı en yüksek olan seviyeye sınıflandırılır (Chen & de la Torre, 2013).

### fA-M (Fully Additive Models)

Tez kapsamında önerilen fA-M adından da anlaşılacağı üzere A-CDM tabanlıdır. Yeni bir model kullanılmasındaki en önemli faktör mevcut BTM'lerin ÇBMTK formüllerine çok yakın olmamasıdır. Önerilen yeni formülle bu iki akım birbirine daha yakın hale getirilmeye çalışılmıştır. FA-M'in çok kategorili niteliklere uygun A-CDM'ye göre farkı maddenin gerektirdiği seviyeden daha yüksek nitelik seviyesine sahip bireylerin doğru yanıt olma olasılığının artmasıdır. Model bu yönüyle MTK'ye yaklaşmıştır. Daha fazla bilgi, yetenek, beceriye sahip olan bireyin daha yüksek doğru yanıt olma olasılığına sahip olması ayrıca tüm MTK modelleri için de geçerlidir. Ancak diğer hiçbir BTM'de bu durum söz konusu değildir. Model mevcut haliyle sadece çok kategorili niteliklere uygundur.

$$q_{jk}^* = \begin{cases} 0 & \text{eğer } q_{jk} = 0 \\ 1 & \text{değilse} \end{cases}$$

$$\alpha_{lk}^{**} = \begin{cases} 0 & \text{eğer } \alpha_{lk} < q_{jk} \\ 1 & \text{değilse} \end{cases}$$

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j} q_{jk}^* \sum_{m=1}^M \delta_{jm}^k \alpha_{lk}^{**}$$

Formüle göre  $j$  maddesi için hesaplanacak  $\alpha_l$  örtük sınıfındaki grubun doğru yanıt olma olasılığı için öncelikle Q matrisindeki  $j$  maddesine ait vektördeki 0 olmayan elemanlar  $q_{jk}^*$  için 1 olarak ele alınır. Bu sayede maddede yoklanmayan nitelikler için etki puanı alınmasının önüne geçilmiş olur.  $\alpha_{lk}^*$ 'nın  $\alpha_{lk}$  grubun ilgili niteliğinin seviyesi orijinal Q matrisi  $q_{jk}$ 'deki seviyesinden küçük ise 0, eşit veya büyükse 1 değeri alması sağlanır. Bu eşitlikle grup seviyesinin erişemediği nokta için etki puanını engeller. Model pG-DINA ile kıyaslandığında, pG-DINA her çökmüş grup için ürettiği parametre ile doygun yapıya sahip iken, fA-M düşmüş her grup için parametre hesaplayacaktır.

fA-M, pG-DINA ve G-DINA gibi 3 ayrı fonksiyon altında yazılabilir. Logit fonksiyonu altında:

$$\text{logit}[P(\alpha_{ij}^*)] = \lambda_{j0} + \sum_{k=1}^{K_j} q_{jk}^* \sum_{m=1}^M \lambda_{jm}^k \alpha_{lk}^{**}$$

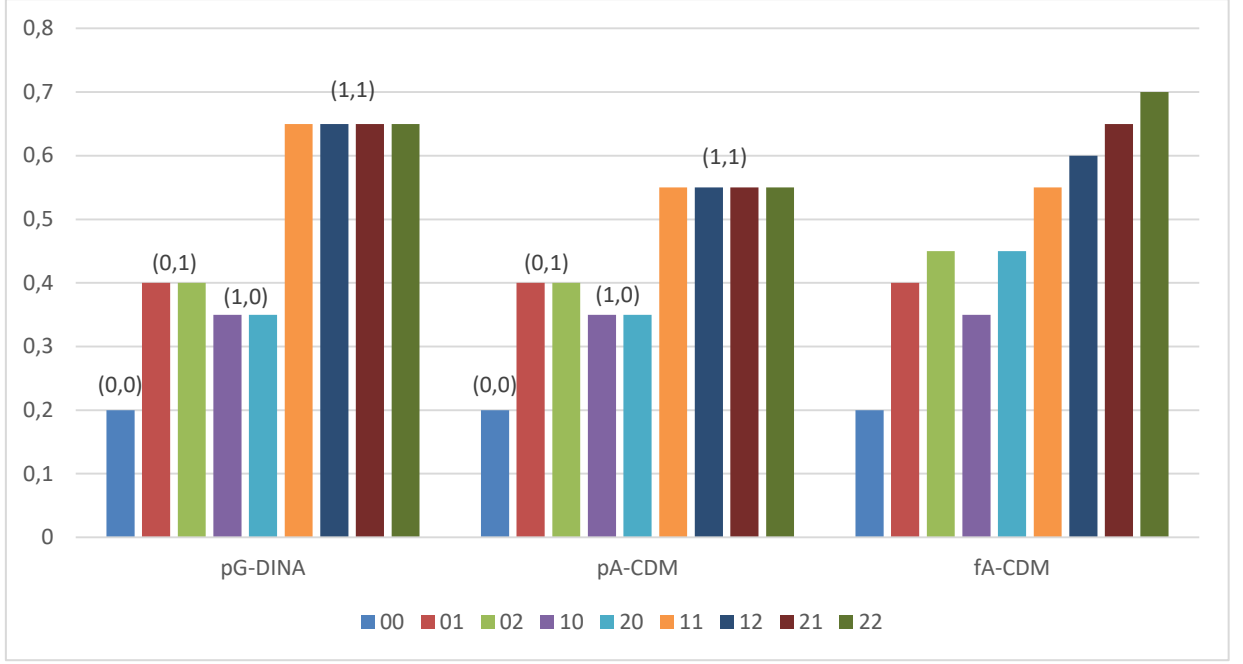
logit fonksiyonunda fA-M'in her ikisi modelinde aynı fonksiyonu paylaşması nedeniyle telafi edici ÇBMTK'ya yakın olması beklenmektedir.

Log fonksiyonu altında:

$$\log[P(\alpha_{ij}^*)] = v_{j0} + \sum_{k=1}^{K_j} q_{jk}^* \sum_{m=1}^M v_{jm}^k \alpha_{lk}^{**}$$

Log fonksiyonu ile gerçekleştirilen fA-M'in ise log toplamlarının çarpım etkisi nedeniyle, olasılıkların çarpımı ile hesaplanan telafi edici olmayan ÇBMTK'ye yakın olması beklenmektedir.

Örnek madde üzerinden örtük gruplardaki bireylerin maddeyi doğru yanıt olma olasılıkları modellere göre Şekil 1.8'deki gibi gösterilebilir.



**Şekil 1.8. Çok Kategorili Nitelik İçin Model Örnek Olasılıkları**

$\delta_0 =$  Temel etki= ,2

$\delta_1^1 =$  İlk niteliğin birinci düzey temel etkisi = ,15

$\delta_2^1 =$  İlk niteliğin ikinci düzey temel etkisi = ,10

$\delta_1^2 =$  İkinci niteliğin birinci düzey temel etkisi= ,20

$\delta_2^2 =$  İkinci niteliğin ikinci düzey temel etkisi= ,05

$\delta_{12} =$  İlk ve ikinci niteliğin etkileşim etkisi= ,10

Örnek grafik nitelikler için en yüksek düzeyin 2 olduğu Q matrisinde (1,1) seviyesinde nitelik gerektiren madde için verilmiştir. FA-M, pG-DINA ve çok kategorili A-CDM'den (pA-CDM; Chen & de la Torre, 2013) farklı olarak  $\delta_1^2$  ve  $\delta_2^2$  etkilerini de kullanmaktadır. Bunun nedeni pG-DINA ve pA-CDM'nin maddenin gerektirdiği nitelik düzeyinden fazla seviyedeki niteliğe sahip bireylere daha fazla parametre vermemesinden kaynaklanmaktadır. pG-DINA ise pA-CDM ve fA-M farklı olarak  $\delta_{12}$  etkileşim etkisine sahiptir.

Grafikte göze çarpan ilk durumlardan biri pG-DINA ve pA-CDM'de örtük sınıflar gruplanmış halde görünmektedir. Bu gruplanmalar sırayla (0,0), (0,1), (1,0) ve (1,1) çökmüş grupları içindir. Buradan da anlaşılacağı üzere pG-DINA ve pA-CDM çökmüş gruplar temelinde parametre ve bilgi üretmektedir. FA-M'a baktığımızda ise



örtük grupların birbirinden farklı olasılıkları görülmektedir. Bu da modelin düşmüş grup üzerinden parametre ürettiğini göstermektedir. Bu sayede farklı nitelik örüntüsündeki grupları birleştirmek zorunda kalınmamaktadır. Ayrıca bilgi kaybının da önüne geçilmektedir. FA-M'ın MTK'ye benzer yönü grafikte de görülmektedir. MTK'de bulunan yetenek arttıkça doğru yanıt verme olasılığı artmasını gerektiren formülü, fA-M'a bireyin niteliği arttıkça doğru yanıt verme olasılığı artar olarak yansıtılmıştır.

## 2. İLGİLİ ARAŞTIRMALAR

Bu bölümde araştırma konusuyla ilgili olan ÇBMTK ve BTM çalışmalarına ve bu modellerin uyarlama çalışmalarına değinilmiştir.

### 2.1. ÇBMTK ile İlgili Çalışmalar

Bu bölümde çalışma konumuzun ÇBMTK kısmına yakınlığı görülen araştırmalardan bahsedilecektir.

Spray, Dawey, Reckase, Ackerman ve Carlson (1990) çok boyutlu 2 parametrelili lojistik modelin telafi edici ve telafi edici olmayan modellerini karşılaştırmışlardır. İki boyutlu olarak tasarlanan simülasyon çalışmasında her model için madde parametreleri eşitlenmesi veya eşleştirilmesi durumunda madde yanıtı ve test puanı düzeyinde modeller arasında farklılığın tespiti üzerinde durmuşlardır. Bu amaçla 2000 birey için 0, 0,25, 0,50 ve 0,75 korelasyona sahip yetenek parametreleri üretilmiş ve 20 madde için verilen yanıtlama olasılık farkları modeller için en aza indirgeyecek madde parametreleri üretilmiştir. Burada yapılan işlem yetenek ve yanıtlama olasılığının sabitlenip, madde güçlük ve ayırt ediciliklerin modeller arasında serbest hale getirilerek modelleri daha iyi kıyaslanabilir hale getirmektir. Araştırma sonucunda toplam test puanında modeller arası farklılık madde yanıtlama olasılıklarının modeller için aynı olması nedeniyle zor gözlenebilirken, madde seviyesinde farklılık çok daha belirgindir. Madde veya yetenek parametrelerine ait öncül bilgiler olmaksızın modeller arası farkın ayırt edilemeyeceği, hatta özellikle gerçek test uygulamalarında beklendiği üzere yetenekler arası yüksek korelasyonun bulunduğu durumlarda öncül bilgilerin dahi model ayırımında yeterli olmayacağı sonucuna varılmıştır.

Gerçek veri seti üzerinden Bolt ve Lall (2003) tarafından yapılan çalışmada ise tek parametrelili telafi edici ve telafi edici olmayan modeller Markov Chain Monte Carlo tekniği ile karşılaştırılmıştır. Ayırt edicilik parametresi bire eşitlenmiş telafi edici M2PL modeli ile (McKinley ve Reckase, 1982) telafi edici olmayan Multicomponent latent trait model (MLTM; Whitely, 1980) kıyaslanmıştır. M2PL analizinde madde parametreleri kabul edilebilir oranda doğru tahmin edilirken, MLTM'de parametrenin yeniden elde edilme oranı ise örneklem büyüklüğünden, toplam madde sayısından ve yetenekler arası korelasyondan gözle görünür şekilde etkilendiği belirtilmiştir. MLTM başarılı olduğu yüksek korelasyon ve büyük örneklem durumlarında dahi

M2PL'den parametrenin yeniden elde edilmesinde daha düşük performans göstermiştir. Çalışmanın uygulandığı testin telafi edici modele daha fazla uyum sağladığı görülmüştür. Ayrıca bu test temel alınarak her iki model için veri üretilmiş telafi edici modelde üretilen veride telafi edici olmayan modelin uyumu düşükken, telafi edici olmayan model için üretilen verinin telafi edici modele olan uyumunun MLTM'ye yakın düzeyde olduğu görülmüştür.

Adams, Wilson ve Wang (1997) Multidimensional Random Coefficients Multinomial Logit Model (MRCMLM) ile pek çok tek parametrelili çok boyutlu modeli kapsayan genel bir model ortaya koymuşlardır. Modelin özelliklerini ortaya koymak için yaptıkları çalışmada boyut sayısı, madde yapısı, yanıt kategorisi, madde sayısı, örneklem büyüklüğü ve boyutlar arası korelasyon koşullarını veri üretiminde kullanmışlardır. Parametrik olmayan ve normal analiz sonuçlarına göre madde parametrelerinin tahminindeki yanlılık düşük düzeyde görünmektedir. Örneklem dağılımı normal olduğunda parametre dağılım tahmini iyileşmektedir. Ayrıca karmaşık yapıdaki madde tipinde görülen parametre tahmin yanlılığı basit yapı için gözlenen yanlılıktan fazladır. Bu sonuç çok kategorili yanıt örüntüsüne sahip veriler için çok daha belirgindir. Boyutlar arası korelasyonun ise parametre yanlılığının farklılaşmasında önemli bir faktör olmadığı görülmüştür.

Maddelerin basit yapıda olması yani madde arası boyutluluk göstermesi durumunda çok boyutlu testteki her boyut için tek boyutlu MTK uygulaması akla gelebilir. De la Torre ve Patz (2005) bu çalışmayı 2 ve 5 boyut, 10, 30 ve 50 madde, 0, 0,40, 0,70, 0,90 boyutlar arası korelasyon ve basit yapıdaki maddeler üzerinden gerçekleştirmişlerdir. Çalışmada birey parametresinin tahmini üzerinde durulmuştur. Tek boyutlu ve çok boyutlu MTK uygulamaları beklendiği üzere korelasyonun 0 olduğu koşullarda aynı sonuçları verirken, korelasyonun arttıkça MTK karşısında ÇBMTK'nin etkinliği daha görünür hale gelmiştir. Aynı durum az sayıda madde barındıran testler için de geçerlidir. Araştırmacılar tek seferde ölçülen birden fazla yeteneğin birbirinden bağımsız olmasının pek mümkün olmadığını aralarında belirgin bir ilişki olabileceğini hatta birbirleri arasında korelasyon olan yeteneklere ilişkin testlerin ayrı ayrı yapılması durumunda dahi daha güvenilir sonuçlarına elde edilmesi amacıyla ÇBMTK analizlerinin kullanılmasını önermektedirler.

Telafi edici olmayan modele yönelik Chalmers ve Flora (2014) tarafından yapılan veri üretimine dayalı çalışmada 2 ve 3 boyut, 2 ve 3 PL modeller, 1000, 2500 ve 4000 örneklem büyüklüğü, boyut başına 5, 10 ve 15 madde, 0, 0,2, 0,4, 0,6, 0,8 boyutlar arası korelasyon katsayıları koşul olarak kullanılmıştır. Parametre kestiriminde Metropolis-Hasting Robbins-Monro (MHRS) karma metodunun incelendiği çalışma sonuçlarına göre genel olarak parametrelerin kabul edilebilir düzeyde yeniden elde edilmesi için ciddi çaba gerektirmektedir. Örneklem büyüklüğü ve boyuta yük veren madde sayısının artırılması daha kesin tahminler için gerekli görülmüştür. Örneklem büyüklüğünün ancak on binleri bulması durumunda yeterli güven aralığında tahmin edilebileceği belirtilmiştir. Boyut sayısının artırılması tahmin edilen değerlerin gerçek değerlerden uzaklaşmasına neden olmuş bunun önüne geçilebilmesi için daha geniş örneklem büyüklüğüne ihtiyacı ifade edilmiş, ayrıca şans parametresinin de yer aldığı modelde ayırt edicilik ve konum parametrelerinin gerçek düzeyden çok daha uzak tahmin edildiği görülmüştür. Boyutlar arası korelasyonun parametre tahmin doğruluğunda olumsuz bir etken olmuştur.

## **2.2. BTM ile İlgili Çalışmalar**

Bu bölümde çalışmanın BTM kısmına yakınlığı görülen araştırmalarından bahsedilecektir.

Sinharay, Puhan ve Haberman (2009) bilişsel tanı skorlarının kullanımına ilişkin hazırladıkları raporda üniversite ve enstitü gibi akademik kuruluşların kendi eğitimlerini ve yoğunlaştıkları alanları değerlendirmek adına mezunlarının performans profillerini ortaya çıkarmak istediklerini belirtmişlerdir (Haladyna & Kramer, 2004). Aynı zamanda Öğrencilerin Ne Bildiğini Bilmek (Pellegrino ve diğerleri, 2001) raporundaki değerlendirmenin amacı öğrencinin belirli bilgi, beceri ve yetenekleri hakkında bilgi sağlamak olmasına vurgu yaparak tanısız skorların bu tür bilgiyi sağlayacağını belirtmiştir. Alt alanlara ait skorların kullanımını göstermek adına gerçekleştirdikleri simülasyon çalışmasında alt puan sayısı, alt görevlerin uzunluğu, alt görevler arası korelasyonlar ve örneklem büyüklüğünü koşul olarak kullanmışlardır. Sonuçlarda BTM'lerin tanı puanı elde etmede kullanılması halinde kullanılan model parametrelerin güvenilir şekilde elde edilmesi gerektiğini belirtmişlerdir. Bunun içinde yeterli sayıda madde bulunması gerektiğini

vurgulamışlardır. Ayrıca basit düzeydeki BTM'lerin bilişsel profillerinin, alt görev puanına göre bilgi olarak daha fazla katkıda bulunacağını ifa etmişlerdir.

Geniş kapsamda uygulanan sınavlara BTM'lerin uygulanmasına yönelik Chen ve de la Torre (2014) tarafından yapılan çalışmada 2000 yılında yapılan Uluslararası Öğrenci Değerlendirme Programı (Programme for International Student Assessment – PISA) Okuma bölümü BTM'lerle değerlendirilmiştir. Birleşik Krallık verilerinin uygulama ve Amerika Birleşik Devletleri verilerinin doğrulama çalışması için kullanıldığı araştırmada 2814 öğrencinin 26 maddeye verdiği yanıtlar değerlendirmeye alınmıştır. Q matrisi geliştirme çalışmasında 8 nitelik analizler sonucunda 6'ya indirilerek son haline getirilmiştir. DINA, DINO, A-CDM, Lineer Logistic Model (LLM; Maris, 1999) ve Reduced Reparametrized Unified Model (R-RUM; DiBello, Roussos, & Stout, 2007; Hartz, 2002) modellerinin denendiği çalışma sonucunda eklemeli modellerden LLM en iyi mutlak ve bağıl uyumu gösterirken diğer bir eklemeli R-RUM ikinci en iyi uyuma sahip olduğu görülmüştür. Yaygın olarak kullanılan DINA ve DINO ise sırasıyla bu test için en kötü uyumu göstermişlerdir. Doğrulama verisi ile gerçekleştirilen analizde de aynı sonuçlar edinilmiştir. Ayrıca daha fazla sayıda madde kullanımı ile birey profillerinin daha doğru şekilde tahmin edileceğini ve maddelere sonradan Q matrisi hazırlanmaktansa testin orijinal olarak BTM tabanlı hazırlanması halinde daha ideal sonuçlara ulaşılabileceğini bildirmişlerdir.

Çok kategorili Q matrisi uygulamasına ilişkin rastlanılan ilk çalışma 2004 yılında Karelitz tarafından gerçekleştirilmiştir. Çalışmanın simülasyon kısmında nitelik sayısı 3-5 ve nitelik kategorisi 4 ve 5 olarak kullanılmıştır. Her veri seti için 2000 yanıt örüntüsü örneklem büyüklüğü olarak belirlenmiş ve MCMC algoritması kullanılmıştır. BTM analizinde DINA modelin kullanıldığı çalışmada niteliklerin bağımsız olması halinde nitelik sayısının artması tahmin doğruluğunu etkilemezken, kategori sayısının artması tahmin doğruluğunu düşürmüştür. Çarpık dağılım kullanıldığında da yetenek profillerinin doğru tahmin oranı azaldığı görülmüştür. Uygulamalı analizde ise 40 madde 3 nitelik ve 4 kategori barındıran test 200 öğrenciye uygulanmıştır. Bireyler test uygulaması öncesi 10 gruba ayrılmış, 10 nitelik örüntüsüne göre eğitime tabi tutulmuştur. Böylece bireylerle nitelik örüntüleri ilişkilendirilmiştir. Test uygulaması sonucunda elde edilen parametre doğrulama oranları simülasyon sonuçlarına göre çok düşük çıkmıştır.

PG-DINA modelinin ortaya konduğu çalışmanın (Chen & de la Torre, 2013) simülasyon veri ile gerçekleştirilen kısmında 3 seviyeli 5 nitelik içeren 15 ve 30 maddelik Q matrislerine için 500, 1000 ve 2000 kişilik veriler üretilmiştir. Çok kategorili A-CDM ve DINA modellerine göre üretilen verilerde üretilen madde parametreleri standart hatalar eşliğinde değerlendirildiğinde iyi bir düzeyde elde edildiği görülmüştür. Çok kategoriye sahip Q matrisi pG-DINA ile ve aynı matrisi ikili kategoriye dönüştürülmüş hali G-DINA ile analizi sonucunda pG-DINA'nın tüm özelliklerde ve vektör sonuçlarında birey parametreleri çok daha iyi bir oranda doğru tahmin ettiği görülmüştür. Gerçek veri analiz sonuçlarında ise pG-DINA'nın G-DINA'dan daha iyi model veri uyumuna sahip olduğu ortaya çıkmıştır.

Templin ve Bradshaw (2014) yaptıkları çalışmada daha önce Buck ve Tatsuoka (1998) ilgili test için hazırlanan Q-matrisini uygulaması sonucunda ortaya çıkan profil örüntülerinden yola çıkarak yeni bir model sunmuşlardır. 28 madde ve 3 nitelik üzerinden yapılan Log-linear Cognitive Diagnosis Model (LCDM; Henson, Templin, & Willse, 2009) ile analiz edilen Q-matrisinde lineer hiyerarşi bulunmamasına rağmen birey profillerinin %95'i lineer hiyerarşiye uygun şekilde (0,0,0), (0,0,1), (0,1,1) ve (1,1,1) gruplarında görülmesi, araştırmacılar tarafından 3. niteliğin 2. niteliğe, 2. niteliğin ise ilk niteliğe ön koşul olduğu şeklinde yorumlanmıştır. Analiz araştırmacıların geliştirdikleri Hierarchical Diagnostic Classification Model (HDCM) ve aynı Q matrisi ile gerçekleştirildiğinde tüm bireyler lineer hiyerarşiye uygun şekilde yukarıda sıralanan 4 profil örüntüsüne yerleşmişlerdir. HDCM etkinliğini araştırmak için yapılan simülasyon çalışma sonucunda ise modelin hiyerarşik nitelik yapısını tespit etmede kullanılabileceği ve DINA ve DINO gibi çok kullanılan iki BTM ile karşılaştırıldığında daha iyi sonuç verdiği belirtilmiştir. Hiyerarşik yapılara hitap eden diğer modeller olan RSM ve AHM Attribute Hierarchy Method (AHM; Leighton, Gierl, & Hunka, 2004) ile karşılaştırıldığında bu modellerin hiyerarşik yapıyı tespit etmek için kullanılamamasının HDCM'nin artısı olduğu ifade edilmiştir.

### **2.3. ÇBMTK ve BTM Uyarlaması ile İlgili Çalışmalar**

Bu bölümde çalışma konumuzun geneline yakınlığı görülen araştırmalarından bahsedilecektir.

BTM ile MTK karşılıklı uyarlanmasını amaçlayan çalışma de la Torre ve Karelitz (2009) tarafından DINA model ile tek boyutlu MTK arasında gerçekleştirilmiştir. Araştırma simülasyona dayalıdır. -2, -1, 0, 1 ve 2 düzeylerinde güçlük parametresine

sahip 5 madde tipi beşer kez tekrarlanarak 25 maddelik MTK testi düşük, orta ve yüksek düzeyde ayırt ediciliğe sahip 5000 bireylik 3 ayrı veri setinin üretiminde kullanılmıştır. Birey parametreleri ortalaması 0 standart sapması 1 olan normal dağılıma sahiptir. MTK verisinin madde karakteristik eğrisinden yararlanarak her madde için güçlük parametresinin altında ve üstünde olan birey sayısı hesaplanarak DINA model için  $g$  ve  $s$  parametreleri hesaplanmıştır. Madde zorlukları BTM için lineer hiyerarşiye sahip  $(1,0,0,0,0)$ ,  $(1,1,0,0,0)$ ,  $(1,1,1,0,0)$ ,  $(1,1,1,1,0)$  ve  $(1,1,1,1,1)$  5 maddeye denk olarak görülmüştür. Bu maddeler 5'er kez kullanılarak 25 maddelik test ve Q matrisi BTM için hazırlanmıştır. Madde parametreleri olan  $g$  ve  $s$  parametreleri de kullanılarak BTM verisi üretilmiştir. Her iki veri seti de 100 tekrar ile üretilip analiz edilmiştir. Sonuçlarda düşük tanı gücüne sahip testlerde uyum indeksleri birbirine yaklaşırsa da beklendiği üzere sürekli olarak verinin üretildiği model daha iyi uyum göstermiştir. Ayrıca madde parametreleri de doğru modelde daha iyi tahmin edilmiştir. Birey profil tahminlerine bakıldığında BTM verisi çok daha iyi tahmin ederken MTK verisi bireylerin ortalama %60'ını gerçek değerden farklı tahmin etmiştir. MTK madde parametrelerinin tahminine bakıldığında MTK verisi için 2PL tahminleri gayet iyi iken BTM verisinden elde edilen hata düzeyleri çok yüksektir. Birey yetenek parametresi güçlük parametreleri kesme puanı olarak kullanılarak 6 kategorili (0-5) kesikli hale getirilerek modeller arasında karşılaştırmaya da gidilmiştir. MTK verisine 2PL uygulanarak elde edilen yetenek ile gerçek yetenek ortalama %74 oranında doğru tahmin edilmiştir. Bu oran BTM verisine 2PL uygulandığında ise %68'lerde olduğu görülmüştür. BTM verisinde uç değerlerdeki bireyleri tahmin daha sorunlu iken, MTK verisinde orta değerlerdeki bireyler daha sorunlu olarak tahmin edilmektedir. Genel olarak kaliteli maddeler kullanılarak uyarılama çalışmasının daha verimli gerçekleştirileceği vurgulanmıştır.

Diğer bir çalışma Wang tarafından 2009 'da doktora tezi olarak gerçekleştirilmiştir. Çalışmada BTM'lerden R-RUM modeli ile ÇBMTK modellerinden telafi edici model karşılıklı olarak uyarlanmıştır. Modellerin test kalitesini, madde ve birey parametrelerini ne düzeyde birbirine yakın kestirdiği araştırılmıştır. Analizlerin Monte Carlo Markov Chain yöntemi ile yapıldığı çalışmanın veri üretiminde BTM için basit ve karmaşık Q matrisi, özellikler arası korelasyon katsayıları (0,2, 0,5 ve 0,9) ve madde kalitesi zayıf, normal ve kaliteli olarak kontrol altına alınmıştır. Bu çalışmada R-RUM madde parametreleri ÇBMTK madde parametreleri gibi

yorumlanabileceği düşüncesiyle herhangi bir dönüştürme yapılmamıştır. Test uzunluğu 40, örneklem büyüklüğü ise 2000 olarak belirlendiği çalışmada 10 tekrar ile gerçekleştirilmiştir. İşlem aşamalarına bakıldığında, Q matrisi ve çok değişkenli normal dağılıma dayanan yetenek parametreleri üretiminin ardından, yetenek parametreleri 0 kesme puanına göre 0-1 dönüştürülmüş ve madde parametrelerinin üretimi ardından R-RUM veri seti elde edilmiştir. Bu veri seti R-RUM ve 2PL ÇBMTK'de tahmin edilmiştir. 2PL ÇBMTK tahmini madde parametrelerine göre ÇBMTK verisi üretilmiş ve R-RUM ile analiz edilmiştir. Birey parametreleri ile ilgilenen kısımda ise her model kendi parametresi ile elde edildikten sonra her iki modelde de analiz edilmiştir. Korelasyonun artması birey parametrelerini çok az düşürse de genel olarak uyum hem basit hem de karmaşık yapı için iyi düzeyde bulunmuştur. Madde kalitesine göre uygun bir şekilde uyum oranının farklılaştığı sonuçlar arasındadır. Madde parametreleri açısından modellerin madde kalitesinin tespitinde farklı sonuçlara ulaştığı görülmüştür. BTM verisinde birbiri ile daha uyumlu sonuçlar elde edildiği, çoğu basit yapı matriste yüksek düzey uyum olduğu, ÇBMTK verisinde daha yüksek düzeyde hata görüldüğü ve ÇBMTK verisinin BTM verisi kadar başarılı sınıflama yapmadığı bildirilmiştir.

Bradshaw ve Templin (2014) önerdikleri *Scaling Individuals and Classifying Misconceptions* adını verdikleri model ile genel olarak MTK'den gelen birey parametresi ile LCDM'den yararlanarak BTM'nin verdiği birey nitelik profil örüntüsünü birlikte değerlendirip yeteneği ölçeklemek ve profil örüntüsünden yola çıkarak kavram yanlışlarını göstermeyi amaçlamıştır. Modelin etkinliğini test etmek için 3 ve 6 nitelik, 3000 ve 10000 örneklem büyüklüğü, 30 ve 60 madde, düşük ve yüksek kalitedeki test koşulları çaprazlanarak simülasyon çalışması yapılmıştır. Bulgularda nitelik sayısı artışında tahmin hatası artmakta iken, örneklem büyüklüğü ve madde sayısı artışında hatalar azalmaktadır. Güvenirlik tahmininde de benzer sonuçlara ulaşılmıştır. Gerçek veriye dayalı çalışmada ise fizik maddeleri üzerinden kavram yanlışlarının düzeyleri ve nedenleri üzerinde durulmuştur.

ÇBMTK'deki sürekli birey parametresi ile BTM'deki kesikli birey parametresini birleştirerek yeni bir model önerme çalışması da Hong, Wang, Lim ve Douglas diğerleri tarafından yapılmıştır (2014). Çalışmada kesikli yapıda ve sürekli yapıda sonuç üreten 2 farklı model önerilmiştir. Telafi edici olmayan ÇBMTK modeli ile bağlayıcı olmayan DINA modelini içeren yeni modelin etkinliğini MCMC algoritması



ile 20 ve 40 madde sayısı 500 ve 2000 örneklem büyüklüğü koşulları çaprazlanarak denemişlerdir. Tek boyut üzerinde yapılan çalışmada özellikle kısa testlerde yetenek parametresinin hatası yüksek iken, güçlük ve ayırt edicilik parametresi tahmin hatası test uzunluğu arttıkça azaldığı görülmüştür. Hatalar niteliğin iki, boyutun tek olduğu durumda biraz daha artmıştır. Doğru sınıflama yüzdesi ise kısa testlerde %82, uzun testlerde %92 düzeylerinde olduğu görülmüştür. Modelin sürekli yetenek parametresi üreten kısmına bakıldığında hataların madde sayısı ile ters orantılı şekilde azaldığı, boyut sayısı ile doğru orantılı şekilde arttığı görülmüştür. Ancak tahmini hataların düşük olarak gerçekleştiği görülmüştür. Model ayrıca BTM uygulamalarında sıkça kullanılan Tatsuoka'nın (1990) uygulamasına ait kesirlerde çıkarma testi verisine de uygulanmıştır. Testte yer alan 8 niteliğin ayrı ayrı ve değişik şekilde birlikte ele alındığı bölümde en iyi uyum sürekli yapıda parametre üreten testi toplu olarak değerlendiren araştırmacıların önerdiği modelde görülmüştür. Önerilen sürekli yapıda çıktı veren modelin testte özellikle 3'den fazla boyut olması durumunda ve ÇBMTK'nin model uyumu sağlayamaması durumunda iyi bir alternatif olduğu üzerinde durulmuştur.

ÇBMTK'ye dayalı testin BTM uygulaması üzerine çalışma yapan Bradshaw, Izsák, Templin ve Jacobson (2014) öğretmenlerin rasyonel sayılara ilişkin bilgilerini ölçen 4 nitelik içeren bir test hazırlamışlardır. Çalışmanın temel amacı çok boyutlu yapılara ilişkin BTM'ye dayalı testin nasıl hazırlanacağı ve analiz edilmesi gerektiğini göstermek olarak görebiliriz. Her niteliğin bir boyuta karşılık geldiğini kabul edilen çalışmada boyutlar arası korelasyonların 0,70 düzeylerinde olmasına rağmen test çok boyutlu olarak kabul edilmiştir. Sadece bir boyut yapı olarak basit, diğer üçü karmaşıktır. Analiz sonuçları ile alan eğitimcilerine matematik öğretmenlerinde bulunan eksikler gösterilmek istenmiştir. Ayrıca BTM ile elde edilen güvenilirlik düzeyinin ÇBMTK'de çok daha fazla soru ile elde edilebileceğinden testin altyapısının BTM'ye dayalı olmasının avantajı üzerinde durulmuştur.

Çok boyutlu test geliştirip, BTM ve ÇBMTK'ye yakın faktör analizine göre birey parametrelerini karşılaştırmayı amaçlayan Kunina-Habenicht, Rupp ve Wilhelm'in (2009) çalışması matematik alanındaki dört temel işlem üzerinden gerçekleştirilmiştir. BTM analizlerinde genel modellerden olan GDM'den elde edilen sonuçlar, çok boyutlu doğrulayıcı faktör analiz (DFA) sonuçları ile karşılaştırılmıştır. 3. ve 4. sınıftaki 464 öğrenci üzerinde 87 madde ile uygulanan test doğrulayıcı faktör

analizinde deęişik modellerde test edilmiş ve en iyi modelin tüm boyutlar arasında korelasyon olan model olduęu görülmüştür. DFA ve GDM birey çıktılarının benzer korelasyonlara sahip olduęu görülmüştür. Tüm niteliklere sahip olan ve hiçbirine sahip olmayan profiller her iki sınıf için de en yüksek orana sahiptir. Bekleneceęi üzere 4. sınıfı için tüm niteliklere sahip olma oranı daha yüksek gerçekleşmiştir. DFA ve GDM birey sonuçlarının boyutlar için korelasyonu yaklaşık 0,85 düzeylerinde gerçekleşmiştir.

Lee, Park ve Taylan (2011) Uluslararası Matematik ve Fen ve Alanlarında Yönelim (TIMSS) testinin 25 maddelik matematik bölümünü BTM'de DINA modele uygulamışlardır. Testte testi hazırlayanlarca önceden belirlenmiş 15 nitelięe göre Q matris uygulamaya konulmuştur. Örneklem olarak ABD tamamı 564 birey ile ve 132 ve 127 bireylik iki ayrı eyaleti ayrı ayrı ele alınmıştır. Büyük örnekleme 3PL, küçük örneklemlerde ise DINA model daha iyi uyum gösterdięi görülmüştür. Ancak DINA model madde parametrelerine baktığımızda özellikle g parametrelerinde çok yüksek parametreler elde edildięi görülmüştür. Bu parametreler eyalet bazındaki analizler için çok daha yüksektir. Kaydırma (s) parametrelerinde de kimi maddelerde 0,60'ları bulan deęerlere ulaşılmıştır. 1-s-g ayırt edicilik parametresi olarak görülebilecek parametreler ise ortalama 0,45 düzeylerinde gerçekleşmiştir. Niteliklerin bireylerde bulunma durumlarından yararlanarak maddelere doęru yanıt vermelerini lojistik regresyon ile inceleyen analiz ise pek çok maddede başarısızlıkla sonuçlanmıştır. Bu araştırma özellikle örneklem büyüklüęü, Q matris belirleme ve nitelik sayısı bakımından analiz sonucunda ortaya çıkan parametreler sağlıklı kabul edilemeyeceęi için sakıncalı görülmektedir.

İlgili çalışmalara özet olarak bakıldığında, ÇBMTK araştırmalarında telafi edici modelin daha sık kullanıldıęı gözlenmiştir. Telafi edici model ve telafi edici olmayan model karşılaştırmalarında telafi edici modelin daha iyi performans göstermesi telafi edici olmayan modelin geri planda kalmasına neden olmuş olabilir. Çalışmaların büyük çoğunluęunun simülasyon veri üzerine kurulması ÇBMTK'nin hala uygulamada yeterince yerini almadıęını göstermektedir.

BTM alanında yapılan çalışmalara bakıldığında uygulamaya dönük çalışmaların ön planda olduęu görülmektedir. Yeni modellerin ortaya çıkışının hala devam ettięi gözlenmiştir. Bu modeller arasında bu tezin araştırma konusu araştırma konusu içinde yer alan çok kategoriye sahip Q matrisleri ve hiyerarşik ilişkiye sahip Q

matrisleri de yer almaktadır. Bu çalışmalara yakın olarak daha iyi model veri uyumuna sahip olabilecek yeni bir model önerilmiştir.

Uyarlama çalışmalarında dikkat çeken husus BTM'nin MTK veya ÇBMTK verisine uyarlama girişimlerinin daha sık olarak görülmesidir. Ayrıca bu girişimlerin gerçek veriye dayalı olarak gerçekleştirildiği görülmüştür. MTK/ÇBMTK altyapısında hazırlanan testler sonradan hazırlanan Q matrisi yardımıyla BTM'lerde analiz edilmiştir. Ancak bu yoldan elde edilen sonuçların iyi düzeyde olmadığı görülmüştür. Bu yüzden BTM uygulamalarının tamamen BTM'ye uygun hazırlanan verilerle gerçekleştirilmesi tavsiye edilmektedir.

Yapılan araştırmalar eşliğinde çalışmaya bakıldığında hem önerilen model ile hem de kullanılacak diğer modeller ve farklı koşullar ile alana yenilikler kazandırması amaçlanmaktadır. Bilişsel tanı bilgisi ve madde tepki kuramından gelecek bilgilerin birbirini desteklemesi, modeller birbirine yakın olduğu sürece anlamlı olacaktır. Yakın modellerin hem birbiri yerine kullanımı hem de uyarlama sonucu ortaya çıkan yeni bilgiler ile birbirini desteklemesi ortaya çıkacak çıktılar veriden elde edilen toplam bilgiyi artıracaktır.

### 3. YÖNTEM

Bu bölümde araştırmanın türü, verilerin elde edilmesi ve verilerin çözümlenmesi ile bilgiler ilgili alt başlıklarda verilmiştir.

#### 3.1. Araştırmanın Türü

Araştırma kapsamında ele alınan modellere uygun veriler üretilip sonrasında üretilen ve uyarlanan modellerdeki performansa odaklanılmıştır. Verilerin istenilen ölçütlere uygun olarak türetilip ÇBMTK ve BTM çerçevesinde uygulamalar ve karşılaştırmalar yapılacağından araştırma temel araştırma olarak ele alınabilir.

#### 3.2. Verilerin Elde Edilmesi

Çalışma her bir koşulda 25 tekrar ile gerçekleştirilmiştir.

Verilerin üretilmesinde kullanılan koşullar şunlardır:

**Ayirt edicilik:** Uniform dağılıma sahip olan düşük  $a \sim U(0,6 - 0,8)$ , orta  $a \sim U(0,9 - 1,1)$  ve yüksek düzeyde  $a \sim U(1,5 - 1,7)$  ayirt edicilik değerleri kullanılmıştır. Ayirt edicilik kullanan uyarlama çalışmalarının ikisinde de yakın değerlerin ayirt edicilik aralığı olarak kullanıldığı görülmüştür (de la Torre & Karelitz, 2009; Wang, 2009)

**Çok Kategorili Q Matrisi:** BTM parametresinin ÇBMTK parametresini kontrol ettiği değişken olarak analizde yer almıştır. 4 kategoriye (0, 1, 2 ve 3) sahip Q matrisindeki değerler 0,67 ile çarpım sonucundan 1,34 çıkartılarak madde güçlük parametresi elde edilmiştir. Buna göre Q matrisinde 0, 1, 2 ve 3 değerlerine karşılık güçlük parametresinde sırasıyla -1,34, -0,67, 0 ve 0,67 değerleri karşılık gelmiştir. Güçlük parametresindeki son 3 değer aynı zamanda yetenek puanlarını kesikli hale getirmek için kullanılan kesme noktalarıdır. Bu kesme noktaları çok değişkenli normal dağılıma sahip grubu yaklaşık olarak 4 eşit parçaya bölmektedir. Kesme noktaları olarak bu değerlerin seçimi ile yetenekler arası korelasyonun olduğu durumlarda dahi bir boyuttan yüksek değerinden düşük yetenek parametresine sahip bireylerin görülmesi sağlanmıştır.

Q matrisinde 0 değerine karşılık gelen -1,34 değeri simetrik yapıyı bozuyor görünse de bu değer yer aldığı madde sadece diğer boyutu ölçtüğü için ayirt edicilik değeri 0'a eşitlenmiştir. Böylece ölçülmeyen boyuta ilişkin güçlü değeri 0 ile çarpımından dolayı etkisiz hale gelmiştir.

**Yetenekler Arası Korelasyon:** Temel etkinin görülmesi adına korelasyonun 0 olduğu koşul ve korelasyon etkisinin anlaşılması adına korelasyonun 0,6 olduğu değer tüm koşullar için çaprazlanarak kullanılmıştır. Bireyleri ölçülen çoğu bilişsel özellikler birbiri ile ilişkili olduğundan boyutlar arası korelasyonun 0 olması pratik görünmemektedir. Boyutlar arası korelasyonun 0,5-0,6 düzeylerinde kullanılması uygun görülebilir (Reckase, 2007).

**Madde Yapı Türlerinin Testteki oranı:** Maddelerin birden fazla boyuta yük verip vermemelerine göre değişen madde yapısı tüm testi kapsayacak şekilde değil, orantısal olarak farklılaştırılmıştır. Buna göre 15 maddelik testte maddelerin, çoğunlukla karmaşık yapı için 3'ü basit ve 12'si karmaşık, eşit oran için 7'si basit ve 8'si karmaşık, çoğunlukla basit yapı için 12'si basit ve 3'ü karmaşık yapıdadır. 30 maddelik testte ise yapılar aynı sırayla 6-24, 15-15, 24-6 madde ile 60 maddelik testte ise yapılar yine aynı sırayla 12-48, 30-30, 48-12 madde ile temsil edilmektedir. Diğer bir ifade ile madde yapılarının testte %20-%80, %50-%50 ve %80-%20 oranlarda temsil edilmiştir. 6 maddelik (01, 02, 03, 10, 20, 30) basit yapıya sahip maddelerin aynı oranda temsil edilmesine özen gösterilmiştir. Aynı durum 9 maddelik (11, 12, 13, 21, 22, 23, 31, 32, 33) karmaşık yapıdaki maddeler için de geçerlidir. Tekrarlı yapılan işlemde maddelerin yaklaşık aynı oranda temsil edilmesi sağlanmıştır.

**Madde sayısı:** 3 kesme noktası ve dolayısıyla 4 kategoriden oluşan maddelerin Q matrisinde tam olarak temsil edilebilmesi için  $2^k-1=2^4-1=15$  maddeye ihtiyaç duyulmaktadır. Bu sayı kısa test uzunluğu olarak belirlenmiştir. İki katı olan 30 madde orta uzunlukta, 60 madde ise uzun test olarak belirlenmiştir.

**ÇBMTK Modelleri:** Çaprazlanan her koşul için 2PL Telafi edici ÇBMTK ve 2PL Telafi edici Olmayan ÇBMTK verisi üretilerek analizler gerçekleştirilmiştir. İncelenen pek çok çalışmada 2 parametrelili yöntemlerin kullanıldığı görülmüştür (Wang, 2009; Hartig & Höhler, 2008; Chalmers & Flora, 2014; Bolt & Lall, 2003; Spray ve diğerleri 1990; Reckase, 1985; Reckase & McKinley, 1991).

**BTM Modelleri:** İlk alt problemde pG-DINA ve fA-M özdeş, log ve logit fonksiyonlarıyla tahmin için kullanılmıştır. İkinci alt problemde ise bu modeller aynı zamanda veri üretimi için de kullanılmıştır.

İlk alt problem için:

1. Birey parametreleri arasında korelasyonun 0 ve 0,6 olduğu durumlar için ayrı ayrı çok değişkenli normalliğe sahip iki boyutlu 5.000 bireylik yetenek parametresi üretilmiştir.
2. Telafi edici ve telafi edici olmayan ÇBMTK 2PL modelleri için test uzunluğu, madde yapısının testteki oranı ve madde ayırt ediciliği sırasıyla çaprazlanarak veriler üretilmiştir. Böylece toplamda 108 veri seti üretilmiştir.

İkinci alt problem için:

1. İlk alt problem için üretilen verilerin aynısı 100.000 birey için üretilmiştir. Bunu yapmaktaki amaç BTM'ler için veri üretiminde temel teşkil edecek madde parametrelerini birey sayısından fazla etkilenmeden istikrarlı olarak elde etmektir. 5.000'den küçük grupları araştırmamızda bulunan 16 örtük gruba yerleştirmek her seferinde farklı ve düzensiz grup büyüklüklerine yol açabileceğinden 100.000 birey parametresi üzerinden madde parametresinin elde edilmesi yoluna gidilmiştir.
2. Her iki boyut için birey parametresinin 0,1,2,3 kesikli hale getirilmesi ardından her örtük grupta kaç bireyin bulunduğuna ve bunlardan kaçının doğru yanıt verdiği bilgileri elde edilmiştir. Grup büyüklükleri toplam birey sayısına, doğru sayıları ise grup büyüklüklerine bölünerek orantısal hale getirilmiştir.

**Tablo 3.3: Örnek Örtük Sınıflar İçin ÇBMTK Grup Büyüklükleri ve Doğru Yanıtlama Olasılıkları**

<i>Örtük Sınıflar</i>	<i>(0,0)</i>	<i>(0,1)</i>	<i>(0,2)</i>	<i>(0,3)</i>	<i>(1,0)</i>	<i>(1,1)</i>	<i>(1,2)</i>	<i>(1,3)</i>
<b>Grup Büyüklüğü</b>	0,067	0,063	0,056	0,062	0,059	0,059	0,063	0,061
<b>Tür Yapı</b>								
<b>TE Basit</b>	0,13	0,1	0,12	0,11	0,38	0,39	0,38	0,39
<b>TEO Basit</b>	0,14	0,14	0,1	0,1	0,36	0,36	0,36	0,34
<b>TE Karmaşık</b>	0,07	0,22	0,42	0,71	0,21	0,52	0,74	0,93
<b>TEO Karmaşık</b>	0,05	0,15	0,21	0,27	0,07	0,26	0,38	0,57
<i>Örtük Sınıflar</i>	<i>(2,0)</i>	<i>(2,1)</i>	<i>(2,2)</i>	<i>(2,3)</i>	<i>(3,0)</i>	<i>(3,1)</i>	<i>(3,2)</i>	<i>(3,3)</i>
<b>Grup Büyüklüğü</b>	0,066	0,062	0,069	0,070	0,062	0,060	0,066	0,056
<b>Tür Yapı</b>								
<b>TE Basit</b>	0,64	0,66	0,6	0,63	0,88	0,91	0,87	0,88
<b>TEO Basit</b>	0,65	0,59	0,63	0,62	0,9	0,85	0,89	0,87
<b>TE Karmaşık</b>	0,42	0,73	0,91	0,97	0,74	0,96	0,98	1
<b>TEO Karmaşık</b>	0,09	0,33	0,52	0,74	0,12	0,29	0,6	0,81

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 3.3 ve 3.4'deki örnek maddeler Q matrisinde (2,0) basit yapıdaki ve (1,2) karmaşık yapıdaki, ÇBMTK'de (0, -1,34) ve (-0,67, 0) madde zorluğu, (1, 0) ve (1, 1) madde ayırt edicilik parametrelerine sahiptir. Tablo 3.3'den örtük grupların birbirine benzer büyüklükte olduğu görülmektedir. ÇBMTK'de gerçekleşen maddeyi doğru yanıtlama olasılıklarına bakıldığında basit yapıdaki madde için telafi edici ve telafi edici olmayan model arasında ciddi farklar görülmemektedir. Küçük farklar ise veri üretiminden kaynaklanmaktadır. Ayrıca ilk niteliğin ve boyutun maddede yoklanmaması nedeniyle madde yanıtlama olasılığı ilk örtük sınıfa duyarlı değildir. Modeller arası fark karmaşık yapıdaki maddenin doğru yanıtlama olasılıklarına bakıldığı zaman kendini göstermektedir. Örneğin (3,1) örtük sınıfındaki bireyin telafi edici modelde bu soruya doğru yanıt verme olasılığı 0,96 iken telafi edici olmayan modelde bu oran 0,29'a düşmektedir.

3. pG-DINA verisi üretimi için 16 grup basit yapıdaki maddeler için 2'ye, karmaşık yapıdaki maddeler için 4'e indirgenmiştir. Yapılan işlem birey parametreleri maddenin güçlük parametreleri ile karşılaştırıp, her boyut için bireylere 0 veya 1 değeri vermektir. Birey parametresi güçlük parametresinden büyükse maddeyi yapacağı düşüncesiyle 1 aksi takdirde 0 değeri verilmiştir. Böylece iki boyutunda yoklandığı sorular için (00, 10, 01, 11) gruplarının, tek boyutlu maddeler içinse (0,1) gruplarının doğru yanıt verme olasılıkları hesaplanmıştır. PG-DINA için hesaplanan madde parametreleri örnek maddeler üzerinden gösterilecek olursa;

**Tablo 3.4: pG-DINA İçin Örnek Grupların Doğru Yanıtlama Olasılıkları**

<i>Örtük Sınıflar</i>		<i>00</i>	<i>10</i>	<i>01</i>	<i>11</i>
<b>Tür</b>	<b>Yapı</b>				
<b>TE</b>	<b>Basit</b>	0,25	0,75	-	-
<b>TEO</b>	<b>Basit</b>	0,24	0,74	-	-
<b>TE</b>	<b>Karmaşık</b>	0,14	0,59	0,58	0,92
<b>TEO</b>	<b>Karmaşık</b>	0,10	0,19	0,24	0,60

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

pG-DINA veri üretimi için çökertilmiş örtük sınıfların maddeyi doğru yanıtlama olasılıkları Tablo 3.4'de görülmektedir. Karmaşık yapıdaki maddenin telafi edici modeldeki 1,0 ve 0,1 örtük grupların birbirine olan benzerliği dikkat

çekmektedir. Ayrıca modeller arası fark yine karmaşık yapıdaki maddede belirgindir.

4. Q matrisi ve elde edilen grupların doğru yanıt verme olasılıkları kullanılarak, R programında (R Core Team, 2015) yazılan GDINA (Ma & de la Torre, 2016) paketi aracılığıyla pG-DINA verisi elde edilmiştir. Pakette yapılan işlemler yine R üzerinde yayınlanan CDM (Robitzsch, Kiefer, George ve Uenlue, 2016) paketi ile ve Ox-Edit üzerinde Chen ve de la Torre tarafından yazılan kod ile doğrulanmıştır.
5. FA-M'ler için üretilecek veriler için 2. aşamada elde edilen orantısal grup büyüklüklerini dönüştürme yoluna gidilmiştir. İşlem sadece karmaşık yapıdaki maddeler için gerçekleştirilmiştir. Basit yapıdaki maddeler için madde parametresi olarak ÇBMTK verisinden gelen 4 grubun maddeyi doğru yanıtlama olasılığı düşürme işlemi de gerçekleştirilerek kullanılmıştır. İlk niteliğin yoklandığı basit madde için örneğin (0,0 – 0,1 – 0,2 – 0,3) grubundaki bireyler 0 grubunda toplanmıştır. Fully Additive Model'ler eklemeli modeller olduğundan çok kategorili Q matrisi için niteliklerde bulunan en yüksek seviye ve temel etki parametresi gerekmektedir. Her iki nitelik için de en yüksek seviye derecesi 3 olduğundan  $3+3+1=7$  temel parametre her iki niteliğinde yoklandığı maddede bulunacaktır. Dönüştürme işleminde yapılan temel işlem 7 parametrenin gerekli olanlarının toplandığı zaman 16 grubun doğru yanıtlama olasılıklarına en yakın sonucu verecek parametreleri bulmaktır. Bu işlemler özdeş, log ve logit fonksiyonları için ayrı ayrı yapılmıştır. Özdeş fonksiyonunda etki büyüklükleri doğrudan toplanmakta, log fonksiyonunda etki büyüklükleri log işlemi altında toplanmakta ve logit fonksiyonunda ise etki büyüklükleri logit işlemi altında toplanmaktadır.

FA-M için hesaplanan madde parametreleri Q matrisinde (2,0) basit yapıdaki ve (1,2) karmaşık yapıdaki, ÇBMTK'de (0, -1,34) ve (-0,67, 0) madde zorluğu, (1, 0) ve (1, 1) madde ayırt edicilik parametrelerine sahip olan 2 örnek madde üzerinden gösterilecek olursa;



**Tablo 3.5: fA-M Örnek Madde Parametreleri**

		Temel Etki	1. Nitelik 1. Etki	1. Nitelik 2. Etki	1. Nitelik 3. Etki	2. Nitelik 1. Etki	2. Nitelik 2. Etki	2. Nitelik 3. Etki
Tür	Link							
TE	Özdeş	,07	,26	,17	,05	,27	,16	,04
TE	log	-1,54	,55	,19	,04	,55	,18	,03
TE	logit	-2,83	1,43	1,04	1,52	1,51	1,03	1,31
TEO	Özdeş	,02	,07	,10	,03	,14	,17	,18
TEO	log	-3,21	,65	,28	,08	1,14	,52	,34
TEO	logit	-3,60	,98	,54	,15	1,41	,80	,76

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Tablo 3.6: Örnek Örtük Sınıflar İçin ÇBMTK Grup Büyüklükleri ve Doğru Yanıt Olasılıkları**

Örtük Sınıflar	0,0	0,1	0,2	0,3	1,0	1,1	1,2	1,3
Grup Büyüklüğü	0,07	0,06	0,06	0,06	0,06	0,06	0,06	0,06
Tür	Link							
TE	Özdeş	0,07	0,33	0,49	0,53	0,32	0,59	0,75
TE	log	0,21	0,37	0,44	0,46	0,37	0,64	0,79
TE	logit	0,06	0,21	0,43	0,73	0,20	0,53	0,92
TEO	Özdeş	0,02	0,16	0,32	0,50	0,09	0,22	0,39
TEO	log	0,04	0,13	0,21	0,29	0,08	0,24	0,56
TEO	logit	0,03	0,10	0,20	0,35	0,07	0,23	0,59
Örtük Sınıflar	2,0	2,1	2,2	2,3	3,0	3,1	3,2	3,3
Grup Büyüklüğü	0,07	0,06	0,07	0,07	0,06	0,06	0,07	0,06
Tür	Link							
TE	Özdeş	0,49	0,76	0,92	0,95	0,54	0,80	0,96
TE	log	0,45	0,78	0,93	0,96	0,47	0,81	0,97
TE	logit	0,41	0,76	0,90	0,97	0,76	0,94	0,99
TEO	Özdeş	0,18	0,32	0,48	0,66	0,22	0,35	0,52
TEO	log	0,10	0,32	0,53	0,74	0,11	0,35	0,58
TEO	logit	0,11	0,34	0,53	0,71	0,13	0,38	0,74

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Örnek üzerinden gidecek olursak fA-M'da (2,1) örtük sınıfa sahip bireylerin Q matrisinde (1,2) ile temsil edilen telafi edici modeldeki madde için doğru yanıt verme olasılıkları temel etki, 1. nitelik 1. etki, 1. nitelik 2. etki ve 2. nitelik 1. etki toplamı ile elde edilir.

Özdeş fonksiyonu altında:  $\delta_0 + \delta_1^1 + \delta_2^1 + \delta_1^2 = 0,07 + 0,26 + 0,17 + 0,27 = 0,77$

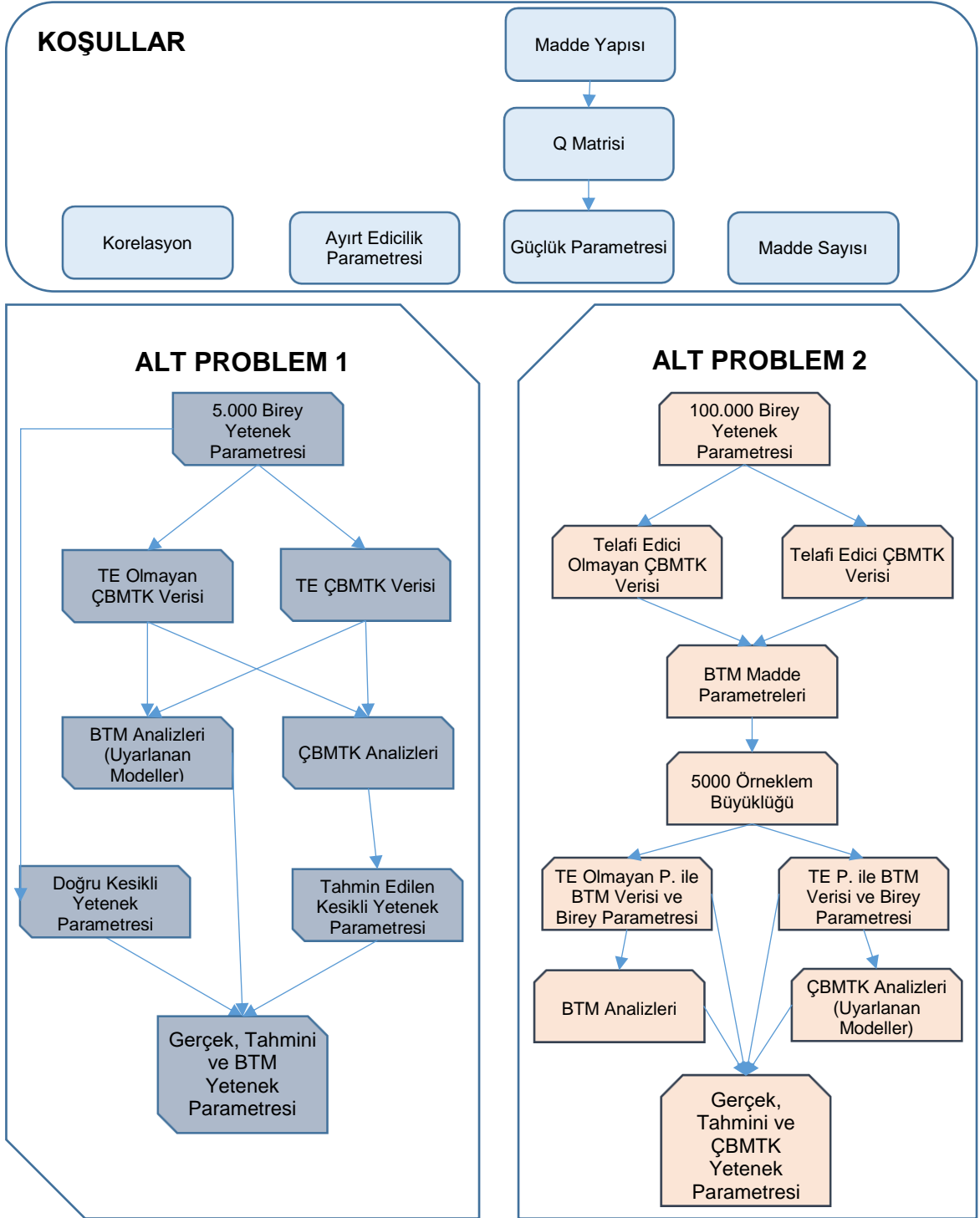
log fonksiyonu altında:  $v_0 + v_1^1 + v_2^1 + v_1^2 = \exp(-1,54 + 0,55 + 0,19 + 0,55) =$

$$\exp(-1,54) * \exp(0,55) * \exp(0,19) * \exp(0,55) = 0,78$$

logit fonksiyonu altında:  $\lambda_0 + \lambda_1^1 + \lambda_2^1 + \lambda_1^2 = \exp(-2,83 + 1,43 + 1,04 + 1,51) / (1 + \exp(-2,83 + 1,43 + 1,04 + 1,51)) = 0,76$

Küsuratlar nedeniyle küçük farklar göz ardı edildiğinde elde edilen sonuçların Tablo 3.6 ile aynı olduğu görülmektedir. Karmaşık yapıdaki maddeler için gerçekleştirilen temel işlem Tablo 3.5 ile Tablo 3.6 değerleri arasındaki farkları en aza indirecek parametreleri elde etmektir. Bu işlem R yazılımı üzerindeki `nloptr` paketindeki (Johnson) `slsqp` fonksiyonu ve yazılan kodlarla gerçekleştirilmiştir.

Her 3 fonksiyon için bulunan maddeye doğru yanıt verme olasılıkları yine R yazılımı üzerinde yazılan fonksiyon ile veri seti üretiminde kullanılmıştır.



**Şekil 3.9. Araştırma İşlem Basamakları**

### 3.3. Verilerin İşlenmesi ve Çözülmesi

Araştırmada türetilen yetenek parametreleri ile modellere uygun üretilen veriler analiz sonucu ortaya çıkan birey parametreleri karşılaştırılmıştır. Modellerin madde parametreleri birbirine tam olarak çevrilemeyeceği göz önünde bulundurularak

sadece birey parametreleri üzerinde durulmuştur. Ölçme işleminde bireyin durumu hedef alınarak buna aracılık eden madde parametreleri ikinci planda tutulmuştur.

Verilerin analizi yine R yazılımı üzerinden gerçekleştirilmiştir. ÇBMTK tahmini için mirt (Chalmers, 2012) paketi, pG-DINA tahmini için GDINA (Ma & de la Torre, 2016) paketi, fA-M tahminleri için R yazılımı üzerinde hazırlanan kodlar kullanılmıştır. Yapılan işlemler şu şekilde sıralanabilir:

- 1- ÇBMTK modellerinin tahminleri mirt paketi üzerinden elde edilmiştir. EM algoritmasının kullanıldığı analizde tahmin döngüsünün kesme noktası için 0,005 yakınsama puanı ve EAP birey parametresi kestirme yöntemi kullanılmıştır.
- 2- Analiz sonucu elde edilen birey parametreleri güçlük parametresinin dönüşümünde olduğu gibi -0,67, 0 ve 0,67 kesme puanları kullanılarak (0, 1, 2, 3) formunda kesikli hale getirilmiştir. Aynı şekilde veri üretiminde kullanılan birey parametreleri de kesikli hale getirilmiştir.
- 3- Çalışmada önemli role sahip doğru sınıflama oranları (DSO) vektör, nitelik ve yetenek düzeyi için ayrı ayrı hesaplanmıştır. DSO'lar analiz sonucu gelen birey parametrelerinin gerçek birey parametrelerini ne oranda doğru tahmin ettiğini gösteren ortalamalardır. Bu ortalamalar her bir modelin 5000 kişilik bireylerin yüzde kaçının doğru bir şekilde tahmin edildiğini gösterir.

Vektör bazında tam DSO'lar bireylerin her iki yetenek parametresinin de doğru tahmini edilmesine bağlı ortalamalardır. Grafikte çubuklarının alt kısmında V ile temsil edilmişlerdir.  $V \pm 1$  ise birey parametrelerinin bir yetenek için doğru, diğer yetenek için bir düzeylik hata ile elde edilmesi durumunda hesaplanan ortalamadır. Örnek verecek olursak (1,1) birey parametresindeki birey analiz sonucu (1,1) birey parametresine sahipse tam doğru vektör (V) grubunda, (1,0),(1,2),(0,1) veya (2,1) birey parametresine sahipse  $\pm 1$  vektör ( $V \pm 1$ ) grubunda yer alır. Sonra bu grupların büyüklükleri örneklem sayısı olan 5000'e bölünerek bireylerin ne oranda bir kısmının doğru tahmin edildiği bulunur.

Nitelik bazındaki DSO'lar ise, test dizaynının simetrik yapısı nedeniyle iki yetenek/nitelik arasında fark bulunmaması nedeniyle iki nitelik için ayrı ayrı değil, ikisinin ortalaması alınarak tek değer olarak verilmiştir. Bireyin ilk veya

ikinci yetenek değeri analiz sonucu doğru olarak kestirildiyse tam, bir yetenek düzeyi hata ile kestirildi ise  $\pm 1$  nitelik grubunda olduğu değerlendirilmiştir. Örneğin ilk yetenek parametresi 0 olan bireyin yine ilk yeteneği 0 olarak kestirildiyse tam nitelik, 1 olarak kestirildiyse  $\pm 1$  nitelik grubunda değerlendirilir. Bu bireylerin toplam sayısı örnekleme bölünerek nitelikler için ayrı ayrı oranlar elde edilir. Ancak nitelik/yetenekler arası ciddi fark görülmediğinden tam nitelik ve  $\pm 1$  nitelik için iki oranın ortalaması raporlanmıştır.

Yetenek düzeyine göre DSO'lar ise modellerin kestirimlerinde yetenek düzeylerini hangi oranda doğru kestirdiğine ilişkin bilgi taşımaktadır. Nitelik düzeylerine göre bireyler 0,1,2,3 değerleri almaktadır. Modellerin bu değerlerde nitelik düzeyine sahip bireyleri hangi oranda doğru sınıfladığını ilişkin bilgi bu istatistikte bulunur. Örneğin ilk nitelik için 0 değeri taşıyan 1200 bireyin a modelince 800'ü yine 0 değerinde kestirildi ise modelin 0 düzeyindeki başarısı  $800/1200=0,67$  olarak hesaplanır. İkinci nitelik için de benzer işlem gerçekleştirildikten sonra, niteliklerin simetrik yapısı nedeniyle ilgili düzeyi için elde edilen iki oranın ortalaması raporlanmıştır.

- 4- ÇBMTK verilerine pG-DINA analizi Q matrisi aracılığıyla GDINA paketinde uygulanmıştır. 0,001 yakınsama puanına ulaşıldığında sonlandırılan analizde birey puanlarının elde edilmesinde expected a posteriori (EAP) metodu kullanılmıştır. Metodun seçilmesindeki amaç BTM'lerde yöntemin diğer yöntemlere göre daha iyi sonuç vermesidir (Heubner & Wang, 2011). Analiz sonucu elde edilen birey parametreleri 3. adımda olduğu gibi değerlendirilmiştir.
- 5- Yine ÇBMTK verilerine her üç fA-M uygulanmıştır. 250 döngü sonrasında 0,005'lik kesme kriterine ulaşamayan analizler raporlanmamıştır. Analizin daha hızlı ve sağlıklı gerçekleştirilmesi adına ÇBMTK verilerinden elde edilen örtük sınıfların doğru yanıtlanma olasılıkları analize başlama noktası girdisi olarak verilmiştir. Birey parametresi yine EAP yöntemi ile elde edilmiştir.
- 6- 2. alt problem için elde edilen BTM verileri ilgili BTM ile ve verinin alt yapısında bulunan telafi edici veya telafi edici olmayan ÇBMTK ile analiz edilmiştir. Verinin yapısı değiştiğinden dolayı ÇBMTK analizi 2PL ve 3PL

modellerinden hangisinde daha iyi uyum sağladı ise o modelde gerçekleştirilmiştir. Çok büyük bir oranda 3PL modeli veri için daha iyi uyumu sağladığı görülmüştür.

## 4. BULGULAR VE TARTIŞMA

Bu bölümde birinci alt problemdeki BTM'lerin ÇBMTK verisine uyarlanmasını ve ikinci alt problemdeki ÇBMTK'nin BTM verilerine uyarlanmasını içeren analiz sonuçları yer almaktadır.

### 4.1. ÇBMTK Verisi Analizlerine İlişkin Bulgular ve Tartışma

Bu kısımda ÇBMTK verisinin ÇBMTK, pG-DINA ve özdeş, log ve logit fonksiyonlarındaki fA-M'lara uygulanması ile elde edilen sonuçlar sırasıyla madde ayırt edicilik, madde yapısı, test uzunluğu ve yetenekler arasındaki korelasyon koşullarına göre ele alınmıştır. Nitelik için tabloda verilen doğru sınıflama oranı (DSO) değerleri, vektör için tablonun yanında karşılaştırmayı kolaylaştırmak adına şekil üzerinde de gösterilmiştir.

#### 4.1.1. Madde Ayırt Edicilik Koşullarına Göre Doğru Sınıflama Oranları

Birinci alt problemin ilk koşulu olan madde ayırt edicilik indeksi için 3 ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. Düşük madde ayırt edicilik için  $a \sim U(0,6-0,8)$ , orta düzey için  $a \sim U(0,9-1,1)$  ve yüksek düzey içinse  $a \sim U(1,5-1,7)$  aralıkları kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda düşük düzey değerleri ele alınmış, orta düzey için elde edilen değerler düşük düzey değerleri ile ve yüksek düzey için elde edilen değerler ise orta düzey değerleri ile karşılaştırılarak değerlendirilmiştir.

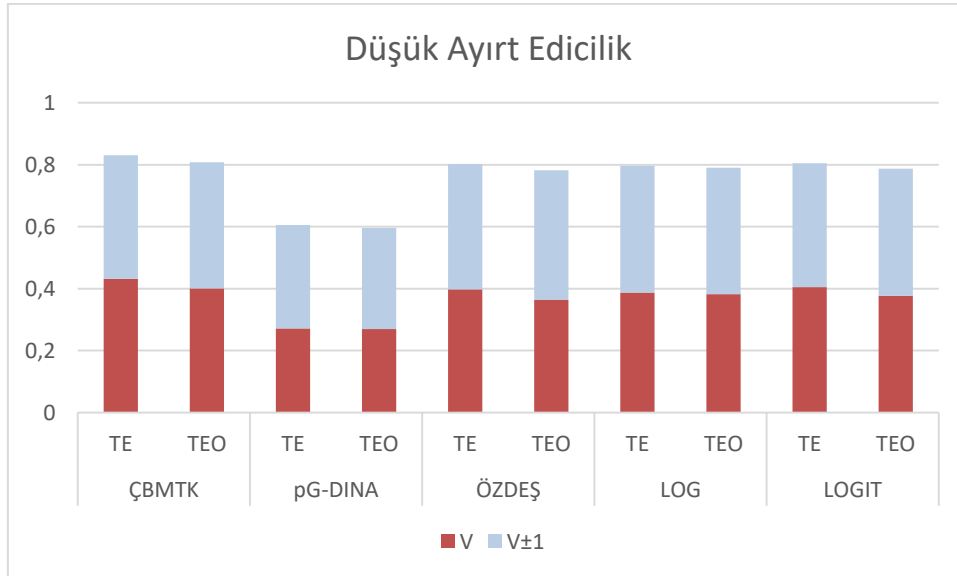
Tablo 4.7: Madde Ayırt Ediciliğe Göre ÇBMTK Verisi DSO'ları

		ÇBMTK		pG-DINA		ÖZDEŞ		LOG		LOGIT		
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	
Vektör	Düşük	TE	0,43	0,40	0,27	0,33	0,40	0,40	0,39	0,41	0,41	0,40
		TEO	0,40	0,41	0,27	0,33	0,36	0,42	0,38	0,41	0,38	0,41
	Orta	TE	0,50	0,38	0,29	0,36	0,45	0,39	0,44	0,40	0,48	0,38
		TEO	0,48	0,39	0,30	0,36	0,43	0,41	0,47	0,39	0,46	0,40
	Yüksek	TE	0,58	0,34	0,33	0,40	0,48	0,38	0,46	0,41	0,56	0,34
		TEO	0,57	0,35	0,38	0,39	0,49	0,39	0,58	0,34	0,56	0,35
Nitelik	Düşük	TE	0,64	0,33	0,50	0,35	0,62	0,35	0,61	0,35	0,62	0,35
		TEO	0,62	0,35	0,50	0,35	0,59	0,36	0,61	0,35	0,60	0,36
	Orta	TE	0,70	0,29	0,53	0,36	0,66	0,32	0,65	0,32	0,67	0,31
		TEO	0,68	0,30	0,53	0,36	0,65	0,33	0,67	0,31	0,67	0,32
	Yüksek	TE	0,76	0,24	0,57	0,35	0,68	0,29	0,67	0,30	0,73	0,26
		TEO	0,75	0,24	0,61	0,33	0,70	0,28	0,76	0,24	0,75	0,25

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.7 madde ayırt edicilik koşullarına göre vektör ve nitelik bazında elde edilen DSO'ları göstermektedir. Vektörün iki nitelikten oluşması ve niteliklerin testte eşit ve simetrik olarak temsil edilmesinden dolayı, tam vektör sonuçları, tam nitelik sonuçlarının karesine çok yakın olarak elde edilmiştir. Bu durum  $\pm 1$  seviye hatasıyla elde edilen değerler için geçerli değildir. Bu yüzden tam ve  $\pm 1$  DSO'lar toplandığında vektör için elde edilen değerler nitelik değerlerinin karesinden küçük olacaktır. Çünkü vektör için verilen  $\pm 1$  değerleri sadece bir niteliğin  $\pm 1$  hata ile elde edildiği durumda geçerlidir. Tablo üzerinden nitelik için elde edilen değerler üzerinde durulacaktır.

Nitelik için verilen değerler incelendiğinde madde kalitesi artışının tüm modellerde DSO değerlerini artırdığı görülmektedir. Tam ve  $\pm 1$  sütunları toplandığında ise pG-DINA haricindeki modellerde ciddi bir değişim olmadığı görülmektedir. Madde kalitesi ile tam vektör sonuçlarında görülen artış,  $\pm 1$  seviye hatası ile yapılan tahminlerin bir kısmının tam doğru olarak tahmininden kaynaklanmaktadır. pG-DINA'da artış direk olarak tam DSO'larda görülmektedir. Ancak bu artışa rağmen pG-DINA DSO'larının diğer modellerden belirgin bir şekilde düşük olduğu görülmektedir.



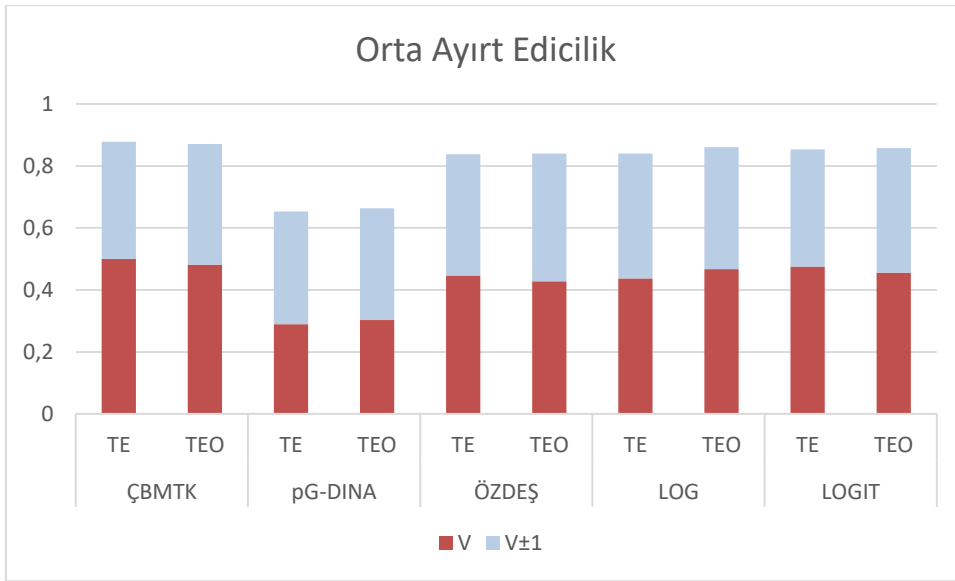
Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.10. Düşük Düzey Madde Ayırt Edicilik Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.10 madde ayırt edicilik indeksinin düşük ( $a \sim U(0,6-0,8)$ ) olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. Modellerin hiçbirinde telafi edici ve



telafi edici olmayan verilerden elde edilen sonuçlar arasında ciddi farklar görülmektedir. Modellere göre sonuçlara bakıldığında pG-DINA DSO'ları diğer model sonuçlarından bariz bir şekilde düşük olduğu görülmektedir. Tam vektör DSO'ları tüm modeller için %40'larda iken bu oran pG-DINA için %30 düzeylerine düşmektedir. Aradaki fark bir boyuttan elde edilen sonucun  $\pm 1$  seviye yaklaşık olmasına izin verildiğinde %20'lere çıkmaktadır. Bu durumda pG-DINA için elde edilen DSO 0,60 iken diğer modellerde 0,80 civarlarındadır. Verinin ait olduğu ÇBMTK ile uyarılama yapılan fA-M DSO'ları arasında büyük farklar bulunmaması, tanıtılan modelin başarılı performansını ortaya koymaktadır.

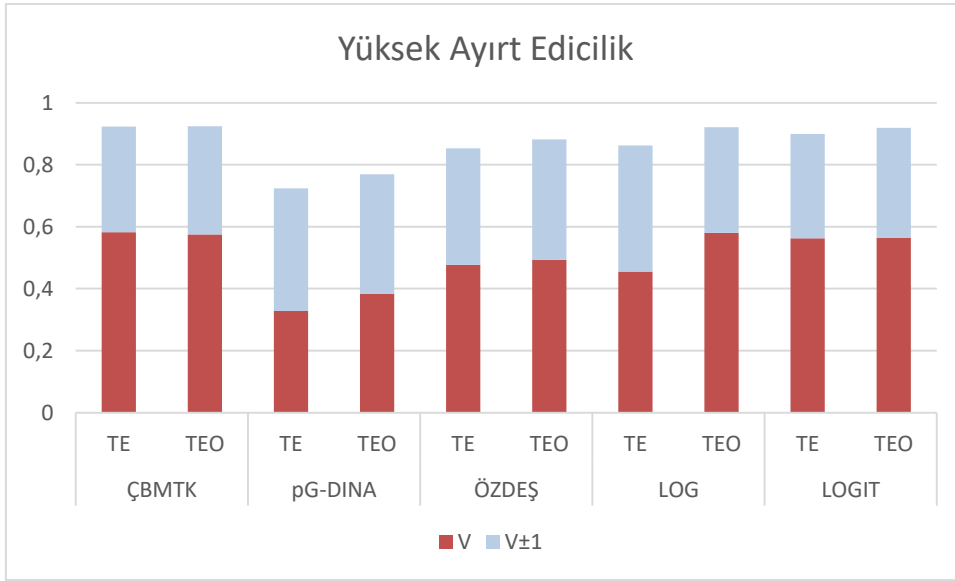


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.11. Orta Düzey Madde Ayırt Edicilik Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.11'deki grafik madde ayırt edicilik indeksinin orta ( $a \sim U(0,9-1,1)$ ) düzeyde olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. Bu koşul için de telafi edici ve telafi edici olmayan verilerden elde edilen sonuçlar arasında farklar önemsizdir. Sonuçlar ayırt edicilik indeksinin düşük olduğu DSO'lar ile kıyaslandığında pG-DINA hariç diğer modellerden elde edilen tam vektör sonuçlarında 0,05-0,09'lük artış görülmektedir. Bu artış verinin ait olduğu ÇBMTK sonuçlarında biraz daha belirgindir. Bu sayede grafikte ÇBMTK sonuçlarının fA-M sonuçlarına göre bir miktar ön plana çıktığı görülmektedir. pG-DINA sonuçlarındaki artış ise 0,02-0,03 düzeyindedir. Madde ayırt ediciliğin düşük olduğu koşulda diğer modellerden çok düşük sonuçlara sahip olan pG-DINA ile diğer modeller arasındaki farkın, orta ayırt edicilikte biraz daha arttığı görülmektedir. Grafik  $V \pm 1$  yönüyle

incelendiğinde pG-DINA'da bir önceki düşük ayırt edicilik koşulundan 0,02-0,03'lük daha iyi sonuçlar elde edilirken, diğer modellerde bu durum tam tersidir, düşük ayırt ediciliğe nazaran 0,02-0,03'lük düşüşler görülmektedir. pG-DINA haricindeki modellerde  $\pm 1$ 'lik hatalı kestirimin bir kısmı, daha iyi düzeydeki madde ayırt edicilik indeksi sayesinde tam doğru vektör kestirimine kaydığı söylenebilir. pG-DINA ile diğer modeller arasındaki DSO farkı  $V \pm 1$  ile birlikte incelendiğinde,  $\pm 1$  değişiminin yönü nedeniyle bir önceki koşula göre değişmediği ve 0,20 seviyelerinde olduğu söylenebilir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.12. Yüksek Düzey Madde Ayırt Edicilik Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.12 madde ayırt edicilik indeksinin yüksek ( $a \sim U(1,5-1,7)$ ) düzeyde olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. Telafi edici ve telafi edici olmayan verilerden elde edilen sonuçlar arasındaki farklar, orta düzeyde ayırt edicilik koşuluna göre biraz daha belirgindir. Özellikle fA-M log ve pG-DINA modellerinde fark diğer modellerden yüksektir ve telafi edici olmayan veride bu modeller daha iyi performans göstermişlerdir. Sonuçlar ayırt edicilik indeksinin orta düzeyde olduğu DSO'lar ile kıyaslandığında, ÇBMTK modellerinin diğer modellerden daha iyi gelişme gösterdiği izlenmiştir. FA-M'lar ise pG-DINA'dan daha iyi gelişme içindedir. Grafikte ÇBMTK sonuçlarının üstünlüğü daha fazla görünür hale gelmiştir. PG-DINA sonuçlarındaki artışın düşük olması, pG-DINA'nın diğer modellerin daha da geride kalmasına neden olmaktadır.  $V \pm 1$  verisi incelendiğinde yine pG-DINA'da bir önceki orta ayırt edicilik koşulundan daha iyi sonuçlar elde

edilirken, diğer modellerde bu durum tam tersidir.  $\pm 1$  dahil vektör DSO'ları incelendiğinde ÇBMTK'de 0,92 seviyesine eriştiği, fA-M'ların ise özellikle de telafi edici olmayan veride log ve logit linkin bu seviyeye çok yakın olduğu göze çarpmaktadır. pG-DINA'nın diğer modellerle arasındaki farkı yüksek düzeydeki ayırt edicilikte bir miktar düşürdüğü görülmektedir, özellikle telafi edici olmayan modelde daha iyi performans artışı göstermiştir. Ancak bu kısmi artış dahi modeller arasındaki en düşük DSO'ya sahip olmasını engelleyememiştir.

#### 4.1.2. Madde Yapısı Oranı Koşullarına Göre Doğru Sınıflama Oranları

Birinci alt problemin ikinci koşulu olan madde yapısı oranı için 3 ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. Karmaşık olarak belirtilen test için %80 karmaşık, %20 basit yapıya sahip maddeler, eşit olarak belirtilen test için %50 karmaşık, %50 basit yapıya sahip maddeler ve basit olarak belirtilen test için %20 karmaşık, %80 basit yapıya sahip maddeler kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda karmaşık değerleri ele alınmış, eşit düzey için elde edilen değerler karmaşık değerleri ile ve basit için elde edilen değerler ise eşit düzey değerleri ile karşılaştırılarak değerlendirilmiştir.

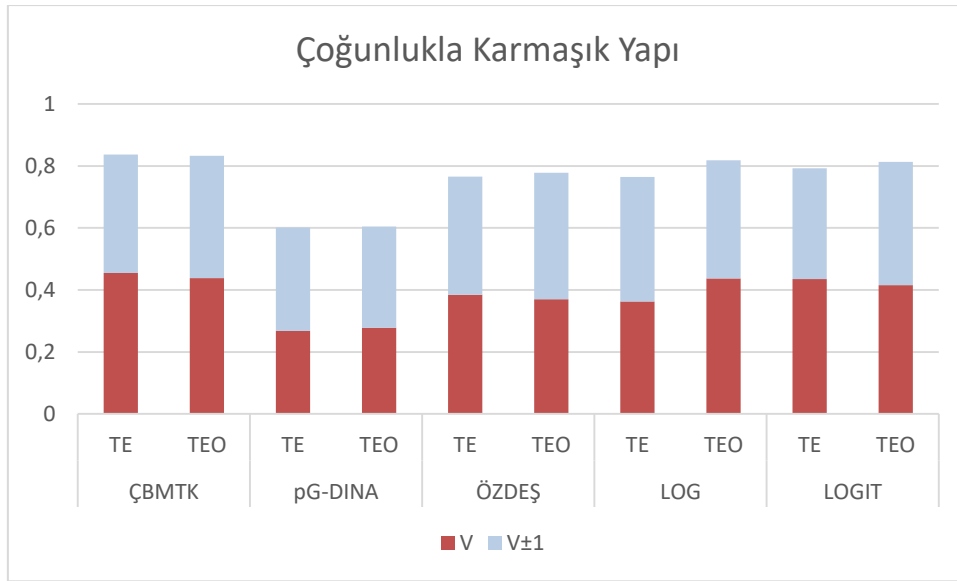
**Tablo 4.8: Madde Yapısı Oranına Göre ÇBMTK Verisi DSO'ları**

		ÇBMTK		pG-DINA		ÖZDEŞ		LOG		LOGIT		
		Tam	$\pm 1$	Tam	$\pm 1$	Tam	$\pm 1$	Tam	$\pm 1$	Tam	$\pm 1$	
Vektör	Karmaşık	TE	0,45	0,38	0,27	0,33	0,38	0,38	0,36	0,40	0,44	0,36
		TEO	0,44	0,39	0,28	0,33	0,37	0,41	0,44	0,38	0,42	0,40
	Eşit	TE	0,52	0,37	0,29	0,37	0,43	0,40	0,43	0,42	0,49	0,38
		TEO	0,49	0,38	0,33	0,37	0,42	0,42	0,48	0,38	0,47	0,39
	Basit	TE	0,54	0,36	0,33	0,39	0,51	0,39	0,48	0,40	0,51	0,38
		TEO	0,53	0,37	0,35	0,38	0,50	0,39	0,51	0,38	0,51	0,38
Nitelik	Karmaşık	TE	0,66	0,32	0,50	0,35	0,60	0,35	0,59	0,36	0,63	0,34
		TEO	0,65	0,32	0,51	0,34	0,60	0,35	0,65	0,32	0,63	0,34
	Eşit	TE	0,71	0,28	0,53	0,36	0,65	0,32	0,65	0,32	0,69	0,30
		TEO	0,69	0,29	0,56	0,35	0,64	0,34	0,68	0,30	0,68	0,31
	Basit	TE	0,73	0,26	0,56	0,36	0,71	0,28	0,69	0,30	0,71	0,28
		TEO	0,72	0,27	0,57	0,35	0,70	0,29	0,71	0,28	0,70	0,28

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.8 madde yapısı oranı koşullarına göre vektör ve nitelik bazında elde edilen DSO'ları göstermektedir. Vektör sonuçları grafikler eşliğinde inceleneceğinden tablo üzerinden nitelik için elde edilen değerler üzerinde durulacaktır.

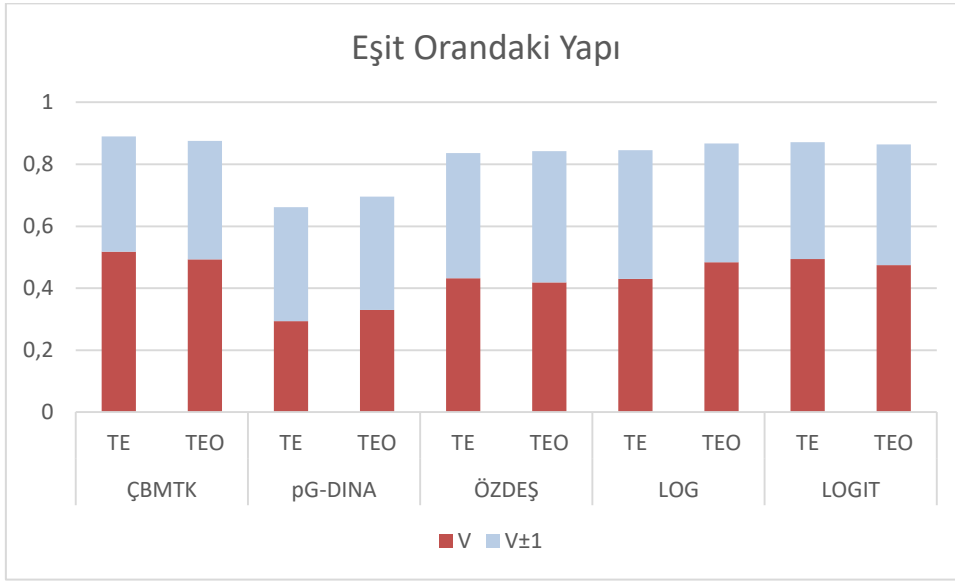
Sonuçları gösteren değerler incelendiğinde testteki maddelerin sadece tek boyutla ilişkili olma oranı artışının tüm modellerde DSO değerlerini artırdığı görülmektedir. Tam ve  $\pm 1$  sütunları toplandığında %100'e yakın doğru sınıflama gerçekleştirilebilmektedir. Ayrıca toplam oranlarda pG-DINA haricindeki modellerde koşullar arası ciddi bir değişim olmadığı görülmektedir. Maddelerin basitleşmesi ile tam vektör sonuçlarında görülen artış,  $\pm 1$  seviye hatası ile yapılan tahminlerin bir kısmının tam doğru olarak tahmini kaynaklıdır. PG-DINA'da artış direk olarak tam DSO'larda görülmektedir. Ancak bu artışa rağmen pG-DINA DSO'larının diğer modellerden belirgin bir şekilde düşük olduğu görülmektedir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

#### **Şekil 4.13. Çoğunlukla Karmaşık Yapı Koşulu ÇBMTK Verisi DSO'ları**

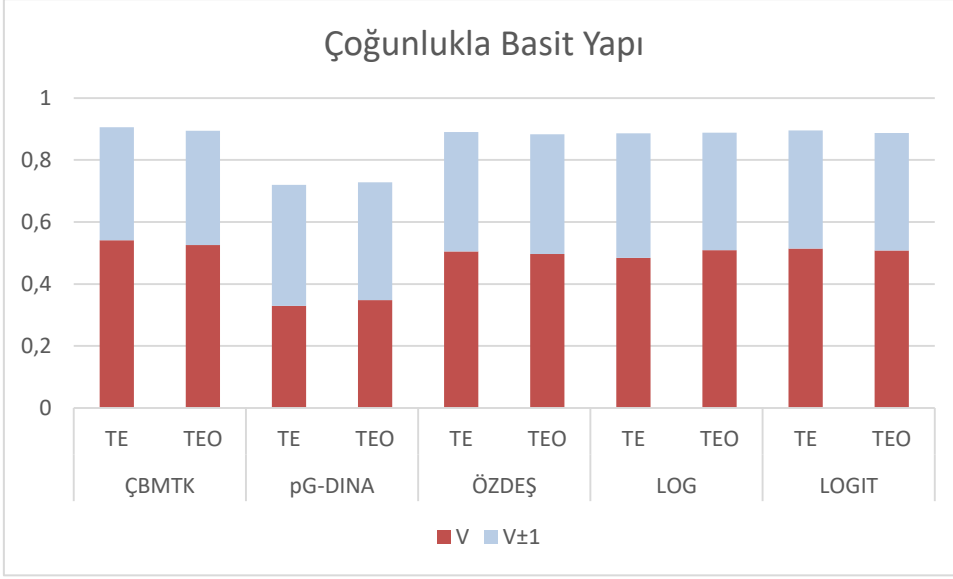
Şekil 4.13'de görülen grafik testteki maddelerin %80'nin karmaşık, %20'sinin basit yapıda olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. FA-M log link haricindeki modellerden telafi edici ve telafi edici olmayan verilerden elde edilen sonuçlar arasında ciddi farklar görülmemektedir. Modellere göre sonuçlara bakıldığında pG-DINA DSO'larının diğer model sonuçlarından bariz bir şekilde düşük olduğu görülmektedir. Tam vektör DSO'ları tüm modeller için %40'larda iken bu oran pG-DINA için %26 düzeylerine düşmektedir. Aradaki fark bir boyuttan elde edilen sonucun  $\pm 1$  seviye yaklaşık olmasına izin verildiğinde %20'lere çıkmaktadır. Bu durumda pG-DINA için elde edilen DSO 0,60 iken diğer modellerde 0,80 civarlarındadır.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

#### Şekil 4.14. Eşit Orandaki Yapı Koşulu ÇBMTK Verisi DSO'ları

Şekil 4.14 testteki maddelerin yarısının karmaşık diğer yarısının ise basit yapıda olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. Maddelerin çoğunlukla karmaşık yapıda olduğu bir önceki grafikte görülen fA-M log link telafi edici olmayan verilerin analizi lehine olan fark, madde yapılarının eşit olarak ele alındığı bu koşulda azalmıştır. pG-DINA'da ise bu farkın telafi edici olmayan veri analizi lehine kısmen de olsa arttığı gözlenmektedir. Sonuçlar maddelerin çoğunlukla karmaşık olduğu DSO'lar ile kıyaslandığında pG-DINA hariç diğer modellerden elde edilen tam vektör sonuçlarında 0,04-0,07'lük artış görülmektedir. Bu artış ÇBMTK sonuçlarında biraz daha belirgindir. Bu sayede grafikte ÇBMTK sonuçlarının fA-M sonuçlarına göre bir miktar ön plana çıktığı görülmektedir. pG-DINA sonuçlarındaki artış ise 0,02-0,03 düzeyindedir. Böylece pG-DINA ile diğer modeller arasındaki farkın biraz daha belirginleştiği görülmektedir. Grafik V±1 yönüyle incelendiğinde, ÇBMTK hariç diğer modellerde önceki koşul olan maddelerin çoğunlukla karmaşık yapıda olduğu koşuldan daha iyi sonuçlar elde edilirken, ÇBMTK'de bu durum tam tersidir. pG-DINA ile diğer modeller arasındaki DSO farkı V±1 ile birlikte genel olarak incelendiğinde, farkın bir önceki koşula göre fazla değişmediği, yine 0,20 seviyelerinde olduğu görülmektedir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.15. Çoğunlukla Basit Yapı Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.15'deki grafik testteki maddelerin %80 basit, %20 karmaşık yapıda olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. Maddelerin yapıya göre eşit sayıda olduğu bir önceki grafikte görülen fA-M log link telafi edici olmayan verilerin analizi lehine olan fark, bu koşulda daha da azalmıştır. PG-DINA'da da bu farkın azaldığı gözlenmiştir. Sonuçlar bir önceki koşul olan maddelerin yapısına göre eşit sayıda yer aldığı sonuçlar ile kıyaslandığında tüm modellerde elde edilen tam vektör sonuçlarında artış görülmektedir. Bu artış fA-M sonuçlarında, özellikle de özdeş linkinde biraz daha belirgindir. Böylece grafikte fA-M sonuçlarının ÇBMTK sonuçlarına göre bir miktar yaklaştığı görülmektedir. PG-DINA sonuçlarındaki artış ise diğer modellerden daha düşük düzeydedir. Grafik  $V_{\pm 1}$  yönüyle incelendiğinde, pG-DINA'da bir önceki maddelerin yapısına göre eşit sayıda olduğu koşuldan daha iyi sonuçlar elde edilirken, diğer modellerde eskiye nazaran düşüşler görülmektedir. PG-DINA haricindeki modellerde  $\pm 1$ 'lik hatalı kestirimin bir kısmı, daha iyi düzeydeki madde ayırt edicilik indeksi sayesinde tam doğru vektör kestirimine kaydığı, böylece tam DSO'larda daha iyi gelişme gösterdiği söylenebilir. PG-DINA ile diğer modeller arasındaki DSO farkı  $V_{\pm 1}$  ile birlikte incelendiğinde, farkın 0,15-0,17 aralığında olduğu görülmektedir.

#### 4.1.3. Test Uzunluğu Koşullarına Göre Doğru Sınıflama Oranları

Birinci alt problemin üçüncü koşulu olan test uzunluğu için 3 ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. 15, 30 ve 60 maddelik test uzunlukları kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda 15 maddelik test değerleri ele alınmış, 30 maddelik test için elde edilen değerler 15 maddelik test değerleri ile karşılaştırılarak değerlendirilmiş ve 60 maddelik test için elde edilen değerler ise 30 maddelik değerleri ile karşılaştırılarak değerlendirilmiştir.

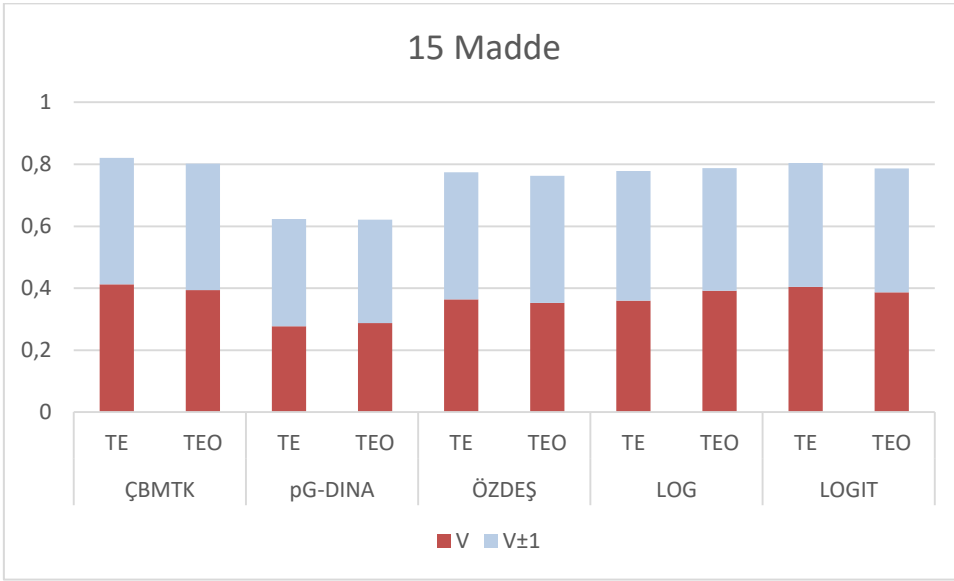
**Tablo 4.9: Test Uzunluğuna Göre ÇBMTK Verisi DSO'ları**

		ÇBMTK		pG-DINA		ÖZDEŞ		LOG		LOGIT	
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1
Vektör	TE	0,41	0,41	0,28	0,35	0,36	0,41	0,36	0,42	0,40	0,40
	15 TEO	0,39	0,41	0,29	0,33	0,35	0,41	0,39	0,40	0,39	0,40
	TE	0,50	0,38	0,30	0,37	0,44	0,40	0,43	0,41	0,48	0,38
	30 TEO	0,49	0,39	0,32	0,36	0,43	0,42	0,48	0,39	0,47	0,40
	TE	0,60	0,33	0,32	0,38	0,52	0,36	0,49	0,38	0,56	0,34
	60 TEO	0,57	0,35	0,35	0,38	0,51	0,39	0,56	0,36	0,54	0,37
Nitelik	TE	0,63	0,34	0,52	0,35	0,60	0,35	0,60	0,36	0,62	0,34
	15 TEO	0,62	0,35	0,52	0,34	0,59	0,36	0,61	0,34	0,61	0,35
	TE	0,70	0,29	0,53	0,36	0,65	0,32	0,65	0,33	0,68	0,31
	30 TEO	0,69	0,29	0,55	0,35	0,65	0,33	0,68	0,30	0,67	0,31
	TE	0,77	0,23	0,55	0,36	0,71	0,28	0,69	0,29	0,73	0,27
	60 TEO	0,75	0,24	0,58	0,35	0,71	0,28	0,74	0,26	0,73	0,27

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.9 test uzunluğu koşullarına göre vektör ve nitelik bazında elde edilen DSO'ları göstermektedir. Vektör sonuçları grafikler eşliğinde inceleneceğinden tablo üzerinden nitelik için elde edilen değerler üzerinde durulacaktır.

Sonuçları gösteren değerler incelendiğinde testteki maddelerin sayısı artışının tüm modellerde DSO değerlerini artırdığı görülmektedir. Tam ve ±1 sütunları toplandığında %100'e çok yakın doğru sınıflama gerçekleşebilmektedir. Ayrıca toplama oranlarda pG-DINA haricindeki modellerde koşullar arası ciddi bir değişim olmadığı görülmektedir. Testin uzaması ile tam vektör sonuçlarında görülen artış, ±1 seviye hatası ile yapılan tahminlerin bir kısmının tam doğru olarak tahmininden kaynaklanmaktadır. pG-DINA'da artış direkt olarak tam DSO'larda görülmektedir. Ancak bu artışın kısıtlı olması pG-DINA DSO'larının diğer modellerden belirgin bir şekilde düşük olmasına neden olmaktadır.

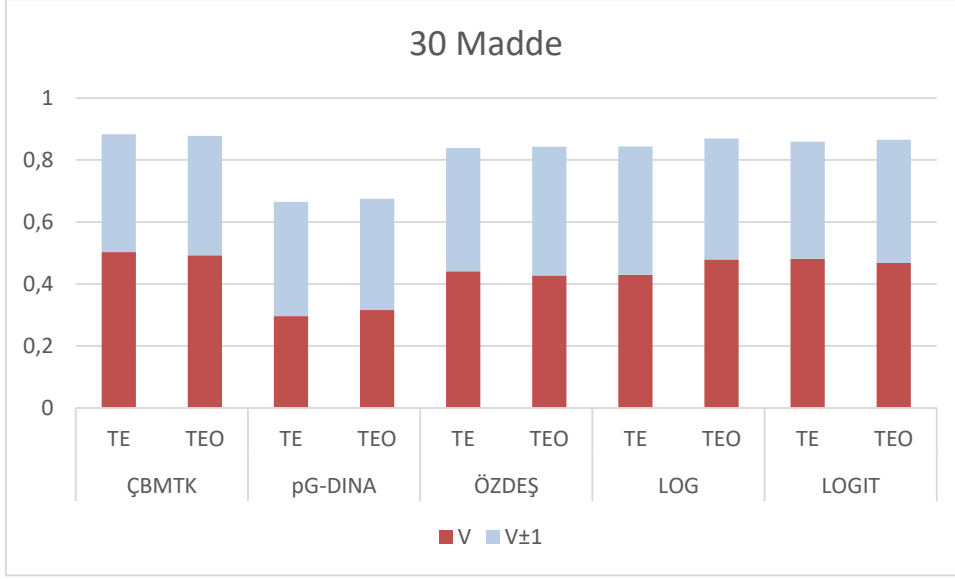


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.16. 15 Maddelik Test Uzunluğu Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.16'da görülen grafik test uzunluğunun 15 madde olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. Daha önce ele alınan diğer koşullarda olduğu gibi fA-M log link haricindeki modellerde telafi edici ve telafi edici olmayan verilerden elde edilen sonuçlar arasında ciddi farklar görülmemektedir. Modellere göre sonuçlara bakıldığında pG-DINA DSO'ları diğer model sonuçlarından bariz bir şekilde düşük olduğu görülmektedir. Tam vektör DSO'ları tüm modeller için %40'larda iken bu oran pG-DINA için %27 düzeylerine düşmektedir. Aradaki fark bir boyuttan elde edilen sonucun  $\pm 1$  seviye yaklaşık olmasına izin verildiğinde %20'lere çıkmaktadır. Bu durumda pG-DINA için elde edilen DSO 0,60 iken diğer modellerde 0,80 civarlarındadır.

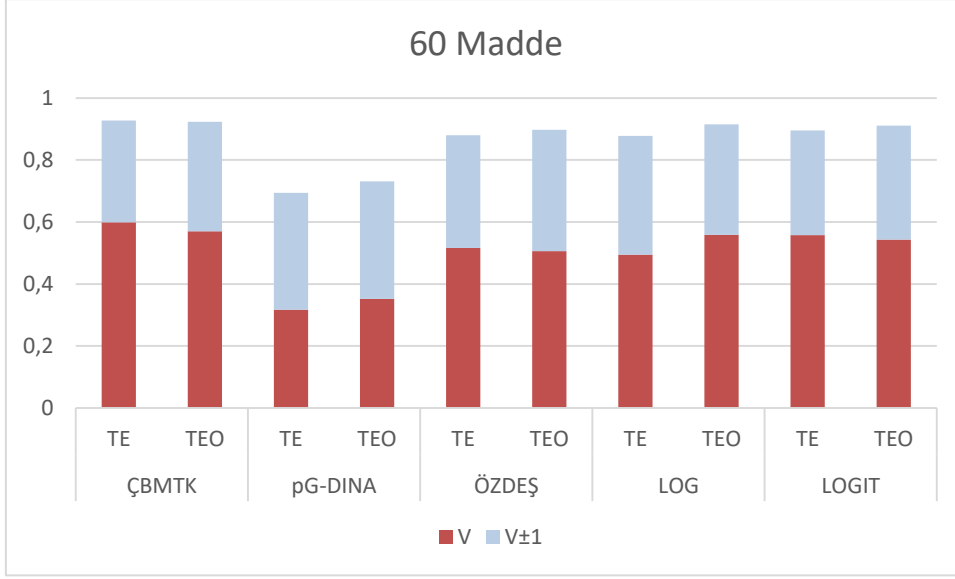




Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.17. 30 Maddelik Test Uzunluğu Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.17 testteki test uzunluğunun 30 madde olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. 15 maddelik testlere ait bir önceki grafikte görülen fA-M log link ve pG-DINA'da telafi edici olmayan verilerin analizi lehine olan fark, biraz daha artmıştır. Sonuçlar 15 maddelik testlere ait DSO'lar ile kıyaslandığında pG-DINA hariç diğer modellerden elde edilen tam vektör sonuçlarında 0,07-0,10'luk artış görülmektedir. Bu artış ÇBMTK sonuçlarında biraz daha belirgindir. Bu sayede grafikte ÇBMTK sonuçlarının fA-M sonuçlarının daha fazla önüne geçtiği görülmektedir. pG-DINA sonuçlarındaki artış ise 0,02-0,03 ile düşük seviyededir. Bu yüzden pG-DINA ile diğer modeller arasındaki farkın biraz daha belirginleştiği görülmektedir. Grafik  $V \pm 1$  yönüyle incelendiğinde, pG-DINA'da önceki koşul olan 15 maddelik test koşulundan daha iyi sonuçlar elde edilirken, ÇBMTK'de bu durum tam tersidir. FA-M'da ise değişimler daha kısmi ve yönü genellemeye imkân vermeyecek şekilde değişkendir. pG-DINA ile diğer modeller arasındaki DSO farkı  $V \pm 1$  ile birlikte genel olarak incelendiğinde, farkın bir önceki koşula göre fazla değişmediği, yine 0,20 seviyelerinde olduğu görülmektedir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.18. 60 Maddelik Test Uzunluğu Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.18'deki grafik test uzunluğunun 60 madde olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. 30 maddelik testlere ait bir önceki grafikte görülen fA-M log link telafi edici olmayan verilerin analizi lehine olan fark, biraz daha artmıştır. pG-DINA'da benzer durum söz konusudur. Sonuçlar 30 maddelik testlere ait DSO'lar ile kıyaslandığında pG-DINA hariç diğer modellerden elde edilen tam vektör sonuçlarında 0,07-0,10'luk artış görülmektedir. Bu artış yine ÇBMTK sonuçlarında biraz daha belirgindir. Bu sayede grafikte ÇBMTK sonuçlarının fA-M sonuçlarına göre bir miktar ön plana çıktığı görülmektedir. pG-DINA sonuçlarındaki artış ise 0,02-0,03 ile düşük seviyedir. Bu yüzden pG-DINA ile diğer modeller arasındaki farkın biraz daha belirginleştiği görülmektedir. Grafik  $V \pm 1$  yönüyle incelendiğinde, pG-DINA'da önceki koşul olan 30 maddelik test koşulundan daha iyi sonuçlar elde edilirken, ÇBMTK ve fA-M'da bu durum tam tersidir. FA-M'daki ise değişimler daha kısmi olduğu gözlenmiştir. pG-DINA ile diğer modeller arasındaki DSO farkı  $V \pm 1$  ile birlikte genel olarak incelendiğinde, farkın bir önceki koşula göre fazla değişmediği, yine 0,17 seviyelerinde olduğu görülmektedir.

#### 4.1.4. Yetenekler Arası Korelasyon Koşullarına Göre Doğru Sınıflama Oranları

Birinci alt problemin dördüncü ve son koşulu olan yetenekler arası korelasyon için iki ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. Yetenekler arası korelasyonun olmadığı durum için 0, yeteneklerin ilişkili olduğu durum için ise 0,6

korelasyon katsayısı kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda korelasyonun olmadığı durum için değerler ele alınmış ve korelasyonun 0,6 olduğu durum için elde edilen değerler ise 0 korelasyon değerleri ile karşılaştırılarak değerlendirilmiştir.

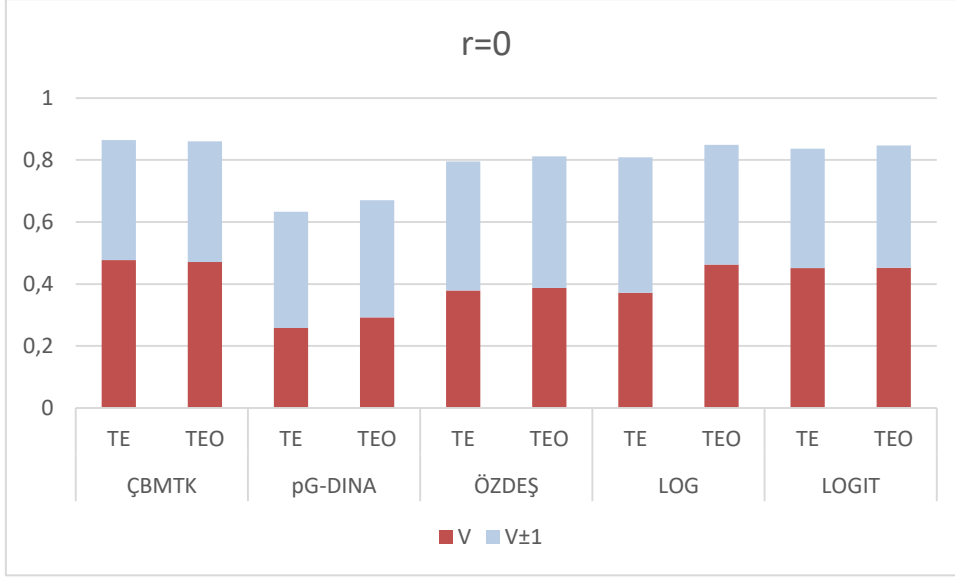
**Tablo 4.10: Yetenekler Arası Korelasyona Göre ÇBMTK Verisi DSO'ları**

<i>r</i>		ÇBMTK		pG-DINA		ÖZDEŞ		LOG		LOGIT	
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1
Vektör	TE	0,48	0,39	0,26	0,38	0,38	0,42	0,37	0,44	0,45	0,39
	0 TEO	0,47	0,39	0,29	0,38	0,39	0,42	0,46	0,39	0,45	0,39
	TE	0,53	0,36	0,33	0,35	0,50	0,36	0,48	0,37	0,51	0,36
	0,6 TEO	0,50	0,37	0,34	0,34	0,47	0,39	0,49	0,38	0,48	0,38
Nitelik	TE	0,68	0,30	0,52	0,36	0,61	0,34	0,61	0,35	0,65	0,32
	0 TEO	0,68	0,30	0,54	0,35	0,62	0,34	0,67	0,30	0,66	0,31
	TE	0,72	0,27	0,55	0,36	0,69	0,30	0,68	0,31	0,70	0,29
	0,6 TEO	0,70	0,29	0,55	0,35	0,67	0,31	0,69	0,30	0,68	0,31

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.10 yetenekler arası korelasyon koşullarına göre vektör ve nitelik bazında elde edilen DSO'ları göstermektedir. Vektör sonuçları grafikler eşliğinde inceleneceğinden tablo üzerinden nitelik için elde edilen değerler üzerinde durulacaktır.

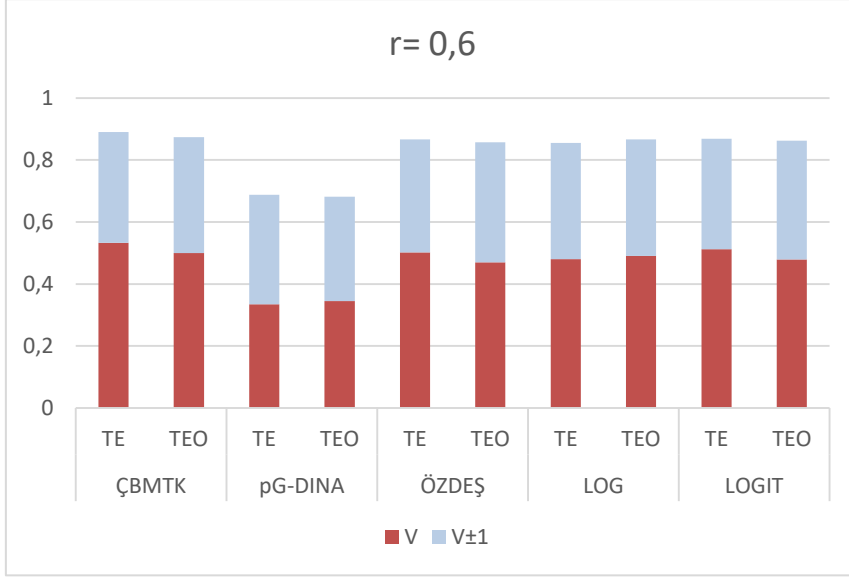
Sonuçları gösteren değerler incelendiğinde yetenekler arası korelasyon artışının tüm modellerde DSO değerlerini artırdığı görülmektedir. Tam ve ±1 sütunları toplandığında %100'e çok yakın doğru sınıflama gerçekleşebilmektedir. Tüm modellerde telafi edici veri analiz sonuçlarının telafi edici olmayan sonuçlara göre korelasyon artışından çok daha etkilendiği görülmektedir, özellikle fA-M'larda görülen artış dikkat çekmektedir. Bunun dışındaki artışlar ise çalışma kapsamındaki diğer koşullarda görülen artışlardan daha düşüktür.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.19. Yetenekler Arası 0 Korelasyon Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.19'da görülen grafik yetenekler arası korelasyonun 0 olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. pG-DINA ve fA-M log link haricindeki modellerden telafi edici ve telafi edici olmayan verilerden elde edilen sonuçlar arasında farkların düşük olduğu görülmektedir. FA-M log linkinde görülen fark, diğer koşulların ilk düzeylerinde görülen farklardan daha yüksektir. Modellere göre sonuçlara bakıldığında pG-DINA DSO'larının yine diğer model sonuçlarından bariz bir şekilde düşük olduğu görülmektedir. Tam vektör DSO'ları tüm modeller için %40 civarlarında iken pG-DINA için bu oran %25-29 düzeylerine düşmektedir. Aradaki fark bir boyuttan elde edilen sonucun  $\pm 1$  seviye yaklaşık olmasına izin verildiğinde %20'lere çıkmaktadır. Bu durumda pG-DINA için elde edilen DSO 0,60 iken diğer modellerde 0,80 civarlarındadır.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.20. Yetenekler Arası 0,6 Korelasyon Koşulu ÇBMTK Verisi DSO'ları**

Şekil 4.20 yetenekler arası korelasyonun 0,6 olduğu koşullar için vektör temelli ortalama DSO'ları göstermektedir. Bu koşulda telafi edici ve telafi edici olmayan verilerden elde edilen sonuçlar arasında farkların telafi edici veri lehine değiştiği görülmektedir. Hatta fA-M özdeş linki 0 korelasyon analizinde az miktarda telafi edici olmayan veri analizi lehine olan fark, 0,6 korelasyonda daha bariz bir şekilde telafi edici veri analizi lehine dönmüştür. Tam vektör sonuçları korelasyonun 0 olduğu DSO'lar ile kıyaslandığında telafi edici verilerin analizlerinde ÇBMTK hariç diğer modellerden elde edilen sonuçlarında 0,07-0,12 arasında ciddi artış görülmektedir. Bu artış verinin ait olduğu ÇBMTK sonuçlarında biraz daha kısıtlıdır. Telafi edici olmayan verilerin analizinde ise 0,03-0,05 düzeyinde artışlar görülmektedir. Bu artış fA-M özdeş linkinde 0,08 ile daha yüksek düzeydedir. Grafik  $V\pm 1$  yönüyle incelendiğinde diğer koşullarda görülmeyen bir sonuç göze çarpmaktadır. Tüm modellerde 0 korelasyon koşuluna göre düşüşler görülmektedir. Bu durum  $\pm 1$ 'lik hatalı kestirimin bir kısmının korelasyonlu yetenek sayesinde tam doğru olarak kestirildiğini göstermektedir. pG-DINA ile diğer modeller arasındaki DSO farkı  $V\pm 1$  ile birlikte incelendiğinde bir önceki koşula göre değişmediği ve 0,20 seviyelerinde olduğu söylenebilir.

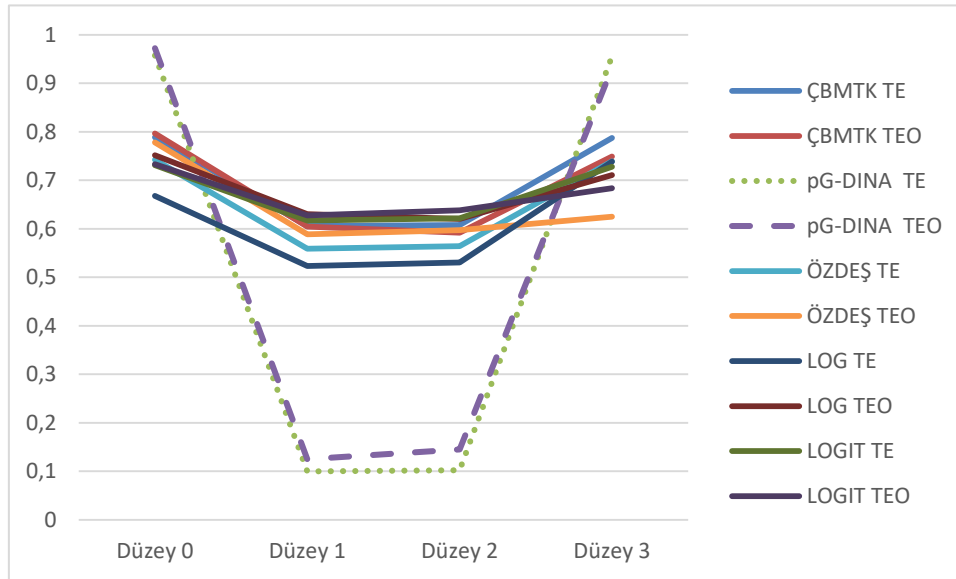
#### 4.1.5. Yetenek Düzeylerinde Doğru Sınıflama Oranları

İlk alt problemin son sorusunda modellerin farklı yetenek düzeylerindeki bireyleri doğru sınıflama oranları incelenmiştir. Çıkan sonuçlar Tablo 4.11’de verilmiştir. Ayrıca sonuçların görselleştirilmesi adına grafik de kullanılmıştır.

**Tablo 4.11: Yetenek Düzeylerinde ÇBMTK Verisi DSO'ları**

		Düzye 0	Düzye 1	Düzye 2	Düzye 3
ÇBMTK	TE	0,79	0,61	0,61	0,79
	TEO	0,80	0,60	0,59	0,75
pG-DINA	TE	0,96	0,10	0,10	0,96
	TEO	0,97	0,12	0,15	0,93
ÖZDEŞ	TE	0,74	0,56	0,56	0,74
	TEO	0,78	0,59	0,60	0,62
LOG	TE	0,67	0,52	0,53	0,74
	TEO	0,75	0,63	0,62	0,71
LOGIT	TE	0,73	0,62	0,62	0,73
	TEO	0,73	0,63	0,64	0,68

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.21. Yetenek Düzeylerine Göre ÇBMTK Verisi DSO'ları**

Şekil 4.21’de görülen sonuçlarda göze çarpan ilk durum pG-DINA’nın düzey 0 ve düzey 3’deki bireyleri çok yüksek bir oranda doğru sınıflamasına rağmen düzey 1 ve düzey 2’de bulunan bireyleri neredeyse hiç doğru sınıflayamamasıdır. pG-DINA haricindeki modellerde de düzey 0 ve 3 deki bireyler düzey 1 ve 2’deki bireylerden daha düşük bir yüzdede doğru sınıflanmaktadır. Ancak pG-DINA’da bu fark adeta

uçuruma dönüşmektedir. Diğer modellerde ortalama 0,18 düzeyinde olan düzey 0 ve 3 ile düzey 1 ve 2 DSO'ları arasındaki fark, pG-DINA'da 0,85'e çıkmaktadır. Burada yer verilmeyen sonuçlarda görülen durumu netliğe kavuşturmaktadır. PG-DINA orta seviyelerdeki bireyleri kendisine yakın olan uç seviyelerde kestirmektedir. Bu durumda seviye 1'deki bireyler çok büyük bir oranda seviye 0'da, seviye 2'deki bireyler ise çok büyük bir oranda seviye 3'de kestirilmektedir. Seviye 0 ve 3'deki bireyler ise mükemmel yakın bir şekilde kendi buldukları uç seviyelerde kestirilmektedir. Grafikte fA-M özdeş linkinin telafi edici olmayan veri kestirimi haricindeki diğer kestirimler simetrik görüntüye sahip oldukları görülmektedir. Bu analiz hariç model analizlerinde seviye 0 seviye 3 ile benzer, seviye 1 seviye 2 ile benzer DSO'lara sahiptir. Grafik izlendiğinde önce düşüş, sonra yatay seyir en son ise artış görülmektedir.

#### **4.2. BTM Verileri Analizlerine İlişkin Bulgular ve Tartışma**

Bu bölümde pG-DINA verisinin pG-DINA, ÇBMTK ve özdeş, log ve logit fonksiyonlarındaki fA-M'lara ve fA-M verilerinin ilgili fA-M fonksiyonu, ÇBMTK ve pG-DINA'ya uygulanması ile elde edilen sonuçlar sırasıyla madde ayırt edicilik, madde yapısı, test uzunluğu ve yetenekler arasındaki korelasyon koşullarına göre ele alınmıştır.

##### **4.2.1. Madde Ayırt Edicilik Koşullarına Göre Doğru Sınıflama Oranları**

İkinci alt problemin ilk koşulu olan madde ayırt edicilik indeksi için 3 ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. Düşük madde ayırt edicilik için  $a \sim U(0,6-0,8)$ , orta düzey için  $a \sim U(0,9-1,1)$  ve yüksek düzey içinse  $a \sim U(1,5-1,7)$  aralıkları kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda düşük düzey değerleri ele alınmış, orta düzey için elde edilen değerler düşük düzey değerleri ile ve yüksek düzey için elde edilen değerler ise orta düzey değerleri ile karşılaştırılarak değerlendirilmiştir.

**Tablo 4.12: Madde Ayırt Ediciliğe Göre pG-DINA Verisi DSO'ları**

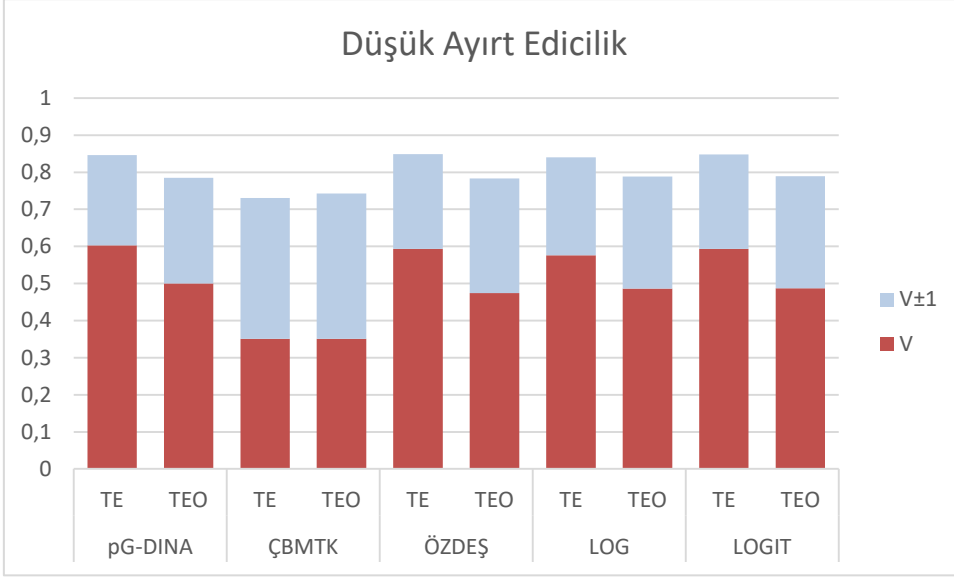
		Vektör						Nitelik					
		Düşük		Orta		Yüksek		Düşük		Orta		Yüksek	
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1
pG-DINA	TE	0,60	0,24	0,72	0,19	0,84	0,12	0,76	0,17	0,84	0,12	0,91	0,08
pG-DINA	TEO	0,50	0,28	0,64	0,23	0,79	0,15	0,70	0,21	0,79	0,16	0,88	0,10
ÇBMTK	TE	0,35	0,38	0,44	0,35	0,54	0,30	0,56	0,37	0,63	0,32	0,70	0,27
ÇBMTK	TEO	0,35	0,39	0,48	0,35	0,60	0,29	0,58	0,36	0,67	0,28	0,76	0,22
ÖZDEŞ	TE	0,59	0,26	0,72	0,20	0,82	0,14	0,76	0,18	0,83	0,13	0,90	0,09
ÖZDEŞ	TEO	0,47	0,31	0,60	0,26	0,71	0,20	0,68	0,24	0,76	0,18	0,83	0,13
LOG	TE	0,58	0,26	0,69	0,21	0,80	0,15	0,75	0,19	0,82	0,14	0,89	0,09
LOG	TEO	0,49	0,30	0,63	0,24	0,78	0,16	0,68	0,23	0,78	0,17	0,87	0,10
LOGIT	TE	0,59	0,25	0,72	0,19	0,83	0,13	0,75	0,18	0,84	0,13	0,91	0,08
LOGIT	TEO	0,49	0,30	0,63	0,25	0,78	0,16	0,68	0,23	0,78	0,17	0,87	0,10

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.12'de farklı madde ayırt edicilik düzeylerine sahip pG-DINA verisinin pG-DINA, ÇBMTK ve fA-M'lara uygulanması sonucu vektör ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiğinde verinin ait olduğu pG-DINA'ya çok yakın bir düzeyde fA-M'ların başarılı kestirimleri dikkat çekmektedir. ÇBMTK ise diğer modellerden daha başarısız olduğu görülmektedir. Madde ayırt edicilik arttıkça DSO'lar tam ve tama ilave edilen  $\pm 1$  hatalı sonuçlarda artmaktadır. Bu şekilde %100'e yaklaşan DSO'lar görmek mümkündür. ÇBMTK haricindeki modellerin telafi edici verilerin analizinden elde edilen sonuçlar biraz daha iyidir.

İlk alt problemde aynı senaryolar için doğru model olan ÇBMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiğinde, pG-DINA verisinin çok daha başarılı bir şekilde kestirildiği görülmektedir. Doğru modeller olan ÇBMTK ve pG-DINA'ya ait tam vektör DSO'larına bakıldığında düşük düzey madde ayırt edicilikte telafi edici veri için 0,17, telafi edici olmayan veri içinse 0,1'lik fark olduğu görülmektedir. Bu farkın madde ayırt edicilik artışında daha da arttığı görülmektedir. Her iki alt problemde de uyarlanan modeller olan fA-M'larda da benzer farklar görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA ve burada en düşük performansı gösteren ÇBMTK'nin arasında da benzer fark bulunmaktadır.

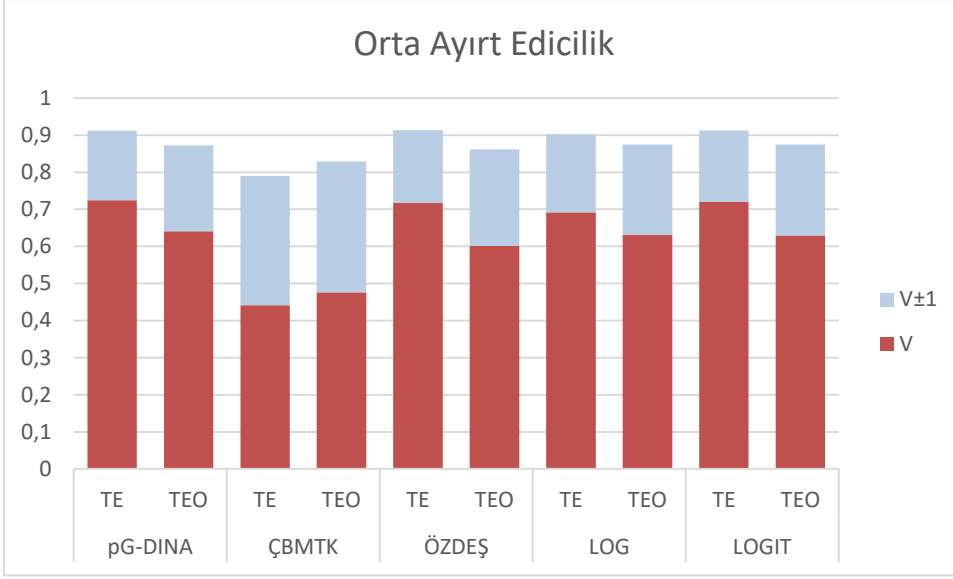




Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

#### **Şekil 4.22. Düşük Düzey Madde Ayırt Edicilik Koşulu pG-DINA Verisi DSO'ları**

Madde ayırt ediciliğin düşük ( $a \sim U(0,6-0,8)$ ) olduğu pG-DINA verisinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.22'de görülmektedir. ÇBMTK hariç diğer modellerin telafi edici veri türünde daha iyi sonuçlar verdiği görülmektedir. ÇBMTK ayrıca elde edilen sonuçlara bakıldığında daha düşük olan DSO'lar nedeniyle de diğer modellerden ayrılmaktadır. Diğer modellerde telafi edici veri için gerçekleşen 0,60'lık tam doğrulama oranı ÇBMTK'de 0,35 olarak gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise ÇBMTK yine 0,35 iken diğer modeller 0,50'lik doğru sınıflamaya sahiptir. Bu sonuçlar  $V \pm 1$  hata payı eklendiğinde ÇBMTK'de 0,74 civarlarındayken diğer modellerde telafi edici veri için 0,84, telafi edici olmayan veri için ise 0,79 dolaylarındadır. Ayrıca  $V \pm 1$  eklendiğinde ÇBMTK haricindeki diğer modellerde aynı veri türünde görülen küçük farklar ortadan kaybolmuştur. Buna ilaveten aynı modelin telafi edici ve telafi edici olmayan veri analizleri arasındaki farkın da azaldığı görülmektedir.

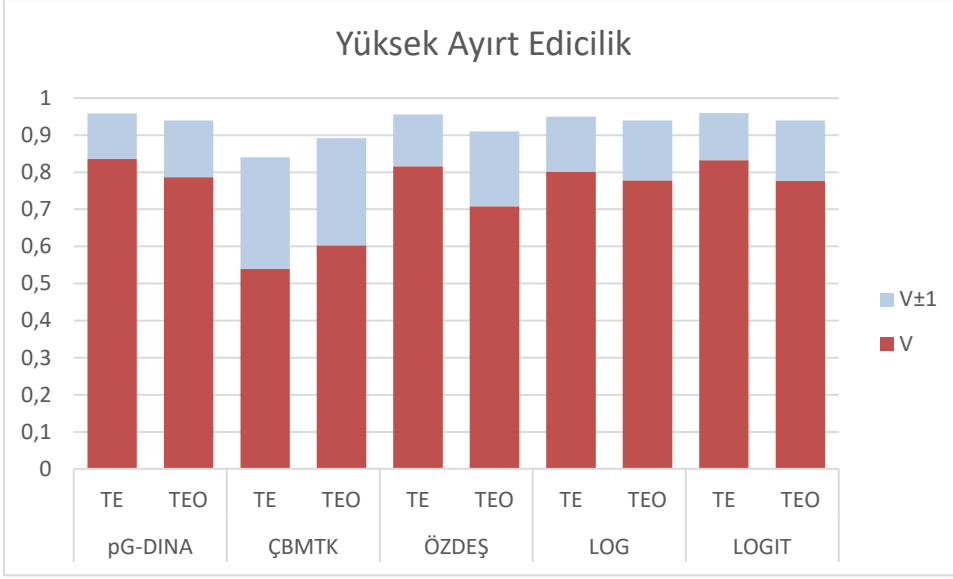


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.23. Orta Düzey Madde Ayırt Edicilik Koşulu pG-DINA Verisi DSO'ları**

Şekil 4.23'deki grafik değerleri ayırt ediciliğin orta ( $a \sim U(0,9-1,1)$ ) olduğu durumlardaki pG-DINA verisinin analizi sonucu elde edilen DSO'lara aittir. Değerler düşük ayırt edicilik koşulunda elde edilen değerlerle karşılaştırıldığında, telafi edici verinin analiz sonucu lehine olan durum farkın bir miktar azalmasına rağmen devam ettiği görülmektedir. ÇBMTK buna istisna olmaya devam etmektedir ve telafi edici olmayan veri lehine olan fark biraz daha belirgin hale gelmiştir. ÇBMTK haricindeki modellerle telafi edici verilerde 0,7 düzeylerinde DSO elde edilirken, telafi edici olmayan verilerde 0,6 düzeyi biraz aşılmaktadır. ÇBMTK'de ise 0,44-0,47 seviyelerine ulaşılmıştır. Tam vektör tahmininde 0,09-0,14 arasında artışlar gözlenmektedir. Artışların tüm modellerde benzer bir oranda gerçekleşmesi dikkat çekmektedir.

$V \pm 1$  değerlerine bakıldığında düşük madde ayırt edicilik sonuçlarına göre 0,03-0,05 düzeyinde düşüş görülmektedir. Madde kalitesinin düşükten orta düzeye çıkmasıyla  $V \pm 1$  değerlerinin bir kısmı tam doğru tahmine kaymakta ve  $V \pm 1$  kestirimi ise kaybettiği rakamdan bir miktar daha az yeni kestirim elde etmektedir. Tam ve  $\pm 1$  değerleri birlikte değerlendirildiğinde ÇBMTK dışındaki modellerde telafi edici veri için 0,9'un biraz üstünde, telafi edici olmayan veri için ise 0,9'un bir miktar altında değerler elde edilmiştir. ÇBMTK da ise bu değerler telafi edici olmayan veri için 0,8'in biraz üstünde, telafi edici veri için ise 0,8'in bir miktar altında gerçekleşmiştir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.24. Yüksek Düzey Madde Ayırt Edicilik Koşulu pG-DINA Verisi DSO'ları**

Şekil 4.24'te yüksek ( $a \sim U(1,5-1,7)$ ) madde ayırt ediciliğe sahip pG-DINA verilerinin analizi sonucu elde edilen DSO'lar bulunmaktadır. Değerler orta düzey madde ayırt edicilik koşulundaki sonuçlar ile karşılaştırıldığında tam ve toplam DSO'ların arttığı görülmektedir. Bu artışın telafi edici olmayan veri türü için daha yüksek olduğu görülmektedir. Böylece ÇBMTK'de telafi edici olmayan veri lehine olan fark biraz daha artarken, diğer modellerde telafi edici veri lehine olan fark azalmıştır. Telafi edici lehine olan bu fark, fA-M özdeş linki dışındaki modellerde çok daha küçük değerlere ulaşmıştır. Tam vektör DSO'larında meydana gelen artış telafi edici veri türünde 0,1 düzeylerinde iken telafi edici olmayan veri türünde 0,15 düzeylerindedir. Artışın tüm modellerde birbirine benzer değerlerde olduğu görülmektedir.

$V \pm 1$  değerlerinde meydana gelen değişim yine tüm modeller için benzerdir, tüm modellerde 0,05-0,07 düşüş görülmüştür. ÇBMTK hariç diğer modellerde 0,83'e ulaşan tam DSO'lar pG-DINA verisinin ÇBMTK hariç modellerde yüksek bir doğrulukla tahmin edilebileceğini göstermiştir.

**Tablo 4.13: Madde Ayırt Ediciliğe Göre fA-M Verileri DSO'ları**

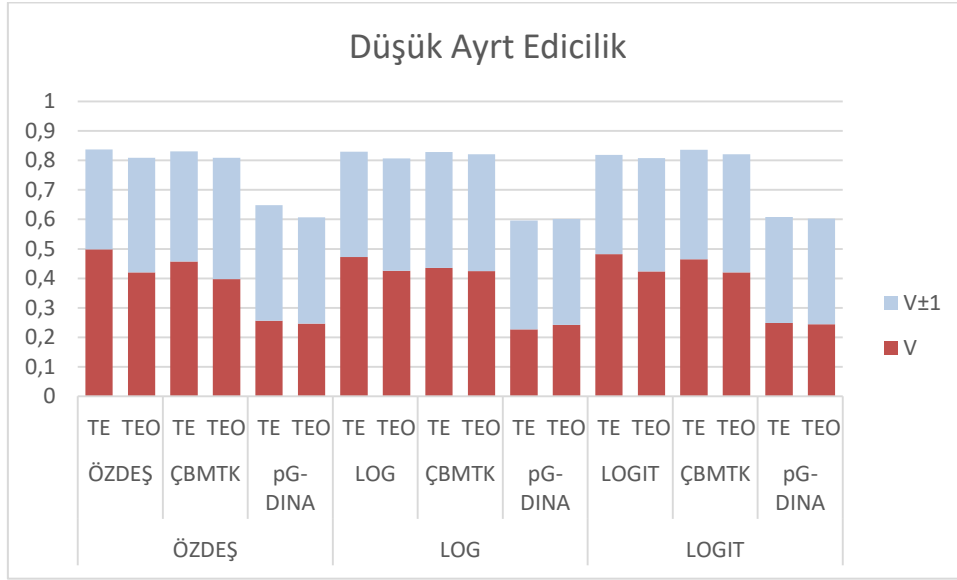
		Vektör						Nitelik						
		Düşük		Orta		Yüksek		Düşük		Orta		Yüksek		
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	
KİMLİK	ÖZDEŞ	TE	0,50	0,34	0,61	0,29	0,74	0,21	0,68	0,29	0,76	0,23	0,84	0,15
	ÖZDEŞ	TEO	0,42	0,39	0,54	0,35	0,68	0,27	0,63	0,32	0,72	0,26	0,81	0,18
	ÇBMTK	TE	0,46	0,37	0,55	0,34	0,65	0,28	0,65	0,33	0,72	0,27	0,79	0,21
	ÇBMTK	TEO	0,40	0,41	0,49	0,38	0,61	0,32	0,62	0,35	0,69	0,29	0,78	0,22
	pG-DINA	TE	0,26	0,39	0,29	0,44	0,34	0,47	0,51	0,37	0,54	0,38	0,59	0,37
	pG-DINA	TEO	0,25	0,36	0,28	0,40	0,37	0,42	0,50	0,36	0,53	0,37	0,61	0,34
	LOG	TE	0,47	0,36	0,58	0,32	0,70	0,25	0,67	0,29	0,75	0,23	0,83	0,17
	LOG	TEO	0,43	0,38	0,55	0,33	0,70	0,25	0,64	0,32	0,73	0,25	0,83	0,17
	ÇBMTK	TE	0,44	0,39	0,52	0,37	0,61	0,32	0,64	0,33	0,71	0,28	0,78	0,22
	ÇBMTK	TEO	0,42	0,40	0,54	0,35	0,68	0,27	0,64	0,33	0,72	0,26	0,82	0,18
LOG	pG-DINA	TE	0,23	0,37	0,24	0,40	0,29	0,42	0,49	0,36	0,51	0,37	0,55	0,37
	pG-DINA	TEO	0,24	0,36	0,28	0,39	0,37	0,41	0,49	0,36	0,53	0,37	0,61	0,34
	LOGIT	TE	0,48	0,34	0,60	0,27	0,73	0,19	0,66	0,31	0,74	0,25	0,83	0,17
	LOGIT	TEO	0,42	0,38	0,55	0,34	0,70	0,24	0,64	0,32	0,72	0,26	0,82	0,17
LOGIT	ÇBMTK	TE	0,47	0,37	0,57	0,32	0,68	0,26	0,66	0,32	0,73	0,26	0,81	0,19
	ÇBMTK	TEO	0,42	0,40	0,53	0,36	0,67	0,27	0,63	0,34	0,72	0,27	0,81	0,18
	pG-DINA	TE	0,25	0,36	0,27	0,39	0,30	0,42	0,50	0,36	0,52	0,37	0,56	0,36
	pG-DINA	TEO	0,24	0,36	0,28	0,39	0,37	0,42	0,50	0,36	0,53	0,37	0,61	0,34

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.13'de farklı madde ayırt edicilik düzeylerine sahip özdeş, log ve logit linkindeki fA-M verilerinin pG-DINA, ÇBMTK ve verinin üretildiği fA-M'lara uygulanması sonucu vektör ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiğinde verinin ait olduğu fA-M kestirimine yakın bir düzeyde ÇBMTK'nin başarılı kestirimi dikkat çekmektedir. PG-DINA ise diğer modellerden daha geride olduğu görülmektedir. Madde ayırt edicilik arttıkça DSO'lar tam ve tama ilave edilen  $\pm 1$  hatalı sonuçlarda artmaktadır. %100'e yaklaşan DSO'lar görmek mümkündür. Genel olarak bakıldığında telafi edici verilerin analizinden elde edilen sonuçların telafi edici olmayanlara göre bir miktar daha iyi olduğu görülmüştür.

İlk alt problemde aynı senaryolar için doğru model olan ÇBMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiğinde, fA-M verilerinin daha başarılı bir şekilde kestirildiği görülmektedir. Ancak ikinci alt problemdeki pG-DINA verisi kestirimi ile karşılaştırıldığında ise daha düşük doğrulukta kestirim yapıldığı görülmektedir. FA-M'lara ait tam vektör DSO'larının düşük düzey madde ayırt edicilikte, ilk alt problemdeki ÇBMTK sonuçlarından telafi edici veri için 0,04-0,07,

telafi edici olmayan veri içinse 0,02-0,03'lük oranda daha iyi olduğu görülmektedir. Yine aynı koşul için ikinci alt problemdeki pG-DINA kestirimi fA-M'lardan telafi edici veri için 0,1-0,13, telafi edici olmayan veri içinse 0,07-0,08'lik oranda daha iyi olduğu görülmektedir. Bu farkların madde ayırt edicilik artışında ÇBMTK için daha da arttığı, pG-DINA için pek değişmediği görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA fA-M verilerinde de en düşük performansı göstermiştir.

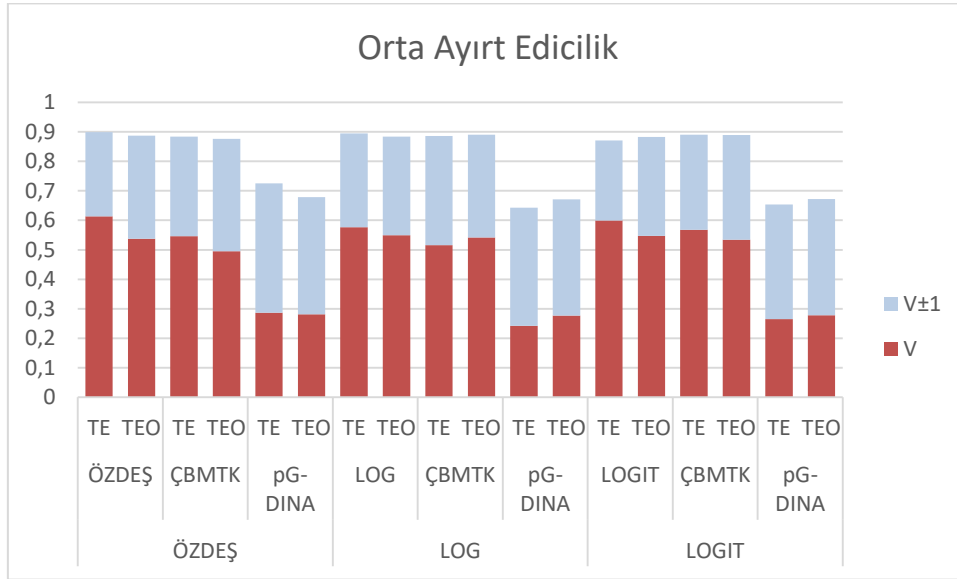


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.25. Düşük Düzey Madde Ayırt Edicilik Koşulu fA-M Verisi DSO'ları**

Madde ayırt ediciliğin düşük ( $a \sim U(0,6-0,8)$ ) olduğu fA-M verilerinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.25'de görülmektedir. FA-M modellerinde daha belirgin şekilde telafi edici veri lehine olan fark, ÇBMTK analizlerinde daha küçüktür. Bu fark pG-DINA analizlerinde daha da azalmakta ve log linki analizinde telafi edici olmayan veri, daha iyi sonuç vermektedir. ÇBMTK'nin verinin üretildiği fA-M model sonuçlarına yakın sonuçlar ürettiği görülmektedir. FA-M modelleri ile ÇBMTK arasındaki fark son derece küçüktür. FA-M modelleri kendi aralarında değerlendirildiğinde sonuçlar herhangi bir linki ön plana çıkarmamaktadır. PG-DINA ise her üç modele ait veride de en kötü oranlara sahiptir. Diğer modellerde telafi edici veri için gerçekleşen 0,48'lik tam doğrulama oranı pG-DINA'da 0,25 olarak gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise pG-DINA yine 0,25 iken diğer modeller 0,42'lik doğru sınıflamaya sahiptir. Bu sonuçlar

$V \pm 1$  hata payı eklendiğinde ÇBMTK'de ve fA-M'da 0,80 civarlarındayken pG-DINA'da ise 0,60 dolaylarındadır.



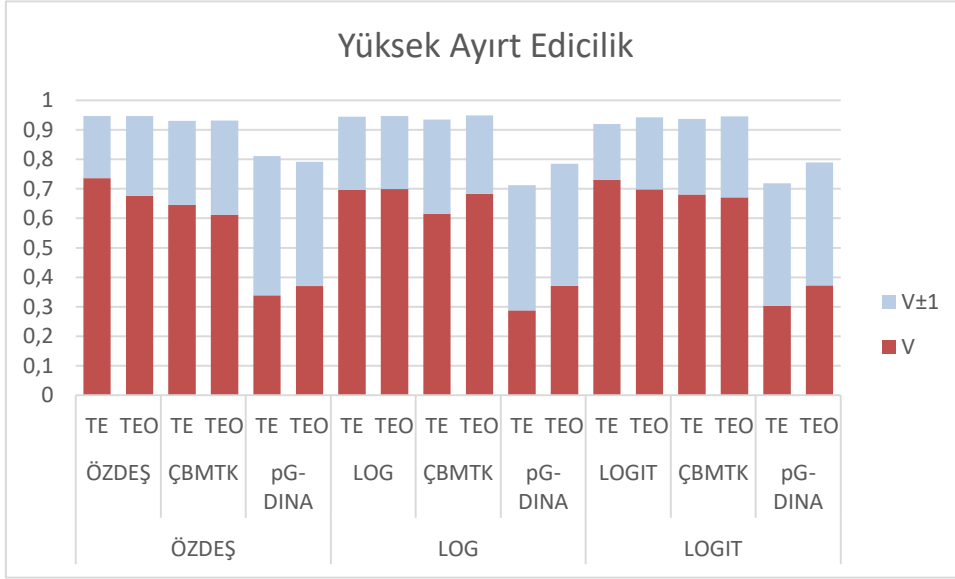
Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.26. Orta Düzey Madde Ayırt Edicilik Koşulu fA-M Verisi DSO'ları**

Şekil 4.26'da görülen grafikteki değerler ayırt ediciliğin orta ( $a \sim U(0,9-1,1)$ ) olduğu durumlardaki fA-M verilerinin analizi sonucu elde edilen DSO'lara aittir. Değerler düşük ayırt edicilik koşulunda elde edilen değerlerle karşılaştırıldığında, telafi edici verinin analiz sonucu lehine olan farkın log linki haricinde devam ettiği görülmektedir. Log linkinde telafi edici olmayan veride daha iyi sonuçlar alınmaya başlanmıştır. Ayrıca pG-DINA'da telafi edici olmayan veri lehine olan fark biraz daha belirgin hale gelmiştir. pG-DINA haricindeki modellerle telafi edici verilerde 0,52-0,62 düzeylerinde DSO elde edilirken, bu rakam telafi edici olmayan verilerde 0,55-0,49 düzeyindedir. pG-DINA'da ise ancak 0,24-0,28 seviyelerine ulaşılmıştır. ÇBMTK ve fA-M'larda tam vektör tahmininde 0,08-0,13 arasında artışlar gözlenmekteyken bu rakam pG-DINA sadece 0,01-0,04 aralığındadır. PG-DINA ile diğer modeller arasındaki farkın arttığı görülmektedir.

$V \pm 1$  değerlerine bakıldığında düşük madde ayırt edicilik sonuçlarına göre ÇBMTK ve fA-M'larda 0,03-0,07 düzeyinde düşüş görülmektedir. PG-DINA'da eskiye nazaran 0,03-0,05'lik artış görülmektedir. Tam ve  $\pm 1$  değerleri birlikte değerlendirildiğinde telafi edici ve telafi edici olmayan sonuçların birbirine yaklaştığı

görülmüştür. Toplamda pG-DINA dışındaki modellerde 0,9'un biraz altında değerler elde edilmiştir. pG-DINA'da ise bu değerler 0,64-0,72 aralığında gerçekleşmiştir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.27. Yüksek Düzey Madde Ayırt Edicilik Koşulu fA-M Verisi DSO'ları**

Şekil 4.27'de yüksek ( $a \sim U(1,5-1,7)$ ) madde ayırt ediciliğe sahip fA-M verilerinin analizi sonucu elde edilen DSO'lar bulunmaktadır. Değerler orta düzey madde ayırt edicilik koşulundaki sonuçlar ile karşılaştırıldığında tam ve toplam DSO'ların arttığı görülmektedir. Bu artışın telafi edici olmayan veri türü için biraz daha yüksek olduğu görülmektedir. Böylece log linkinde ve pG-DINA'da telafi edici olmayan veri lehine olan fark biraz daha artarken, diğer modellerde telafi edici veri lehine olan fark azalmıştır. FA-M ve ÇBMTK'de tam vektör DSO'larında meydana gelen artış telafi edici veri türünde 0,09-0,13 düzeylerinde iken telafi edici olmayan veri türünde 0,12-0,15 düzeylerindedir. pG-DINA'da bu artışın kısıtlı olduğu ve telafi edici veri türünde 0,03-0,05, telafi edici olmayan veri türünde 0,09 düzeylerinde olduğu görülmüştür.

V±1 değerlerinde meydana gelen değişim ise fA-M ve ÇBMTK'de 0,05-0,1 düzeyinde düşüş, pG-DINA'da ise 0,02-0,03'lük artış yönündedir. pG-DINA hariç diğer modellerde 0,7'lere ulaşan-yaklaşan tam DSO'lar, V±1 ile değerlendirildiğinde 0,9'ları geçmektedir. pG-DINA'da ise yükselen V±1 değerleri sayesinde özellikle özdeş linkinde etkili olmakta ve 0,8 seviyesine yaklaşmaktadır.

#### 4.2.2. Madde Yapısı Oranı Koşullarına Göre Doğru Sınıflama Oranları

İkinci alt problemin ikinci koşulu olan madde yapısı oranı için 3 ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. Karmaşık olarak belirtilen test için %80 karmaşık, %20 basit yapıya sahip maddeler, eşit olarak belirtilen test için %50 karmaşık, %50 basit yapıya sahip maddeler ve basit olarak belirtilen test için %20 karmaşık, %80 basit yapıya sahip maddeler kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda karmaşık değerleri ele alınmış, eşit düzey için elde edilen değerler karmaşık değerleri ile ve basit için elde edilen değerler ise eşit düzey değerleri ile karşılaştırılarak değerlendirilmiştir.

**Tablo 4.14: Madde Yapısı Oranına Göre pG-DINA Verisi DSO'ları**

		Vektör						Nitelik					
		Karmaşık		Eşit		Basit		Karmaşık		Eşit		Basit	
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1
pG-DINA	TE	0,73	0,17	0,72	0,19	0,71	0,20	0,83	0,12	0,84	0,12	0,83	0,13
pG-DINA	TEO	0,60	0,23	0,65	0,22	0,67	0,22	0,76	0,17	0,80	0,15	0,81	0,15
ÇBMTK	TE	0,31	0,32	0,50	0,35	0,52	0,36	0,49	0,41	0,69	0,29	0,72	0,26
ÇBMTK	TEO	0,39	0,35	0,48	0,36	0,56	0,32	0,60	0,33	0,68	0,29	0,73	0,24
ÖZDEŞ	TE	0,70	0,19	0,71	0,20	0,70	0,21	0,82	0,13	0,83	0,13	0,83	0,14
ÖZDEŞ	TEO	0,52	0,28	0,61	0,25	0,66	0,24	0,70	0,21	0,77	0,18	0,80	0,16
LOG	TE	0,68	0,20	0,70	0,21	0,70	0,21	0,80	0,14	0,82	0,14	0,83	0,14
LOG	TEO	0,59	0,24	0,64	0,23	0,66	0,23	0,75	0,18	0,79	0,16	0,80	0,16
LOGIT	TE	0,72	0,17	0,72	0,19	0,71	0,21	0,83	0,12	0,84	0,13	0,83	0,14
LOGIT	TEO	0,59	0,24	0,64	0,23	0,66	0,23	0,75	0,18	0,79	0,17	0,80	0,16

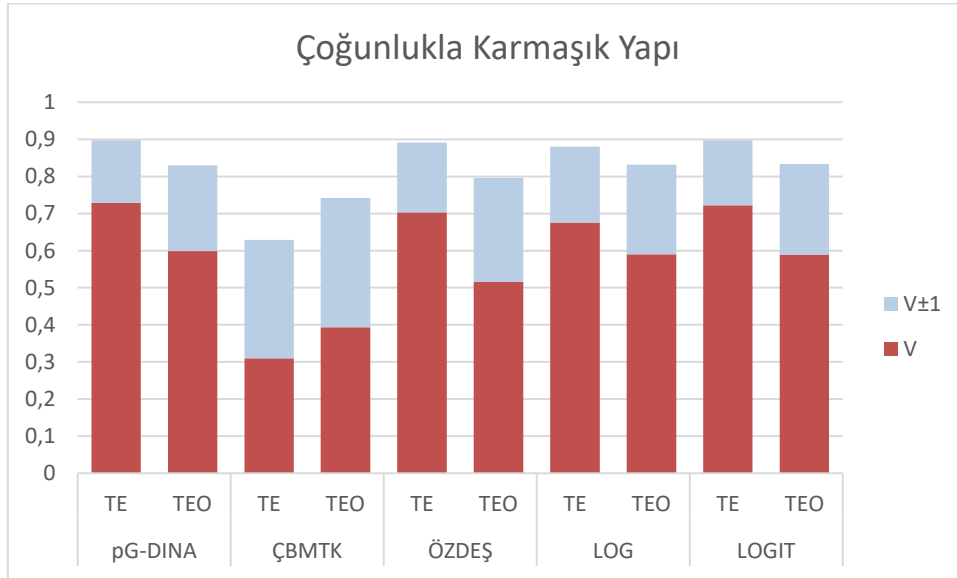
Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.14'de farklı madde yapısı oranlarına sahip pG-DINA verisinin pG-DINA, ÇBMTK ve fA-M'lara uygulanması sonucu vektör ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiğinde verinin ait olduğu pG-DINA'ya çok yakın bir düzeyde fA-M'ların başarılı kestirimleri dikkat çekmektedir. ÇBMTK ise diğer modellerden daha başarısız olduğu görülmektedir. ÇBMTK'de testteki basit yapıdaki madde oranı arttıkça her iki veri türünde de tam DSO'lar artmaktadır. Diğer modellerde artışın çok düşük seviyededir. Küçük de olsa bu artışların telafi edici olmayan veri analizinde olduğu görülmektedir. ÇBMTK haricindeki modellerin telafi edici verilerin analizinden elde edilen sonuçlar biraz daha iyidir.

İlk alt problemde aynı senaryolar için doğru model olan ÇBMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiğinde, pG-DINA verisinin çok daha



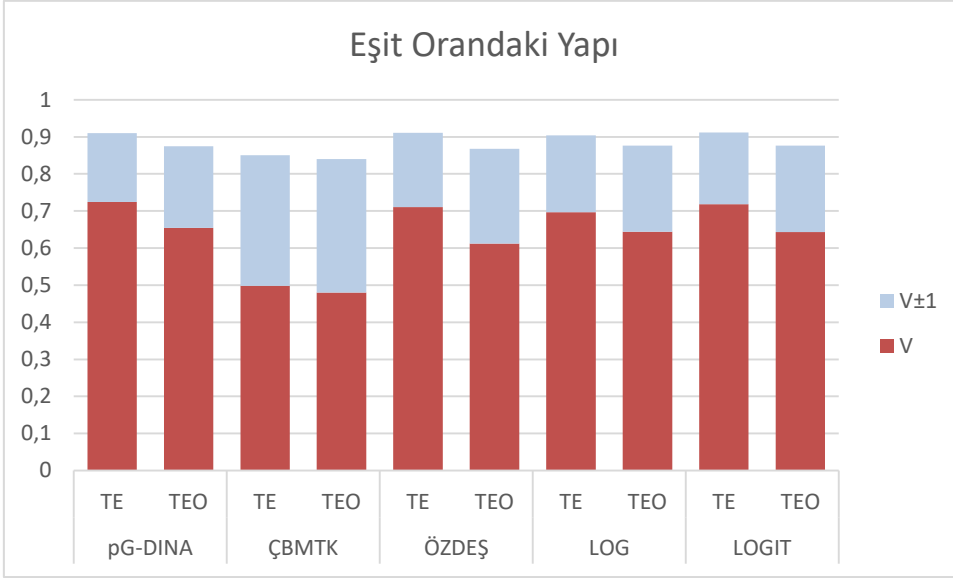
başarılı bir şekilde kestirildiği görülmektedir. Doğru modeller olan ÇBMTK ve pG-DINA'ya ait tam vektör DSO'larına bakıldığında testteki madde yapılarının eşit oranda olduğu telafi edici veri için 0,2, telafi edici olmayan veri içinse 0,16'lık fark olduğu görülmektedir. Her iki alt problemde de uyarlanan modeller olan fA-M'larda da benzer farklar görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA ve burada en düşük performansı gösteren ÇBMTK'nin arasında da benzer fark bulunmaktadır.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.28. Çoğunlukla Karmaşık Yapı Koşulu pG-DINA Verisi DSO'ları**

Testteki maddelerin çoğunun karmaşık yapıda olduğu pG-DINA verisinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.28'deki grafikte görülmektedir. ÇBMTK hariç diğer modellerin telafi edici veri türünde daha iyi sonuçlar verdiği görülmektedir. ÇBMTK ayrıca elde edilen sonuçlara bakıldığında daha düşük olan DSO'lar nedeniyle de diğer modellerden ayrılmaktadır. Diğer modellerde telafi edici veri için gerçekleşen 0,70 civarlarında tam doğrulama oranı ÇBMTK'de 0,31 olarak gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise ÇBMTK 0,39'luk, özdeş linki 0,52'lik ve diğer modeller 0,60 civarında doğru sınıflamaya sahiptir. Bu sonuçlar  $V \pm 1$  hata payı eklendiğinde ÇBMTK'de telafi edici veride 0,63, telafi edici olmayan veri içinse 0,74 düzeyine ulaşmaktadır. Diğer modellerde telafi edici veri için 0,89, telafi edici olmayan veri için ise 0,81 dolaylarındadır. Ayrıca  $V \pm 1$  eklendiğinde aynı modelin telafi edici ve telafi edici olmayan veri analizleri arasındaki farkın da azaldığı görülmektedir.

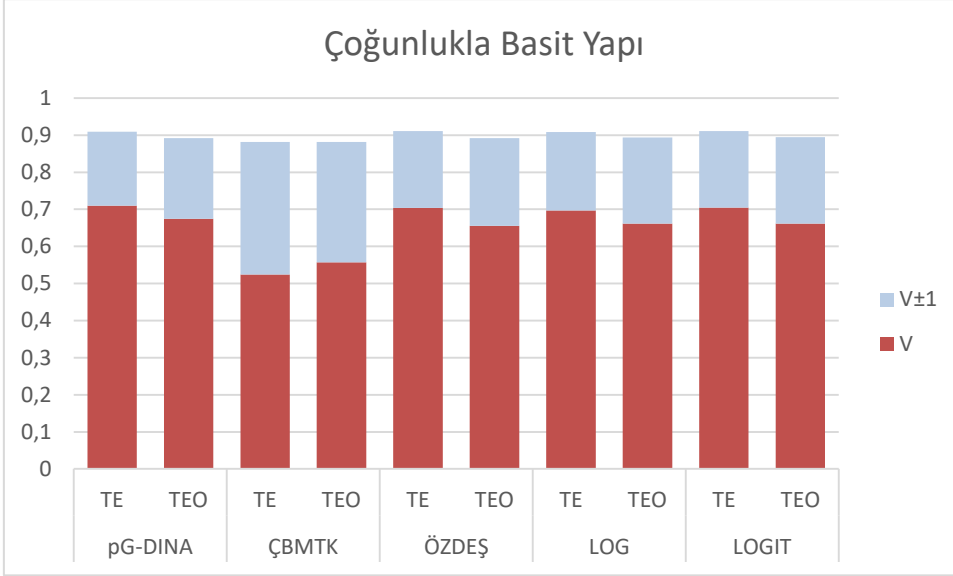


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.29. Eşit Oran Yapı Koşulu pG-DINA Verisi DSO'ları**

Şekil 4.29'da görülen grafikteki yapılarına göre maddelerin testte eşit oranda olduğu durumlardaki pG-DINA verisinin analizi sonucu elde edilen DSO'lara aittir. Değerler çoğunlukla karmaşık yapı koşulunda elde edilen değerlerle karşılaştırıldığında, telafi edici verinin analiz sonucu lehine olan durum farkın bir miktar azalmasına rağmen devam ettiği görülmektedir. Bu gelişmenin tam tersi ÇBMTK'de görülmekte ve telafi edici olmayan veri lehine olan fark kaybolmuş, telafi edici veri oranı bir miktar öne geçmiştir. ÇBMTK haricindeki modellerle telafi edici verilerde 0,7 düzeylerinde DSO elde edilirken, telafi edici olmayan verilerde 0,6 düzeyi biraz aşılmaktadır. ÇBMTK'de ise 0,48-0,5 seviyelerine ulaşılmıştır. Tam vektör tahmininde görülen artışlar ÇBMTK'de özellikle telafi edici olmak üzere her iki veri türünde iken diğer modellerde ciddi değişimlerin telafi edici olmayan veri türünde olduğu görülmektedir.

V±1 değerlerine bakıldığında çoğunlukla karmaşık yapıdaki sonuçlara göre yönleri farklı küçük değişimler görülmektedir. Tam ve ±1 değerleri birlikte değerlendirildiğinde ÇBMTK dışındaki modellerde telafi edici veri için 0,9'un biraz üstünde, telafi edici olmayan veri için ise 0,9'un bir miktar altında değerler elde edilmiştir. ÇBMTK da ise bu değerler 0,85 civarlarında gerçekleşmiştir. ÇBMTK'nin diğer modellerle arasındaki farkı V±1 değerleri ile kapatabildiği görülmektedir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.30. Çoğunlukla Basit Yapı Koşulu pG-DINA Verisi DSO'ları**

Maddelerin çoğunlukla basit yapıda olduğu pG-DINA verilerinin analizi sonucu elde edilen DSO'lar Şekil 4.30'da görülmektedir. Değerler yapısal olarak eşit oran koşulundaki sonuçlar ile karşılaştırıldığında telafi edici olmayan veri için tam DSO'ların arttığı görülmektedir. ÇBMTK'de ise telafi edici veride de artış meydana gelmiştir. Bu artışlarla ÇBMTK'de telafi edici olmayan veri aleyhine olan fark yön değiştirerek lehine dönüşmüş, diğer modellerde ise telafi edici veri lehine olan fark azalmıştır. Ancak görülen farklar çok daha küçük değerlerde bulunmaktadır. Meydana gelen artışların özdeş linki telafi edici ve ÇBMTK telafi edici olmayan veri analizleri için belirgin diğerlerinde ise zayıf olduğu görülmüştür.

V±1 değerlerine bakıldığında oransal olarak eşit yapıdaki sonuçlara göre yönleri farklı, küçük değişimler görülmektedir. Tam ve ±1 değerleri birlikte değerlendirildiğinde tüm modellerin 0,9 seviyesine ulaştığı görülmektedir. Bu açıdan bakıldığında modeller arası ve veri türleri farklar kaybolmuştur.

**Tablo 4.15: Madde Yapısı Oranına Göre fA-M Verileri DSO'ları**

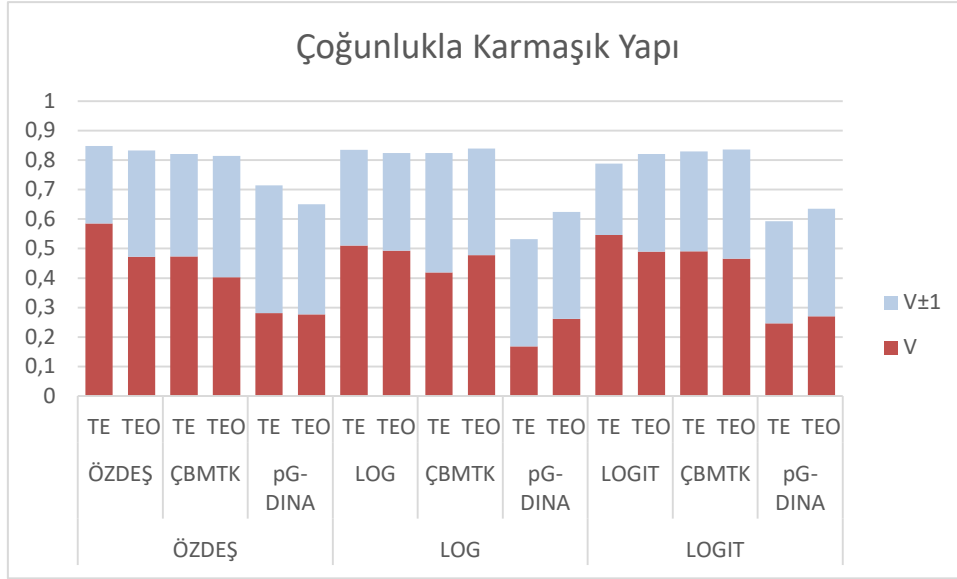
		Vektör						Nitelik						
		Karmaşık		Eşit		Basit		Karmaşık		Eşit		Basit		
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	
KİMLİK	ÖZDEŞ	TE	0,59	0,26	0,63	0,28	0,63	0,29	0,72	0,25	0,77	0,22	0,79	0,21
	ÖZDEŞ	TEO	0,47	0,36	0,56	0,33	0,61	0,31	0,67	0,30	0,73	0,25	0,77	0,22
	ÇBMTK	TE	0,47	0,35	0,57	0,33	0,61	0,32	0,65	0,33	0,74	0,25	0,77	0,22
	ÇBMTK	TEO	0,40	0,41	0,51	0,37	0,59	0,33	0,62	0,35	0,71	0,28	0,76	0,23
	pG-DINA	TE	0,28	0,43	0,29	0,44	0,31	0,43	0,53	0,37	0,55	0,38	0,56	0,38
	pG-DINA	TEO	0,28	0,37	0,30	0,40	0,32	0,41	0,52	0,35	0,55	0,36	0,56	0,36
	LOG	TE	0,51	0,32	0,60	0,31	0,63	0,30	0,70	0,26	0,76	0,22	0,79	0,21
	LOG	TEO	0,49	0,33	0,57	0,32	0,61	0,31	0,68	0,28	0,74	0,24	0,77	0,22
	ÇBMTK	TE	0,42	0,40	0,54	0,36	0,60	0,32	0,64	0,34	0,73	0,26	0,77	0,22
	ÇBMTK	TEO	0,48	0,36	0,56	0,33	0,61	0,32	0,67	0,30	0,74	0,25	0,77	0,22
LOG	pG-DINA	TE	0,17	0,36	0,28	0,41	0,30	0,42	0,46	0,36	0,54	0,37	0,55	0,37
	pG-DINA	TEO	0,26	0,36	0,31	0,40	0,32	0,41	0,52	0,35	0,55	0,36	0,56	0,36
	LOGIT	TE	0,55	0,24	0,62	0,27	0,64	0,29	0,67	0,30	0,76	0,23	0,79	0,20
	LOGIT	TEO	0,49	0,33	0,57	0,32	0,61	0,31	0,67	0,29	0,74	0,24	0,77	0,22
	ÇBMTK	TE	0,49	0,34	0,59	0,31	0,63	0,30	0,66	0,32	0,75	0,24	0,78	0,21
	ÇBMTK	TEO	0,47	0,37	0,56	0,34	0,60	0,32	0,66	0,31	0,73	0,25	0,77	0,22
LOGIT	pG-DINA	TE	0,25	0,35	0,27	0,39	0,30	0,42	0,51	0,34	0,53	0,37	0,55	0,38
	pG-DINA	TEO	0,27	0,36	0,30	0,40	0,32	0,41	0,52	0,35	0,55	0,36	0,56	0,36

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.15'de yapılarına göre maddelerin farklı oranlarda bulunduğu özdeş, log ve logit linkindeki fA-M verilerinin pG-DINA, ÇBMTK ve verinin üretildiği fA-M'lara uygulanması sonucu vektör ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiğinde verinin ait olduğu fA-M kestirimine yakın bir düzeyde ÇBMTK'nin başarılı kestirimi dikkat çekmektedir. pG-DINA ise diğer modellerden daha geride olduğu görülmektedir. Testteki basit yapıdaki madde oranı arttıkça DSO'lar tam sonuçlarda artmaktadır. Genel olarak telafi edici ve telafi edici olmayanlara verilerin analizinden elde edilen sonuçların birbirine yakın olduğu görülmüştür.

İlk alt problemde aynı senaryolar için doğru model olan ÇBMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiğinde, fA-M verilerinin daha başarılı bir şekilde kestirildiği görülmektedir. Ancak ikinci alt problemdeki pG-DINA verisi kestirimi ile karşılaştırıldığında ise daha düşük doğrulukta kestirim yapıldığı görülmektedir. FA-M'lara ait tam vektör DSO'larına bakıldığında yapılarına göre maddelerin eşit oranda olduğu durumda, ilk alt problemdeki doğru model olan

ÇBMTK sonuçlarından telafi edici veri için 0,08-0,11, telafi edici olmayan veri içinse 0,07-0,08'lik oranda daha iyi olduğu görülmektedir. Yine aynı koşul için ikinci alt problemdeki pG-DINA kestiriminin fA-M'lardan telafi edici veri için 0,09-0,12, telafi edici olmayan veri içinse 0,08-0,09'luk oranda daha iyi olduğu görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA ve fA-M verilerinde de en düşük performansı göstermiştir.

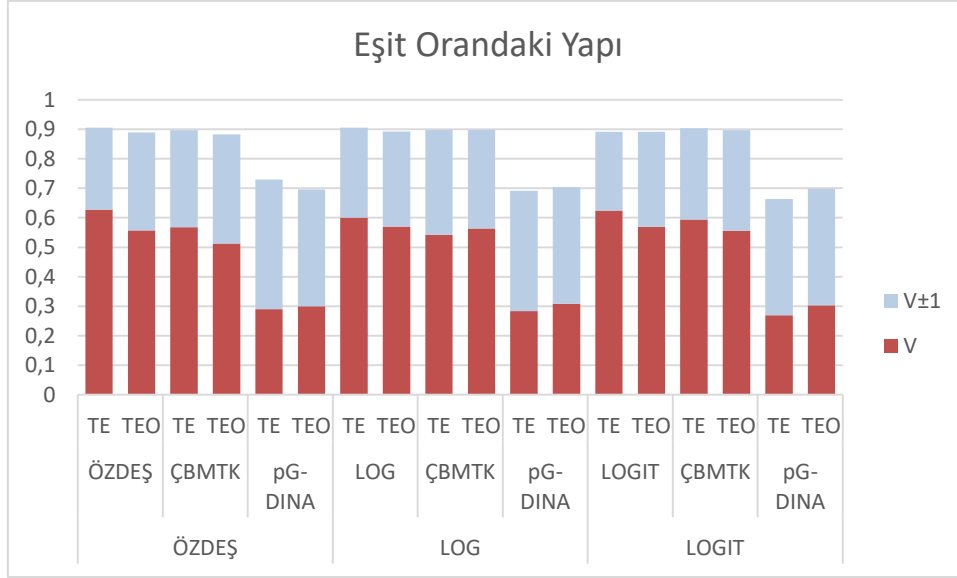


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.31. Çoğunlukla Karmaşık Yapı Koşulu fA-M Verisi DSO'ları**

Testteki maddelerin çoğunun (%80) karmaşık yapıda olduğu fA-M verilerinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.31'de görülmektedir. FA-M özdeş ve logit verisinin doğru model ve ÇBMTK analizinde telafi edici veri lehine fark bulunmaktadır. Log linki verisinin ÇBMTK ve pG-DINA analizi sonucunda ise telafi edici olmayan veri için daha oranlar elde edilmektedir. Aynı modelin verisi için uyarlanan modeller farklı veri türünde daha iyi sonuçlar verebilmektedir. En yüksek farkın özdeş linki verisinde özdeş analizinde telafi edici lehine olduğu görülmektedir. ÇBMTK'nin verinin ürettiği fA-M model sonuçlarına yakın sonuçlar ürettiği görülmektedir. FA-M modelleri ile ÇBMTK arasındaki fark özellikle telafi edici olmayan veride küçüktür. FA-M modelleri kendi aralarında değerlendirildiğinde sonuçlar herhangi bir linki çok fazla ön plana çıkarmamaktadır. pG-DINA ise her üç modele ait veride de en kötü oranlara sahiptir. Diğer modellerde telafi edici veri için gerçekleşen 0,42-0,59 aralığında olan tam doğrulama oranı pG-

DINA'da 0,17-0,28 aralığında gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise pG-DINA 0,25-0,28'e sahipken iken diğer modeller 0,40-0,49'luk doğru sınıflamaya sahiptir. Bu sonuçlar  $V \pm 1$  hata payı eklendiğinde ÇBMTK'de ve fA-M'da 0,80 civarlarındayken pG-DINA'da ise 0,60 dolaylarındadır.



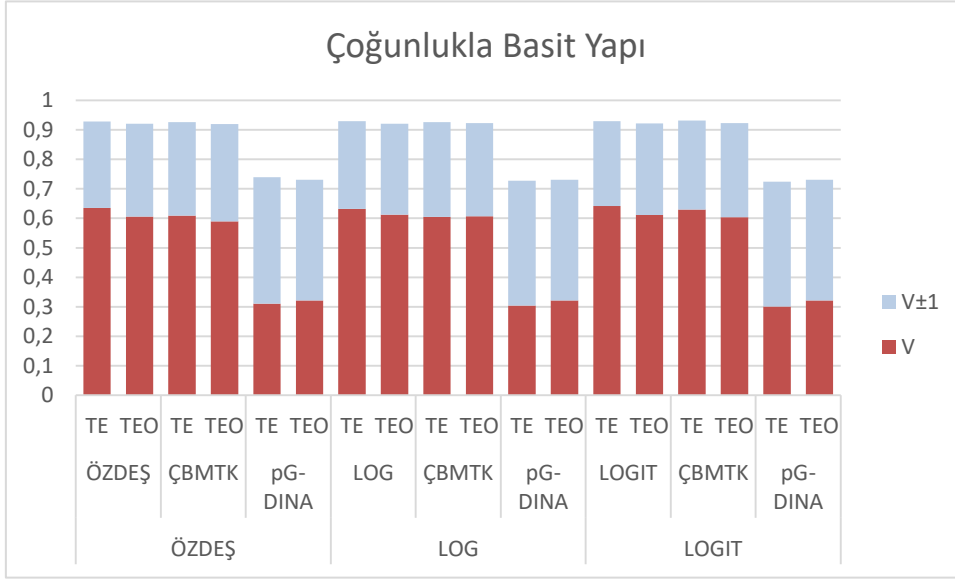
Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.32. Eşit Oran Yapı Koşulu fA-M Verisi DSO'ları**

Şekil 4.32'de görülen grafikteki değerler testteki maddelerin yapılarına göre eşit oranda olduğu durumlardaki fA-M verilerinin analizi sonucu elde edilen DSO'lara aittir. Değerler testin çoğunlukla karmaşık yapıda olduğu durumda elde edilen değerlerle karşılaştırıldığında, telafi edici ve telafi edici olmayan veri analiz sonuçları arasındaki farkların azaldığı görülmektedir. FA-M'larda telafi edici, pG-DINA'da telafi edici olmayan verilerin biraz daha iyi kestirilmiştir. ÇBMTK ise özdeş ve logit linkinde telafi edici, log linkinde ise telafi edici olmayan veride kısmen daha iyidir. pG-DINA haricindeki modellerle 0,54-0,62 düzeylerinde DSO elde edilirken, pG-DINA'da ise ancak 0,27-0,31 seviyelerine ulaşılmıştır. ÇBMTK ve fA-M'larda tam vektör tahmininde gözlenen daha yüksek artışlar nedeniyle pG-DINA ile diğer modeller arasındaki farkın arttığı görülmektedir.

$V \pm 1$  değerlerine bakıldığında maddelerin çoğunlukla karmaşık olduğu sonuçlara göre değişimler modellerde farklı oran ve yöndedir. Tam ve  $\pm 1$  değerleri birlikte değerlendirildiğinde hem modellerin hem de veri türlerinin sonuçların birbirine yaklaştığı görülmüştür. Toplamda pG-DINA dışındaki modellerde 0,9'un altında

değerler elde edilmiştir. pG-DINA'da ise bu değerler 0,7 dolaylarında gerçekleşmiştir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.33. Çoğunlukla Basit Yapı Koşulu fA-M Verisi DSO'ları**

Testteki maddelerin çoğunluğunun (%80) basit yapıda olması durumundaki fA-M verilerinin analizi sonucu elde edilen DSO'lar Şekil 4.33'de bulunmaktadır. Değerler yapısal olarak eşit oran koşulundaki sonuçlar ile karşılaştırıldığında yüksek olan tam DSO'ların artmadığı fA-M ve ÇBMTK düşük olan tam DSO'ların yüksek olanların seviyesine çıktığı görülmektedir. Böylece pG-DINA sonuçları kendi içinde, fA-M ve ÇBMTK sonuçları da kendi içinde verinin türünden bağımsız bir şekilde birbirlerine çok yakın DSO'lar elde etmektedirler.

V±1 değerlerinde meydana gelen değişim sonucu eşitlikler benzer durumdadır. Burada pG-DINA'dan elde edilen sonuçların fA-M ve ÇBMTK'den daha yüksek olduğu görülmektedir. pG-DINA hariç diğer modellerde 0,7'lere ulaşan-yaklaşan tam DSO'lar, V±1 ile değerlendirildiğinde 0,9'ları geçmektedir. pG-DINA'da ise yüksek V±1 değerleri sayesinde 0,7 seviyesini biraz aşmaktadır.

#### 4.2.3. Test Uzunluğu Koşullarına Göre Doğru Sınıflama Oranları

İkinci alt problemin üçüncü koşulu olan test uzunluğu için 3 ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. 15, 30 ve 60 maddelik test uzunlukları kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda 15 maddelik test değerleri ele alınmış, 30 maddelik test için elde edilen değerler 15 maddelik test

değerleri ile karşılaştırılarak değerlendirilmiş ve 60 maddelik test için elde edilen değerler ise 30 maddelik değerleri ile karşılaştırılarak değerlendirilmiştir.

**Tablo 4.16: Test Uzunluğuna Göre pG-DINA Verisi DSO'ları**

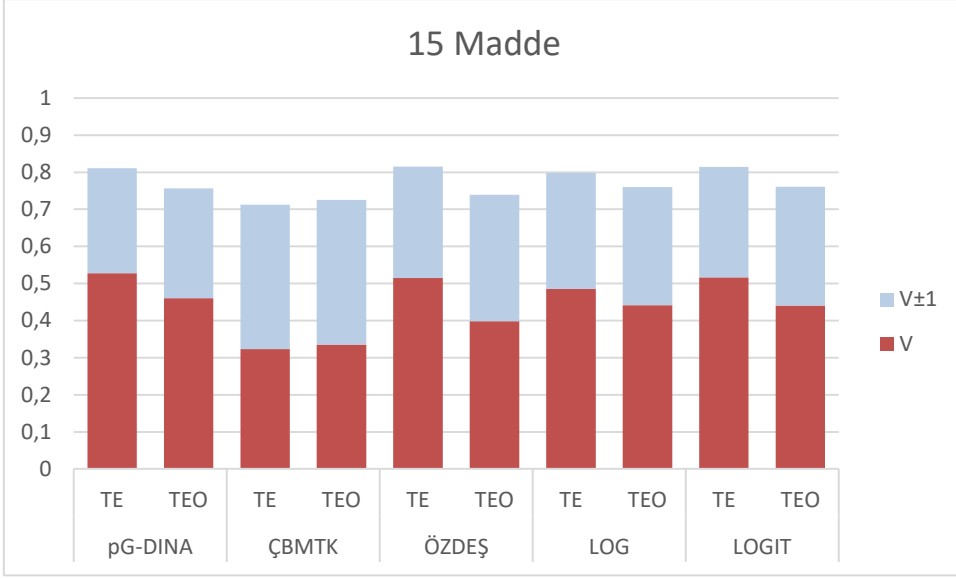
		<i>Vektör</i>						<i>Nitelik</i>					
		<i>15</i>		<i>30</i>		<i>60</i>		<i>15</i>		<i>30</i>		<i>60</i>	
		<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>
<b>pG-DINA</b>	<b>TE</b>	0,53	0,28	0,73	0,19	0,90	0,08	0,72	0,20	0,85	0,12	0,95	0,05
<b>pG-DINA</b>	<b>TEO</b>	0,46	0,30	0,64	0,24	0,83	0,13	0,67	0,23	0,79	0,16	0,90	0,08
<b>ÇBMTK</b>	<b>TE</b>	0,32	0,39	0,45	0,35	0,56	0,29	0,54	0,38	0,64	0,32	0,71	0,26
<b>ÇBMTK</b>	<b>TEO</b>	0,34	0,39	0,48	0,35	0,61	0,29	0,56	0,36	0,68	0,28	0,77	0,22
<b>ÖZDEŞ</b>	<b>TE</b>	0,51	0,30	0,72	0,20	0,89	0,09	0,71	0,22	0,84	0,13	0,94	0,05
<b>ÖZDEŞ</b>	<b>TEO</b>	0,40	0,34	0,60	0,27	0,79	0,16	0,63	0,28	0,77	0,18	0,88	0,09
<b>LOG</b>	<b>TE</b>	0,49	0,31	0,70	0,22	0,88	0,10	0,69	0,23	0,83	0,14	0,94	0,05
<b>LOG</b>	<b>TEO</b>	0,44	0,32	0,63	0,25	0,82	0,14	0,65	0,25	0,79	0,17	0,90	0,08
<b>LOGIT</b>	<b>TE</b>	0,52	0,30	0,73	0,20	0,90	0,08	0,71	0,22	0,84	0,13	0,95	0,05
<b>LOGIT</b>	<b>TEO</b>	0,44	0,32	0,63	0,25	0,82	0,14	0,65	0,25	0,78	0,17	0,90	0,08

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.16'da farklı test uzunluğuna sahip pG-DINA verisinin pG-DINA, ÇBMTK ve fA-M'lara uygulanması sonucu vektör ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiğinde verinin ait olduğu pG-DINA'ya çok yakın bir düzeyde fA-M'ların başarılı kestirimleri dikkat çekmektedir. ÇBMTK'nin diğer modellerden daha başarısız olduğu görülmektedir. Test uzunluğu arttıkça DSO'lar tam ve tama ilave edilen  $\pm 1$  hatalı sonuçlarda artmaktadır. ÇBMTK haricindeki modellerin telafi edici verilerin analizinden elde edilen sonuçlar daha iyidir.

İlk alt problemde aynı senaryolar için doğru model olan ÇBMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiğinde, pG-DINA verisinin çok daha başarılı bir şekilde kestirildiği görülmektedir. Doğru modeller olan ÇBMTK ve pG-DINA'ya ait tam vektör DSO'larına bakıldığında test uzunluğunun 30 (orta) madde olduğu telafi edici veri için 0,23, telafi edici olmayan veri içinse 0,25'lük fark olduğu görülmektedir. Her iki alt problemde de uyarlanan modeller olan fA-M'larda da benzer farklar görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA ve burada en düşük performansı gösteren ÇBMTK'nin arasında da benzer fark bulunmaktadır.

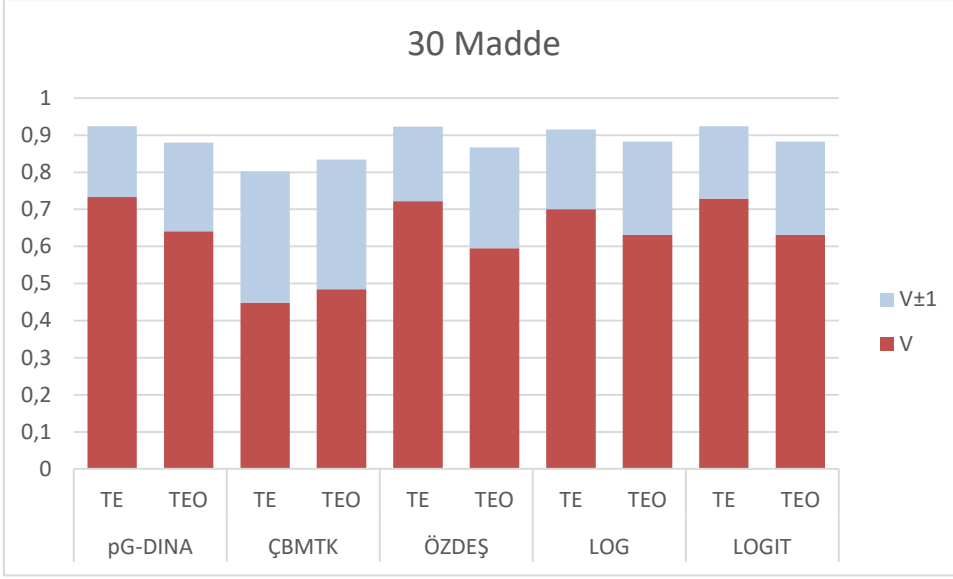




Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.34. 15 Maddelik Test Uzunluğu Koşulu pG-DINA Verisi DSO'ları**

Test uzunluğunun 15 madde olduğu pG-DINA verisinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.34'de bulunan grafikte görülmektedir. ÇBMTK hariç diğer modellerin telafi edici veri türünde daha iyi sonuçlar verdiği görülmektedir. ÇBMTK ayrıca elde edilen sonuçlara bakıldığında daha düşük olan DSO'lar nedeniyle de diğer modellerden ayrılmaktadır. Diğer modellerde telafi edici veri için gerçekleşen 0,50'lik tam doğrulama oranı ÇBMTK'de 0,32 olarak gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise ÇBMTK 0,34 iken diğer modeller 0,40-0,46 arasında doğru sınıflamaya sahiptir. Bu sonuçlar  $V \pm 1$  hata payı eklendiğinde ÇBMTK'de 0,72 civarlarındayken diğer modellerde telafi edici veri için 0,80, telafi edici olmayan veri için ise 0,75 dolaylarındadır. ÇBMTK'nin bu değerle diğer modellere yaklaştığı görülmektedir.

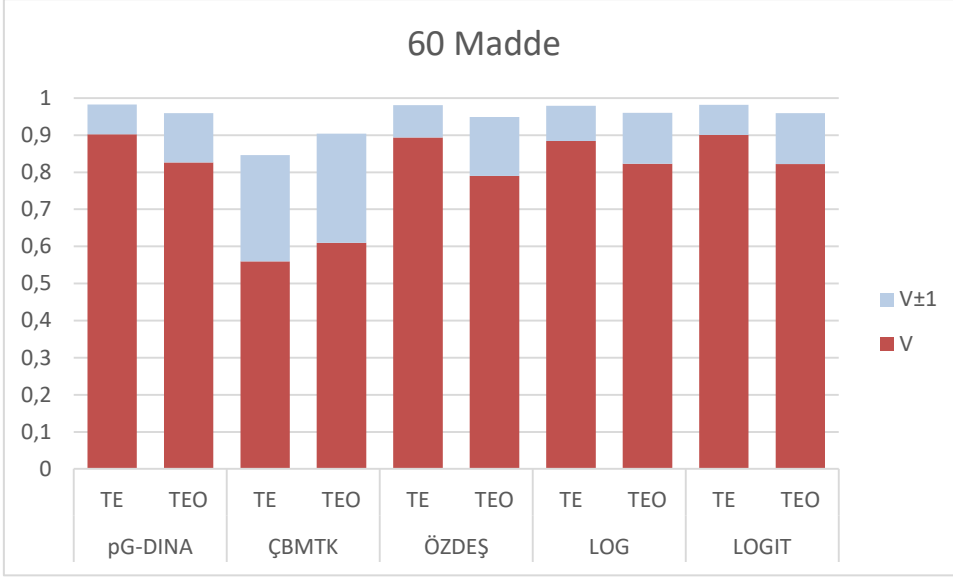


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.35. 30 Maddelik Test Uzunluğu Koşulu pG-DINA Verisi DSO'ları**

Şekil 4.35'deki değerler testin 30 maddeden oluştuğu durumlardaki pG-DINA verisinin analizi sonucu elde edilen DSO'lara aittir. Değerler 15 maddelik test uzunluğu koşulunda elde edilen değerlerle karşılaştırıldığında, telafi edici verinin analiz sonucu lehine olan durum farkların bir miktar artarak devam ettiği görülmektedir. ÇBMTK buna istisna olmaya devam etmektedir ve telafi edici olmayan veri lehine olan fark biraz daha belirgin hale gelmiştir. ÇBMTK haricindeki modellerle telafi edici verilerde 0,7 düzeylerinde DSO elde edilirken, telafi edici olmayan verilerde 0,6 düzeyi biraz aşılmaktadır. ÇBMTK'de ise 0,45-0,48 seviyelerine ulaşılmıştır. Tam vektör tahmininde 0,13-0,21 arasında artışlar gözlenmektedir. ÇBMTK'deki artışların biraz daha az olmakla beraber tüm modellerde benzer bir oranda gerçekleşmesi dikkat çekmektedir.

V±1 değerlerine bakıldığında 15 madde test uzunluğuna kıyasla 0,04-0,1 düzeyinde düşüş görülmektedir. Madde sayısının 15'den 30'a çıkmasıyla V±1 değerlerinin bir kısmı tam doğru tahmine kaymakta ve V±1 kestirimi ise kaybettiği rakamdan bir miktar daha az yeni kestirim elde etmektedir. Tam ve ±1 değerleri birlikte değerlendirildiğinde telafi edici veri için ÇBMTK'de 0,8, diğer modellerde 0,9'un biraz üstünde değerleri elde edilmiştir. Telafi edici olmayan veri için ise tüm modellerde 0,85 dolaylarında değerler elde edilmiştir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.36. 60 Maddelik Test Uzunluğu Koşulu pG-DINA Verisi DSO'ları**

60 maddelik test uzunluğuna sahip pG-DINA verilerinin analizi sonucu elde edilen DSO'lar Şekil 4.36'da görülen grafikte bulunmaktadır. Değerler 30 maddelik test koşulundaki sonuçlar ile karşılaştırıldığında tam ve toplam DSO'ların arttığı görülmektedir. Artışın telafi veri türüne göre birbirine yakın olmakla birlikte telafi edici olmayan veri için bir miktar daha fazla olduğu görülmektedir. Böylece ÇBMTK'de telafi edici olmayan veri lehine olan fark kısmen artarken, diğer modellerde telafi edici veri lehine olan fark azalmıştır. ÇBMTK'de görülen artış telafi edici veri için 0,09 iken telafi edici olmayan veride 0,13 düzeyindedir. Tam vektör DSO'larında meydana gelen artış telafi edici veri türünde 0,17 düzeylerinde iken telafi edici olmayan veri türünde 0,19 düzeylerindedir.

V±1 değerlerinde meydana gelen değişimde ise ÇBMTK 0,06'lık, diğer modellerde 0,11-0,12'lik düşüş görülmüştür. ÇBMTK hariç diğer modellerde tam DSO'lar telafi edici veride 0,9'a, telafi edici olmayan veride 0,79-0,83 aralığındadır. ÇBMTK 'da ise telafi edici veri için 0,56, telafi edici olmayan için 0,61 değerindedir. ±1 eklendiğinde bu değerler ÇBMTK'de 0,86, diğer modellerde ise 0,95 civarlarındadır. Bu oranlar araştırmada elde edilen en iyi oranları göstermektedir. Bu koşulun en iyi doğruluğa sahip koşul olduğu görülmektedir.

**Tablo 4.17: Test Uzunluđuna Gre fA-M Verileri DSO'ları**

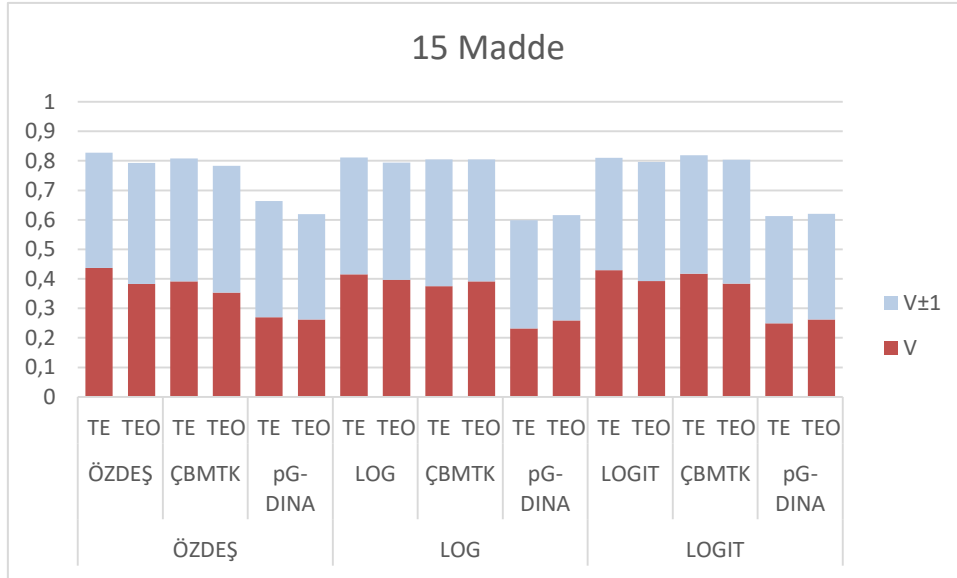
		Vektr						Nitelik					
		15		30		60		15		30		60	
		Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1	Tam	±1
KİMLİK	ZDEŐ TE	0,44	0,39	0,61	0,29	0,80	0,15	0,65	0,32	0,76	0,23	0,87	0,12
	ZDEŐ TEO	0,38	0,41	0,54	0,36	0,71	0,24	0,61	0,34	0,72	0,26	0,83	0,16
	BMTK TE	0,39	0,42	0,55	0,35	0,71	0,23	0,61	0,36	0,72	0,27	0,83	0,17
	BMTK TEO	0,35	0,43	0,50	0,39	0,65	0,30	0,59	0,37	0,70	0,29	0,80	0,20
	pG-DINA TE	0,27	0,39	0,29	0,44	0,32	0,47	0,52	0,36	0,55	0,38	0,57	0,39
	pG-DINA TEO	0,26	0,36	0,30	0,40	0,34	0,42	0,51	0,34	0,54	0,36	0,58	0,36
	LOG TE	0,42	0,40	0,58	0,33	0,75	0,20	0,64	0,31	0,75	0,23	0,85	0,14
	LOG TEO	0,40	0,40	0,55	0,34	0,72	0,23	0,62	0,33	0,73	0,25	0,84	0,16
	BMTK TE	0,38	0,43	0,52	0,38	0,67	0,28	0,61	0,36	0,71	0,28	0,81	0,18
	BMTK TEO	0,39	0,41	0,55	0,35	0,71	0,25	0,62	0,35	0,73	0,26	0,83	0,16
LOG	pG-DINA TE	0,23	0,37	0,25	0,40	0,27	0,43	0,50	0,35	0,52	0,37	0,54	0,38
	pG-DINA TEO	0,26	0,36	0,30	0,39	0,34	0,42	0,51	0,34	0,54	0,36	0,57	0,37
	LOGIT TE	0,43	0,38	0,60	0,28	0,78	0,14	0,63	0,33	0,74	0,25	0,85	0,14
	LOGIT TEO	0,39	0,40	0,55	0,34	0,73	0,22	0,62	0,34	0,73	0,26	0,84	0,16
	BMTK TE	0,42	0,40	0,57	0,33	0,73	0,22	0,63	0,35	0,74	0,26	0,84	0,16
	BMTK TEO	0,38	0,42	0,54	0,36	0,70	0,25	0,61	0,36	0,73	0,26	0,83	0,17
LOGIT	pG-DINA TE	0,25	0,36	0,27	0,39	0,30	0,41	0,51	0,35	0,53	0,37	0,55	0,37
	pG-DINA TEO	0,26	0,36	0,30	0,39	0,34	0,42	0,51	0,34	0,54	0,36	0,58	0,37

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.17'de farklı test uzunluklarındaki zdeŐ, log ve logit linkindeki fA-M verilerinin pG-DINA, BMTK ve verinin retildiđi fA-M'lara uygulanması sonucu vektr ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiđinde verinin ait olduđu fA-M kestirimine yakın bir dzeyde BMTK'nin başarılı kestirimi dikkat ekmektedir. PG-DINA ise diđer modellerden daha geride olduđu grlmektedir. Madde sayısı arttıka DSO'lar tam ve tama ilave edilen ±1 hatalı sonuçlarda artmaktadır. Genel olarak bakıldıđında telafi edici verilerin analizinden elde edilen sonuçların telafi edici olmayanlara gre bir miktar daha iyi olduđu grlmŐtr.

İlk alt problemde aynı senaryolar iin dođru model olan BMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiđinde, fA-M verilerinin daha başarılı bir Őekilde kestirildiđi grlmektedir. Ancak ikinci alt problemdeki pG-DINA verisi kestirimi ile karŐılaŐtırıldıđında ise daha dŐk dođrulukta kestirim yapıldıđı grlmektedir. FA-M'lara ait tam vektr DSO'larına bakıldıđında 30 maddelik test uzunluđunda, ilk alt problemdeki dođru model olan BMTK sonuçlarından telafi

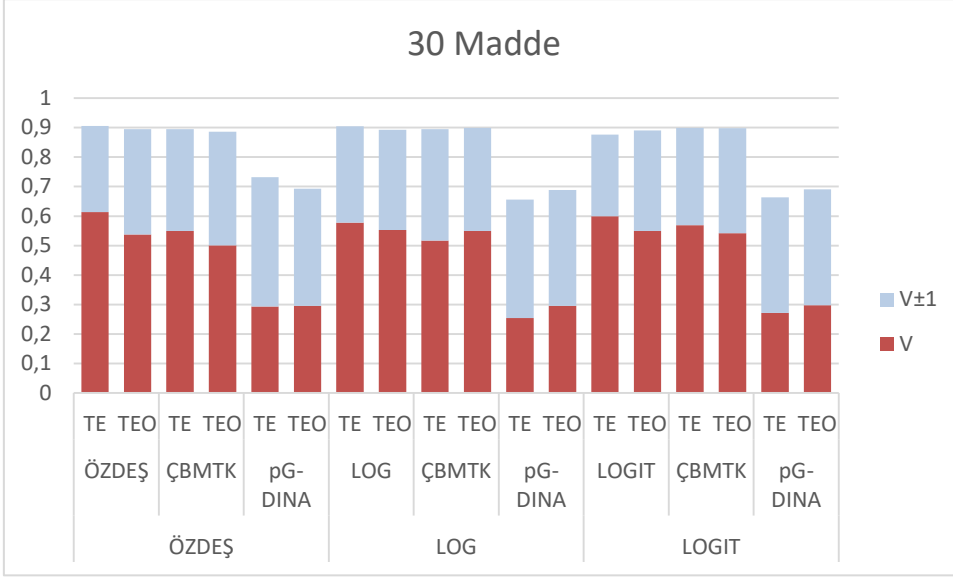
edici veri için 0,08-0,11'lik, telafi edici olmayan veri içinse 0,05-0,06'luk oranda daha iyi olduğu görülmektedir. Yine aynı koşul için ikinci alt problemdeki pG-DINA kestirimi fA-M'lardan telafi edici veri için 0,12-0,18, telafi edici olmayan veri içinse 0,09-0,1'lik oranda daha iyi olduğu görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA fA-M verilerinde de en düşük performansı göstermiştir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.37. 15 Maddelik Test Uzunluğu Koşulu fA-M Verisi DSO'ları**

Testteki madde sayısının 15 olduğu fA-M verilerinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.37'deki grafikte görülmektedir. FA-M analizlerinde telafi edici veri lehine olan fark, ÇBMTK analizlerinde de log linki haricinde aynı yöndedir. pG-DINA analizlerinde daha küçük görülen fark, log ve logit linki analizinde telafi edici olmayan veri lehinedir. ÇBMTK'nin verinin üretildiği fA-M model sonuçlarına son derece yakın sonuçlar ürettiği görülmektedir. FA-M modelleri kendi aralarında değerlendirildiğinde sonuçlar herhangi bir linki ön plana çıkarmamaktadır. pG-DINA ise her üç modele ait veride de en kötü oranlara sahiptir. Diğer modellerde telafi edici veri için gerçekleşen 0,38-0,44 arasındaki tam doğrulama oranı pG-DINA'da 0,23-0,27 arası gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise pG-DINA 0,26 iken diğer modeller 0,35-0,4'lük doğru sınıflamaya sahiptir. Bu sonuçlar  $V \pm 1$  hata payı eklendiğinde ÇBMTK'de ve fA-M'da 0,80 civarlarındayken pG-DINA'da ise 0,60 dolaylarındadır.

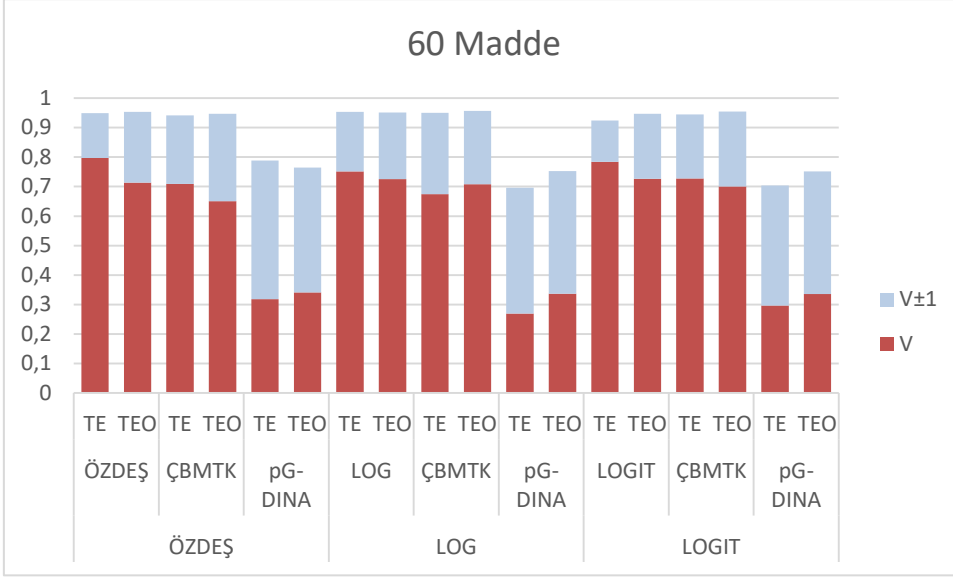


Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.38. 30 Maddelik Test Uzunluğu Koşulu fA-M Verisi DSO'ları**

Şekil 4.38'deki değerler test uzunluğunun 30 madde olduğu durumlardaki fA-M verilerinin analizi sonucu elde edilen DSO'lara aittir. Değerler 15 maddelik test uzunluğu koşulunda elde edilen değerlerle karşılaştırıldığında, küçük de olsa farkların devam ettiği görülmektedir. pG-DINA'da telafi edici olmayan veri lehine olan fark biraz daha belirgin hale gelmiştir. pG-DINA haricindeki modellerle telafi edici verilerde 0,52-0,61 düzeylerinde DSO elde edilirken, bu rakam telafi edici olmayan verilerde 0,5-0,55 düzeyindedir. pG-DINA'da ise ancak 0,25-0,3 seviyelerine ulaşılmıştır. ÇBMTK ve fA-M'larda tam vektör tahmininde 0,14-0,17 arasında artışlar gözlenmekteyken bu rakam pG-DINA sadece 0,02-0,04 aralığındadır. pG-DINA ile diğer modeller arasındaki farkın arttığı görülmektedir.

V±1 değerlerine bakıldığında düşük madde ayırt edicilik sonuçlarına göre ÇBMTK ve fA-M'larda 0,04-0,1 düzeyinde düşüş görülmektedir. pG-DINA'da 15 maddeye nazaran 0,03-0,05'lik artış görülmektedir. Tam ve ±1 değerleri birlikte değerlendirildiğinde telafi edici ve telafi edici olmayan sonuçların birbirine yaklaştığı görülmüştür. Toplamda pG-DINA dışındaki modellerde 0,9'un hemen altında değerler elde edilmiştir. pG-DINA'da ise bu değerler 0,65-0,73 aralığında gerçekleşmiştir.



**Şekil 4.39. 60 Maddelik Test Uzunluğu Koşulu fA-M Verisi DSO'ları**

60 maddelik test uzunluğuna sahip fA-M verilerinin analizi sonucu elde edilen DSO'lar Şekil 4.39'daki grafikte bulunmaktadır. Değerler test uzunluğunun 30 madde olduğu sonuçlar ile karşılaştırıldığında tam ve toplam DSO'ların arttığı görülmektedir. Bu artışın pG-DINA analizlerinde telafi edici olmayan veri türü için biraz daha yüksek olduğu görülmektedir. Bu sayede pG-DINA'da telafi edici olmayan veri lehine olan fark biraz daha belirgin hale gelmiştir. FA-M ve ÇBMTK'de tam vektör DSO'larında meydana gelen artış 0,15-0,19 düzeylerinde iken pG-DINA'da bu artışın kısıtlı olduğu ve 0,02-0,04 düzeylerinde olduğu görülmüştür.

V±1 değerlerinde meydana gelen değişim ise fA-M ve ÇBMTK'de 0,1-0,19 düzeyinde düşüş, pG-DINA'da ise 0,02-0,03'lük artış yönündedir. pG-DINA hariç diğer modellerde 0,7'lere ulaşan-yaklaşan tam DSO'lar, V±1 ile değerlendirildiğinde 0,9'ları geçmektedir. Bu değerlere bakıldığında pG-DINA harici modeller arasında ve veri türleri arasında farklar kaybolmaktadır.

#### **4.2.4. Yetenekler Arası Korelasyon Koşullarına Göre Doğru Sınıflama Oranları**

İkinci alt problemin dördüncü ve son koşulu olan yetenekler arası korelasyon için iki ayrı değerde veriler oluşturulup analizler gerçekleştirilmiştir. Yetenekler arası korelasyonun olmadığı durum için 0, yeteneklerin ilişkili olduğu durum için ise 0,6 korelasyon katsayısı kullanılmıştır. Grafik eşliğinde vektörel sonuçlar için yapılan yorumlarda korelasyonun olmadığı durum için değerler değerlendirilmiş ve

korelasyonun 0,6 olduğu durum için elde edilen değerler ise 0 korelasyon değerleri ile karşılaştırılarak değerlendirilmiştir.

**Tablo 4.18: Yetenekler Arası Korelasyona Göre pG-DINA Verisi DSO'lar**

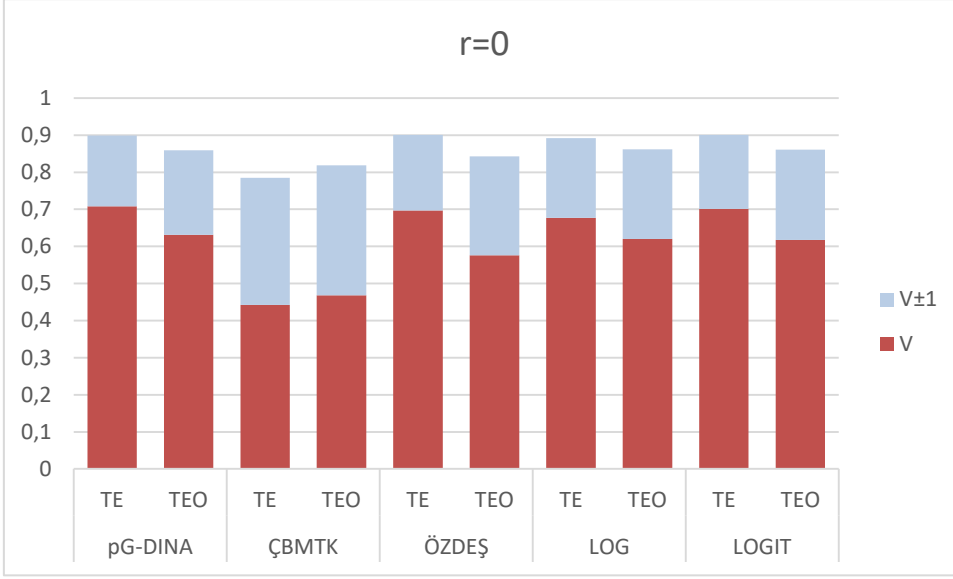
		<i>Vektör</i>				<i>Nitelik</i>			
		<i>0</i>		<i>0,6</i>		<i>0</i>		<i>0,6</i>	
		<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>	<i>Tam</i>	<i>±1</i>
<b>pG-DINA</b>	<b>TE</b>	0,71	0,19	0,73	0,18	0,83	0,13	0,84	0,12
<b>pG-DINA</b>	<b>TEO</b>	0,63	0,23	0,65	0,22	0,78	0,16	0,80	0,15
<b>ÇBMTK</b>	<b>TE</b>	0,44	0,34	0,45	0,34	0,63	0,32	0,63	0,32
<b>ÇBMTK</b>	<b>TEO</b>	0,47	0,35	0,49	0,34	0,67	0,29	0,68	0,28
<b>ÖZDEŞ</b>	<b>TE</b>	0,70	0,20	0,72	0,19	0,82	0,14	0,83	0,13
<b>ÖZDEŞ</b>	<b>TEO</b>	0,58	0,27	0,61	0,25	0,75	0,19	0,77	0,18
<b>LOG</b>	<b>TE</b>	0,68	0,21	0,70	0,20	0,81	0,15	0,83	0,13
<b>LOG</b>	<b>TEO</b>	0,62	0,24	0,64	0,23	0,77	0,17	0,79	0,16
<b>LOGIT</b>	<b>TE</b>	0,70	0,20	0,73	0,18	0,82	0,14	0,84	0,12
<b>LOGIT</b>	<b>TEO</b>	0,62	0,24	0,64	0,23	0,77	0,18	0,79	0,16

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

Tablo 4.18'de farklı yetenekler arası korelasyona sahip pG-DINA verisinin pG-DINA, ÇBMTK ve fA-M'lara uygulanması sonucu vektör ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiğinde verinin ait olduğu pG-DINA'ya çok yakın bir düzeyde fA-M'ların başarılı kestirimleri dikkat çekmektedir. ÇBMTK ise diğer modellerden daha başarısız olduğu görülmektedir. Yetenekler arası korelasyonun varlığı tam DSO'ları artırmaktadır. ÇBMTK haricindeki modellerin telafi edici verilerin analizinden elde edilen sonuçlar daha iyidir.

İlk alt problemde aynı senaryolar için doğru model olan ÇBMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiğinde, pG-DINA verisinin çok daha başarılı bir şekilde kestirildiği görülmektedir. Doğru modeller olan ÇBMTK ve pG-DINA'ya ait tam vektör DSO'larına bakıldığında korelasyonun 0 olduğu telafi edici veri için 0,23, telafi edici olmayan veri içinse 0,16'lık fark olduğu görülmektedir. Her iki alt problemde de uyarlanan modeller olan fA-M'larda da benzer farklar görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA ve burada en düşük performansı gösteren ÇBMTK'nin arasında da benzer fark bulunmaktadır. Ancak korelasyon artışının ÇBMTK verisinin doğru kestiriminde daha etkili olduğu görülmektedir.

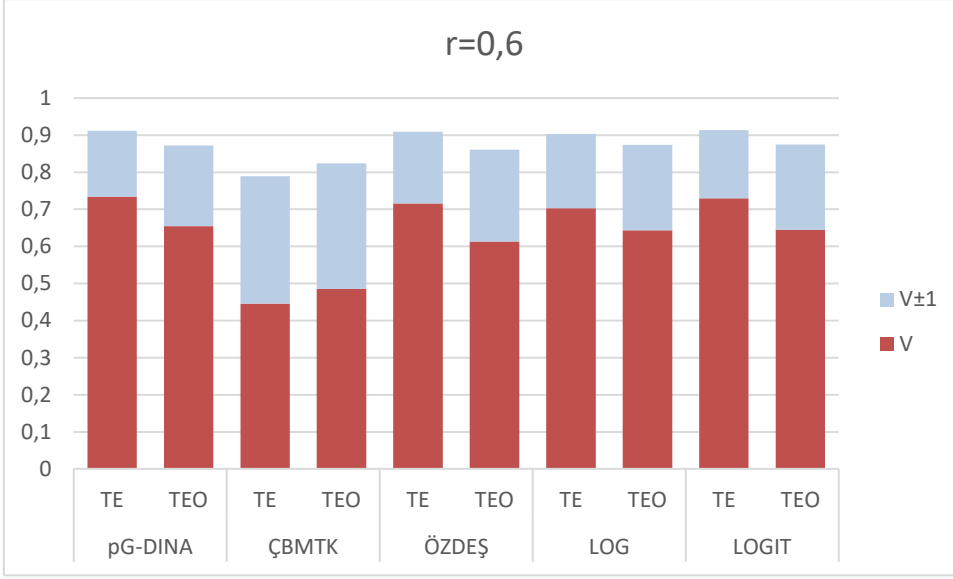




Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.40. Yetenekler Arası 0 Korelasyon Koşulu pG-DINA Verisi DSO'ları**

Yetenekler arası korelasyonun 0 olduğu pG-DINA verisinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.40'daki grafikte görülmektedir. Önceki analizlere benzer ÇBMTK hariç diğer modellerin telafi edici veri türünde daha iyi sonuçlar verdiği görülmektedir. ÇBMTK ayrıca elde edilen sonuçlara bakıldığında daha düşük olan DSO'lar nedeniyle de diğer modellerden ayrılmaktadır. Diğer modellerde telafi edici veri için gerçekleşen 0,7'lik tam doğrulama oranı ÇBMTK'de 0,44 olarak gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise ÇBMTK 0,47 iken diğer modeller 0,58-0,63 arasında doğru sınıflamaya sahiptir. Bu sonuçlar  $V \pm 1$  hata payı eklendiğinde ÇBMTK'de 0,8 civarlarındayken diğer modellerde telafi edici veri için 0,9, telafi edici olmayan veri için ise 0,85 dolaylarındadır. ÇBMTK'nin bu değerle diğer modellere yaklaştığı görülmektedir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.41. Yetenekler Arası 0,6 Korelasyon Koşulu pG-DINA Verisi DSO'ları**

Şekil 4.41'de görülen grafikteki değerler yetenekler arası korelasyonun 0,6 olduğu durumlarda pG-DINA verisinin analizi sonucu elde edilen DSO'lara aittir. Değerler 0 korelasyon koşulunda elde edilen değerlerle karşılaştırıldığında tüm modellerdeki artışın olduğu, ancak bu artışın 0,01-0,03 düzeylerinde son derece küçük olduğu görülmektedir.  $V_{\pm 1}$  değerlerine bakıldığında benzer şekilde çok küçük değişimler görülmektedir. Bu değişimler 0-0,02 birim negatif yöndedir. Tam ve  $\pm 1$  değerleri birlikte değerlendirildiğinde 0 korelasyon şartına göre dikkate değer değişim görülmemektedir.

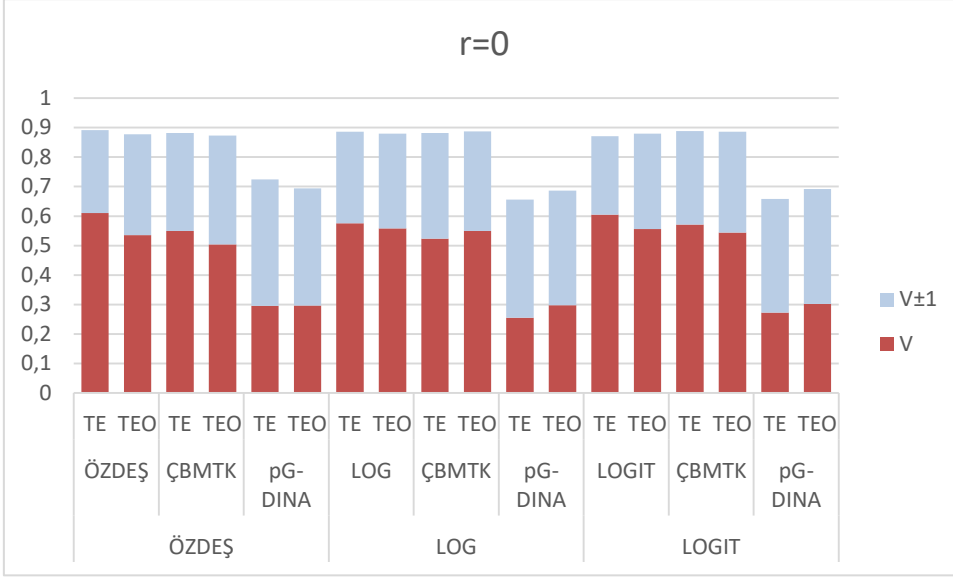
Tablo 4.19'da yetenekler arası korelasyonun olduğu ve olmadığı durumlarda özdeş, log ve logit linkindeki fA-M verilerinin pG-DINA, ÇBMTK ve verinin üretildiği fA-M'lara uygulanması sonucu vektör ve nitelik temelinde elde edilen DSO'lar bulunmaktadır. Niteliklere ait veriler incelendiğinde önceki analizlere benzer şekilde verinin ait olduğu fA-M kestirimine yakın bir düzeyde ÇBMTK'nin başarılı kestirimi dikkat çekmektedir. PG-DINA'ya ait DSO'lara bakıldığında ise diğer modellerden daha düşük performans gösterdiği görülmektedir. Yetenekler arası korelasyon varlığının DSO'ları etkilemediği görülmüştür.

**Tablo 4.19: Yetenekler Arası Korelasyona Göre fA-M Verileri DSO'ları**

			Vektör				Nitelik			
			0		0,6		0		0,6	
			Tam	±1	Tam	±1	Tam	±1	Tam	±1
KİMLİK	ÖZDEŞ	TE	0,61	0,28	0,62	0,27	0,76	0,23	0,76	0,22
	ÖZDEŞ	TEO	0,54	0,34	0,55	0,33	0,72	0,26	0,73	0,25
	ÇBMTK	TE	0,55	0,33	0,55	0,33	0,72	0,27	0,72	0,27
	ÇBMTK	TEO	0,50	0,37	0,50	0,37	0,70	0,28	0,70	0,29
	pG-DINA	TE	0,30	0,43	0,29	0,44	0,55	0,37	0,54	0,38
	pG-DINA	TEO	0,30	0,40	0,30	0,39	0,54	0,36	0,55	0,35
	LOG	TE	0,58	0,31	0,59	0,31	0,74	0,24	0,75	0,22
	LOG	TEO	0,56	0,32	0,56	0,32	0,73	0,25	0,73	0,25
	ÇBMTK	TE	0,52	0,36	0,52	0,36	0,71	0,27	0,71	0,27
	ÇBMTK	TEO	0,55	0,34	0,55	0,34	0,73	0,26	0,73	0,26
LOG	pG-DINA	TE	0,26	0,40	0,25	0,40	0,52	0,37	0,52	0,36
	pG-DINA	TEO	0,30	0,39	0,30	0,39	0,54	0,36	0,54	0,36
	LOGIT	TE	0,60	0,27	0,60	0,27	0,74	0,24	0,74	0,24
	LOGIT	TEO	0,56	0,32	0,56	0,32	0,73	0,25	0,73	0,25
LOGIT	ÇBMTK	TE	0,57	0,32	0,57	0,32	0,73	0,26	0,73	0,26
	ÇBMTK	TEO	0,54	0,34	0,54	0,35	0,72	0,26	0,72	0,26
	pG-DINA	TE	0,27	0,39	0,27	0,39	0,53	0,36	0,53	0,37
	pG-DINA	TEO	0,30	0,39	0,30	0,39	0,55	0,36	0,54	0,36

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

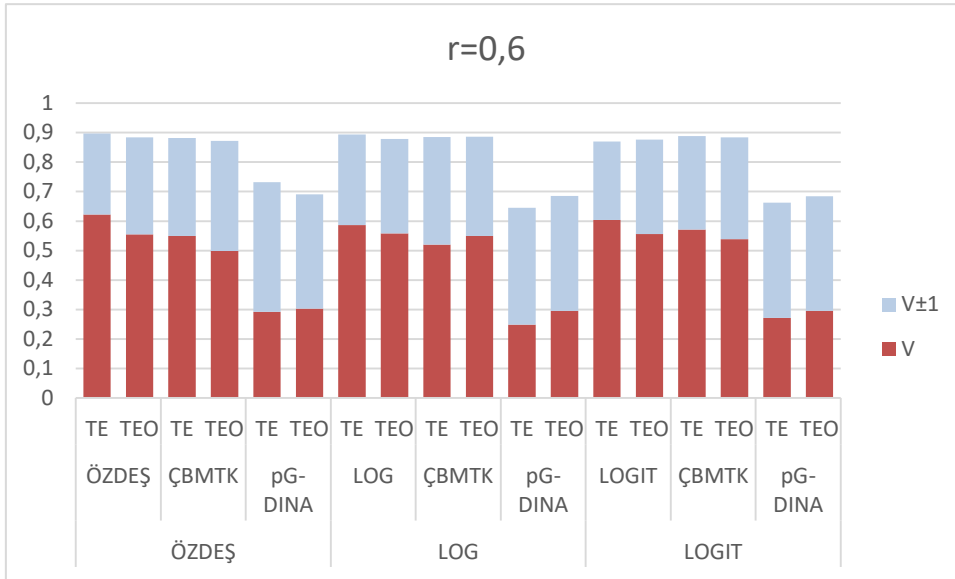
İlk alt problemde aynı senaryolar için doğru model olan ÇBMTK ve uyarlanan modellerden elde edilen DSO'lar incelendiğinde, fA-M verilerinin daha başarılı bir şekilde kestirildiği görülmektedir. Ancak ikinci alt problemdeki pG-DINA verisi kestirimi ile karşılaştırıldığında ise daha düşük doğrulukta kestirim yapıldığı görülmektedir. FA-M'lara ait tam vektör DSO'larına bakıldığında 0 korelasyon koşulunda, ilk alt problemdeki doğru model olan ÇBMTK sonuçlarından telafi edici veri için 0,1-0,13'lük, telafi edici olmayan veri içinse 0,07-0,09'lük oranda daha iyi olduğu görülmektedir. Yine aynı koşul için ikinci alt problemdeki pG-DINA kestirimi fA-M'lardan telafi edici veri için 0,1-0,13, telafi edici olmayan veri içinse 0,07-0,09'lük oranda daha iyi olduğu görülmektedir. İlk alt problemde en kötü doğrulamaya sahip pG-DINA fA-M verilerinde de en düşük performansı göstermiştir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.42. Yetenekler Arası 0 Korelasyon Koşulu fA-M Verisi DSO'ları**

Yetenekler arası korelasyonun 0 olduğu pG-DINA verisinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.42'deki grafikte görülmektedir. ÇBMTK hariç diğer modellerin telafi edici veri türünde daha iyi sonuçlar vermeye devam etmektedir. ÇBMTK ayrıca elde edilen sonuçlara bakıldığında daha düşük olan DSO'lar nedeniyle de diğer modellerden ayrılmaktadır. Diğer modellerde telafi edici veri için gerçekleşen 0,60'lık tam doğrulama oranı ÇBMTK'de 0,3 olarak gerçekleşmiştir. Telafi edici olmayan verinin analizinde ise ÇBMTK yine 0,3 iken diğer modeller 0,55 civarında doğru sınıflamaya sahiptir. Bu sonuçlar  $V_{\pm 1}$  hata payı eklendiğinde ÇBMTK'de 0,71 civarlarındayken diğer modellerde 0,88 dolaylarındadır. Ayrıca  $V_{\pm 1}$  eklendiğinde ÇBMTK haricindeki diğer modellerde aynı veri türünde görülen küçük farklar ortadan kaybolmuştur. Buna ilaveten aynı modelin telafi edici ve telafi edici olmayan veri analizleri arasındaki farkın da azaldığı görülmektedir.



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

#### Şekil 4.43. Yetenekler Arası 0,6 Korelasyon Koşulu fA-M Verisi DSO'ları

Yetenekler arası korelasyonun 0,6 olduğu pG-DINA verisinin analizleri sonucu vektör bazında elde edilen ortalama DSO'lar Şekil 4.43'te görülmektedir. Yetenekler arası korelasyon varlığının nitelik bazında DSO'ları etkilemediği belirtilmişti. Benzer durum, vektör sonuçlarının nitelik sonuçlarının kare değerleri olarak ele alınabileceğinden vektör temelinde de geçerlidir. 0 korelasyon koşulundaki değerler ve grafik fazla değişikliğe uğramadan 0,6 korelasyon koşulunda da elde edilmiştir. Bu açıdan korelasyon koşulu araştırma içindeki en zayıf koşul durumundadır.

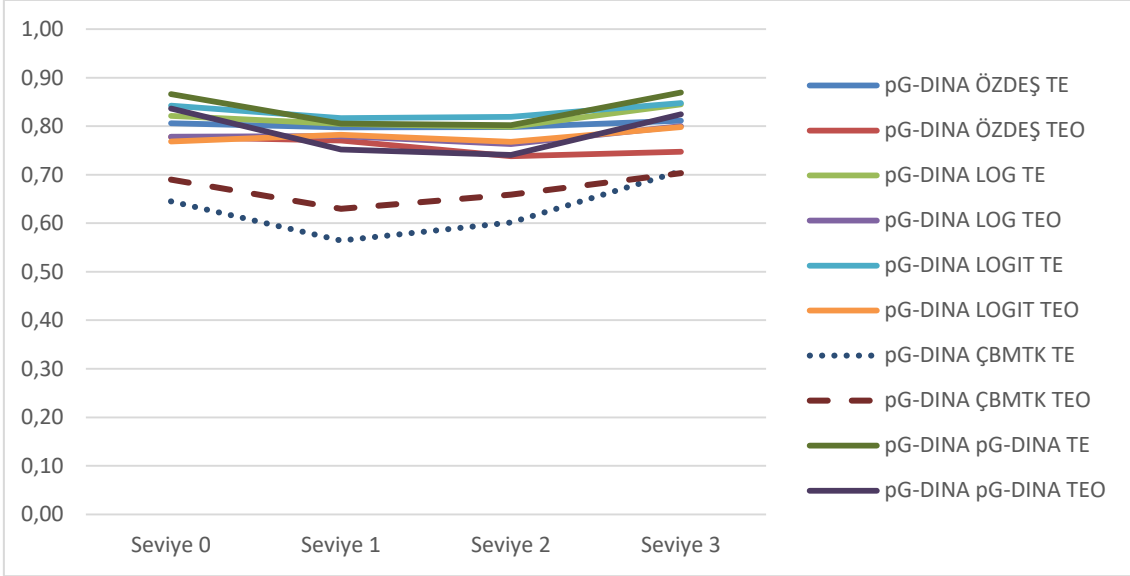
#### 4.2.5. Yetenek Düzeylerinde Doğru Sınıflama Oranları

İkinci alt problemin son sorusunda modellerin farklı yetenek düzeylerindeki bireyleri doğru sınıflama oranları incelenmiştir. Çıkan sonuçlar Tablo 4.20'de verilmiştir. Ayrıca sonuçların görselleştirilmesi adına grafik de kullanılmıştır.

Tablo 4.20: Yetenek Düzeylerinde pG-DINA DSO'ları

		Seviye 0	Seviye 1	Seviye 2	Seviye 3
pG-DINA	TE	0,87	0,81	0,80	0,87
pG-DINA	TEO	0,84	0,75	0,74	0,82
ÇBMTK	TE	0,65	0,56	0,60	0,71
ÇBMTK	TEO	0,69	0,63	0,66	0,70
ÖZDEŞ	TE	0,81	0,80	0,80	0,81
ÖZDEŞ	TEO	0,78	0,77	0,74	0,75
LOG	TE	0,82	0,80	0,80	0,85
LOG	TEO	0,78	0,78	0,76	0,80
LOGIT	TE	0,84	0,82	0,82	0,85
LOGIT	TEO	0,77	0,78	0,77	0,80

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.44. Yetenek Düzeylerinde pG-DINA Verisi DSO'ları**

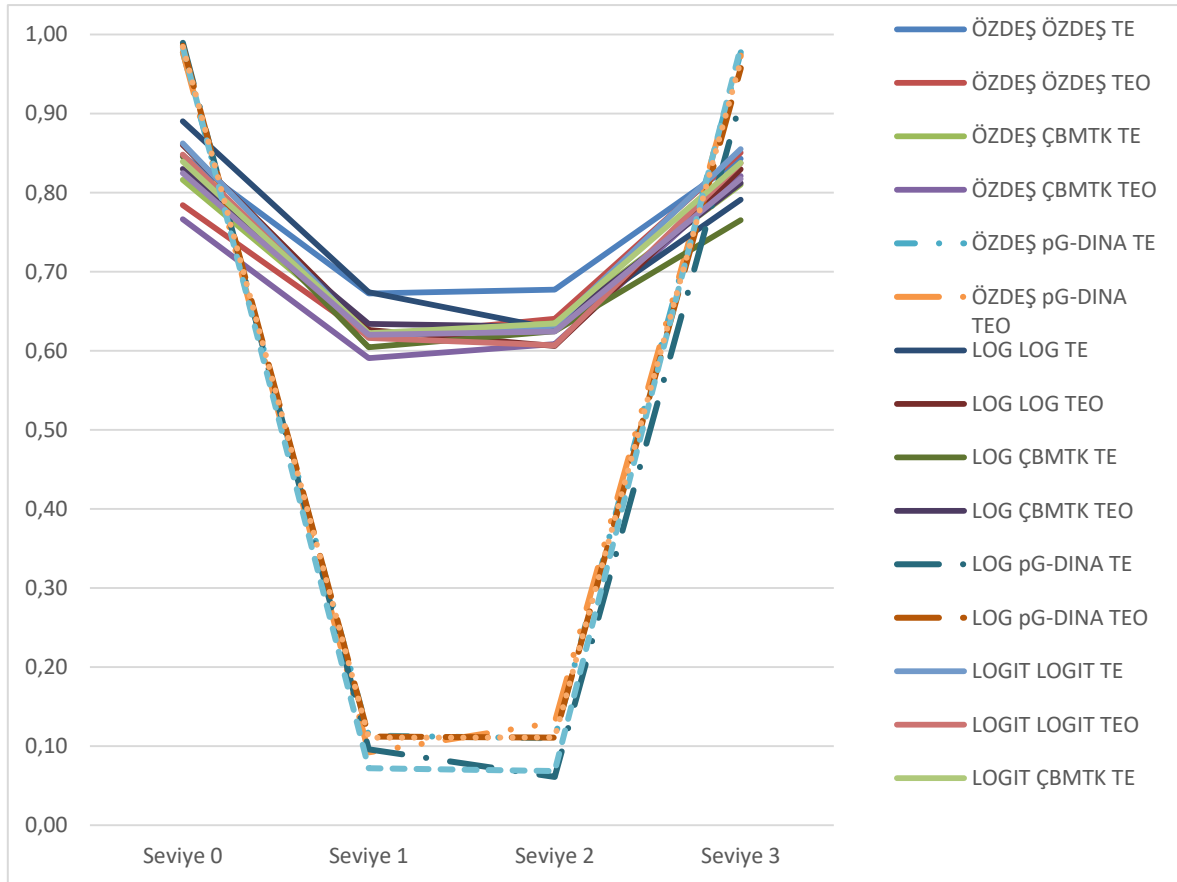
Tablo 4.20 ve Şekil 4.44 pG-DINA verisinin analizi sonucu modellerin yetenek düzeylerine göre DSO'larını göstermektedir. Tablonun görselleştirilmiş hali olan grafiğe bakıldığında iki çizgi haricindeki diğer göstergelerin birbirine çok yakın olduğu hatta iç içe geçtiği görülmektedir. Tablodan bakıldığında gruptan ayrılan ve daha düşük DSO'lara sahip olan analizin ÇBMTK analizi olduğu anlaşılmaktadır. Buradan da genelde ÇBMTK'nin pG-DINA verisi için en zayıf model olduğu çıkarımı yapılabilir. ÇBMTK ayrıca telafi edici olmayan veride daha yüksek doğrulama elde ederek diğer modellerden ayrılmaktadır. ÇBMTK harici model sonuçları daha stabil ve simetrik görünmektedir. ÇBMTK'de seviye 1 öncesi düşüş, sonrasında ise yükseliş eğilimi varken, diğer modellerde seviye 1 öncesi düşüş, sonrası durağan ve en son eski seviyeye kadar yükseliş görülmektedir. Ancak bu değişimlerin yüksek değerlerde olmadığını belirtmek gerekir. Bu yüzden pG-DINA verilerinin her düzeyde doğruluğunun birbirine yakın seviyede kestirildiği söylenebilir.

İlk alt problemdeki ÇBMTK verisinin yetenek düzeylerine göre DSO'larını gösteren tablo ile kıyasladığımızda, pG-DINA verisinde modellerin birbirine daha yakın bir oranda kestirimde bulunduğu görülmektedir. DSO'ların da ilk alt probleme nazaran çok üstünde olduğu ortadadır. Ayrıca ilk alt problemde pG-DINA'nın ortaya koyduğu görüntü gibi çok aykırı bir analiz sonucu olmaması da dikkat çekmektedir.

**Tablo 4.21: Yetenek Düzeylerinde fA-M Verileri DSO'ları**

Üretilen Model	Kestirilen Model	Veri Türü	Seviye 0	Seviye 1	Seviye 2	Seviye 3
ÖZDEŞ	ÖZDEŞ	TE	0,85	0,67	0,68	0,84
		TEO	0,78	0,62	0,64	0,85
	ÇBMTK	TE	0,82	0,62	0,63	0,81
		TEO	0,77	0,59	0,61	0,82
	pG-DINA	TE	0,98	0,11	0,11	0,98
		TEO	0,98	0,09	0,13	0,97
LOG	LOG	TE	0,89	0,67	0,63	0,79
		TEO	0,86	0,63	0,61	0,83
	ÇBMTK	TE	0,85	0,60	0,62	0,76
		TEO	0,83	0,63	0,63	0,81
	pG-DINA	TE	0,99	0,10	0,06	0,92
		TEO	0,99	0,11	0,11	0,96
LOGIT	LOGIT	TE	0,86	0,62	0,63	0,86
		TEO	0,85	0,62	0,61	0,84
	ÇBMTK	TE	0,84	0,62	0,63	0,84
		TEO	0,82	0,62	0,62	0,82
	pG-DINA	TE	0,98	0,07	0,07	0,98
		TEO	0,98	0,11	0,11	0,97

Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan



Not: TE= Telafi Edici, TEO= Telafi Edici Olmayan

**Şekil 4.45. Yetenek Düzeylerinde fA-M Verileri DSO'ları**

Tablo 4.21 ve Şekil 4.45'deki grafik fA-M verilerinin analizi sonucu ortaya çıkan yetenek düzeyine göre DSO'ları göstermektedir. Tablodaki değerleri görselleştiren grafiğe bakıldığında pG-DINA analizlerine ait çizgilerin farklı seviyeleri dikkat çekmektedir. Seviye 0'da %99 civarında DSO'ya sahip pG-DINA analizleri seviye 1'deki sınıflaması%10'lara düşmektedir. Seviye 2 içinde benzeri bir orandaki DSO seviye 3'de yine ilk başta olduğu gibi %98 düzeylerine gelmektedir. Diğer modeller olan fA-M ve ÇBMTK'nin gidişatı izlendiğinde 0,77-0,89 aralığında olan seviye 0'dan sonra her modelde düşüş görülmekte ve DSO'lar 0,59-0,67 aralığına gerilemektedir. Seviye 2'de de benzer oranla sahip olarak yatay olarak ilerleyen seyir seviye 3'de 0,76-0,84 oranına ulaşmaktadır. Yetenek düzeylerine göre bakıldığında pG-DINA ile diğer modeller arasında özellikle orta seviyelerde ciddi farklar bulunmaktadır. pG-DINA dışındaki diğer modeller de orta yetenek seviyelerinin kestiriminde uç seviyelere göre sıkıntılara sahipken bu sıkıntı pG-DINA'da çok daha ileri düzeydedir. Grafik ilk alt problem için aynı amaçla oluşturulan grafikte karşılaştırıldığında 2 farklı gruplaşma ve grupların eğilimleri açısından benzer formdadırlar. PG-DINA analizlerini gösteren değişimin çok olduğu grupların her iki grafikte de yetenek seviyeleri açısından benzer değerler aldığı görülmektedir. Diğer modelleri barındıran pG-DINA'ya göre fazla değişim göstermeyen gruptaki değerlere bakıldığında ise fA-M verilerinin uç değerlerde daha yüksek olmak üzere ÇBMTK verilerinden daha iyi oranda sınıflamaya imkân sağladığı görülmektedir. Grafik, ikinci alt problemin diğer verisi olan pG-DINA verilerinin analizine ait yetenek düzeyine göre DSO'ları gösteren grafikte karşılaştırıldığında ise genel olarak grafiğin daha hareketli ancak özellikle orta düzeylerde daha düşük DSO'lara sahip olduğu görülmektedir.



## 5. SONUÇ ve ÖNERİLER

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

### 5.1. Sonuçlar

İlk alt problemde ÇBMTK verisi üretilmiş, doğru model olarak ÇBMTK'de ve uyarlanan modeller olarak pG-DINA ve fA-M'larda analiz edilmiştir. Bu alt problem BTM'lerin ÇBMTK verisine uyarlanması olarak adlandırılabilir. Bu alt problemde en iyi DSO'lara sahip model, beklendiği üzere doğru model olan ÇBMTK iken fA-M'ların her koşulda ÇBMTK sonuçlarına yakın doğrulama oranı elde ettiği görülmüştür. pG-DINA ise bu iki modelden çok daha düşük performans göstermiştir. Bu sonuçlara göre ÇBMTK yakın bir BTM ailesi önerildiği ortaya konulmuştur.

Çalışmaya uyarlamamızın iyi örnekleri açısından bakıldığında ÇBMTK verisinden çok düzeyli bilişsel tanı bilgisi alınmak istendiğinde fA-M'ların iyi sonuç verdiği görülmüştür. Pratik uygulamaya bakıldığında bu uyarlama işlemi yalnızca yeteneklerin çok düzeyli niteliklere dönüştürülebileceği durumlarda yapılabilir. Çok düzeyli niteliklerin ÇBMTK'ye uyarlamamızında ise fA-M verileri ÇBMTK tarafından yüksek doğrulukla analiz edilmiştir. Bu uyarlama işlemi ise niteliklerin bilinen bir şekilde boyut/boyutlarda yer alabilmesi durumunda uygulanabilir. Bu şartların oluşmadığı durumlarda farklı modellerle uyarlama işlemi yapılmalıdır.

Telafi edici verinin analiz sonuçlarına ait doğru sınıflama oranlarının fA-M log linki haricinde biraz daha yüksek olduğu görülmüştür. Bu sonuç Bolt ve Lall (2003) ve Spray ve diğerleri (1990) tarafından elde edilen sonuçlara benzerdir. FA-M log linki formülünün çarpım etkisi nedeniyle telafi edici olmayan model formülüne yakın olduğu ve bu veriye daha uygun sonuçlar etmesi beklentiler içerisinde. Madde ve testin teorik altyapısının tam olarak bilinmediği durumlarda her iki veri türünde de iyi sonuçlar vermesi nedeniyle fA-M logit linki tercih edilebilir. Özdeş linki ise madde parametrelerinin kolay yorumlanması ile diğer linklerden üstün olsa da veri türlerine göre performansta ön plana çıkmamaktadır.

ÇBMTK'nin analiz sonuçları incelendiğinde telafi edici veri ile telafi edici olmayan veri arasındaki farkın telafi edici veri lehine olduğu görülmüştür. Ancak bu farkın çok yüksek olmadığı görülmüştür. Çoğu koşulda birbirine çok yakın sonuçlar elde

edilmiştir. Literatürde (Chalmers & Flora, 2014) dile getirilen telafi edici olmayan model kestirim sıkıntılarının çok derin olmadığı görülmüştür. Uygulamada görülen telafi edici model üstünlüğünün ve sık kullanımının kestirim başarısından çok testlerin teorik altyapısından kaynaklandığı öne sürülebilir.

Verilerin üretildiği model ile analizleri sonuçlarına bakıldığında ÇBMTK'nin diğer doğru modeller olan pG-DINA ve fA-M'lardan daha düşük doğrulama oranına sahip olduğu görülmüştür. Bu sonuç ÇBMTK verisi kestiriminin diğer model verilerine göre daha zor olduğu sonucunu doğurmaktadır. Verinin üretildiği türdeki modellerle kestirimin gerçekleştirilmesine rağmen ortaya çıkan kestirim kalitesi gerçek veride veri türünün bilinmemesi durumunda daha da düşük olabilir.

ÇBMTK'nin birey parametrelerinin sürekli yapıda olduğu ancak sonuçların karşılaştırabilmek adına kesikli hale getirildiği göz ardı edilmemelidir. Bu dönüşüm nedeniyle bilgi kaybı yaşanmıştır. Ayrıca dönüşümdeki kesme puanları kesin olması nedeniyle gerçekte kesme puanına yakın bir konumda olan birey yeteneği küçük hata payı ile kestirilmesine rağmen kesme puanının diğer tarafında olabilir. Bu şekilde küçük hatalar bile kesikli yapıda büyük görülebilir. Sürekli verinin sürekli veri ile karşılaştırılması daha sonuçları anlamlı kılacaktır.

Araştırmada farklılaştırılan koşullara bakıldığında ÇBMTK analizinde en yüksek verim test uzunluğu koşulundan elde edilmiştir. Bu koşulu sırasıyla madde ayırt edicilik ve madde yapısı oranı izlemiştir. Yetenekler arası korelasyon koşulunun ise diğer koşullara göre etkisi sınırlı kalmıştır. Ancak pG-DINA başta olmak üzere BTM'lerde koşulların etki büyüklüklerinin farklılaşabileceği bulgularda ortaya çıkmıştır.

Koşul değişimlerine diğer modellere göre farklılaşan tepkilere bakıldığında; test uzunluğu ve ayırt edicilik arttıkça modeller özellikle de pG-DINA ve fA-M log linki telafi edici olmayan verilerde daha iyi çalışmıştır. Testin basit madde yapısı oranı arttıkça fA-M log linkinde görülen telafi edici olmayan veri analiz sonuçları lehine olan fark giderek azalmıştır. Korelasyon artışı telafi edici verinin analizinde doğru sınıflamayı daha iyi artırmaktadır, özellikle fA-M'da bu durum daha belirgin olduğu gözlenmiştir.

PG-DINA'nın gösterdiği kötü performansın orta düzeydeki bireylerde görülen çok kötü DSO'lara bağlı olduğu, yetenek düzeylerine göre sınıflama oranlarında ortaya

çıkmiştir. PG-DINA kendi verisi haricindeki diğer verilerde orta düzeydeki bireyleri kestirimde çok büyük oranda en yakın uç düzeyde kestirmiştir. Bu sebeple  $V \pm 1$  hata payı değerleri diğer modellerden yüksek gerçekleşmiştir. PG-DINA'nın kendi verisi dışındaki veri analizlerinde orta düzeydeki bireyleri kestirimde son derece başarısız olduğu, uç düzeylerdeki başarısını ise tüm grubu merkezden uzaklaştırarak uç düzeylere yığmasına sayesinde gerçekleştirmiştir.

Çalışmada normal dağılımdaki yetenekler uniform dağılım oluşturacak şekilde kesme puanları ile kesikli hale getirilmiştir. Buna göre tüm düzeylerde yaklaşık eşit sayıda birey bulunmaktadır. Ancak uygulamada karşılaşılması kuvvetle muhtemel kategorik yeteneklerin de normal dağılıma yakın olması halinde pG-DINA'nın çok daha düşük DSO'lara sahip olacağı ön görülmektedir. Bu yüzden pG-DINA'nın kendi modeline ait olmayan gerçek veride kullanılması daha kötü sonuçlara neden olabilir.

İkinci alt problemde çalışmada ele alınan BTM'ler olan pG-DINA ve fA-M'lara ait veriler üretilmiş ve ÇBMTK, pG-DINA ve fA-M'larda analizleri gerçekleştirilmiştir. Bu alt problem ÇBMTK'nin BTM verisine uyarlanması içermektedir. PG-DINA verisinin analizine bakıldığında, pG-DINA'nın kendi verisini çok iyi bir oranda doğru kestirdiği görülmüştür. Bu oranlar çalışmada elde edilen en yüksek değerlere erişmiştir. FA-M'lardan elde edilen oranların pG-DINA'ya yakın olduğu görülmüştür. ÇBMTK ise bu modellerden daha düşük oranlara sahiptir.

PG-DINA verisinin çalışmada önerilen fA-M'lar tarafından başarılı kestirimi, pG-DINA'da var olan etkileşim etkisinin fA-M'da olmamasına rağmen her düzey için kestirilen etkilerce ikame edilmesi sayesinde gerçekleşmiştir. Modellerde birebir aynı madde parametreleri olmasa dahi farklı parametreler model değişiminde oluşan boşluğu doldurabilmektedir. Bu durum fazla sayıda madde parametresi içeren modellerin bir avantajı olarak değerlendirilebilir.

PG-DINA verisine en kötü kestirimde bulunan ÇBMTK'nin kendi verisinin analizi sonucu elde ettiği değerlere yakın düzeyde DSO'ya sahip olduğu görülmüştür. Bu açıdan bakıldığında ÇBMTK'nin pG-DINA verisinde başarısız olduğu söylenemeyebilir. Ancak ÇBMTK'nin fA-M verilerinde elde ettiği yüksek DSO'lar sorunun ÇBMTK analizinde olmadığını ortaya koymuştur.

PG-DINA ve ÇBMTK karşılıklı uyarlama çalışmalarının her ikisi de düşük doğrulama oranlarına sahiptir. Bu yüzden bu iki modelin birbiri yerine kullanımının sorunlu

olduğu görülmüştür. PG-DINA ve ÇBMTK verilerinin fA-M'lar tarafından başarılı kestirimi fA-M'ların pG-DINA ve ÇBMTK'ya başarılı bir şekilde uyarlanabildiğini ve gerekli durumlarda bu modeller yerine kullanılabileceğini göstermiştir.

PG-DINA verisi analizinde koşullara göre ortaya çıkan farklı sonuçlara bakıldığında, madde ayırt edicilik artışının tüm modellerde telafi edici olmayan DSO'larına daha iyi yansıdığı, pG-DINA telafi edici veri sonuçlarına etki etmediği görülmüştür. Madde yapısındaki basitleşme ÇBMTK'de önce telafi edici sonra telafi edici olmayan veriye avantaj getirirken, diğer modellerde telafi edici olmayan için daha iyi artışı beraberinde getirmektedir. ÇBMTK verisi telafi edici olmayan veride iyi iken diğer modeller telafi edici veride daha iyi performans göstermektedir. Dikkat çeken diğer bir sonuç ise, madde yapısı oranı koşulunun pG-DINA verisinde diğer model verilerine kıyasla etkisini yitirmesidir.

İkinci alt problem için üretilen BTM verilerinden biri de fA-M'dır. Tez kapsamında önerilen fA-M kapsamında 3 link için veriler ayrı ayrı üretilmiştir. Veriler ÇBMTK, pG-DINA ve fA-M'ın üretim için kullanılan link ile analiz edilmiştir. Beklendiği üzere doğru model olan fA-M bu veri için en iyi sınıflama oranına sahiptir. ÇBMTK ise az bir farkla fA-M sonuçlarını izlemektedir. Bu sonuç ÇBMTK verisi analizinde ortaya çıkan duruma benzerlik göstermektedir. Ortaya çıkan sonuçlar gerekli durumlarda ÇBMTK ve fA-M'ların birbiri yerine kullanılabileceğini göstermiştir. pG-DINA sonuçları ÇBMTK verisi analizinde olduğu gibi yine en kötü oranlara sahip olduğu görülmüştür. Bu sonuç ilk araştırma problemindeki sonucu destekler niteliktedir. ÇBMTK ve pG-DINA'nın karşılıklı uyarlanmasında elde edilen sonuçlar istenilen düzeyde değildir.

PG-DINA'daki etkileşim etkisi diğer modellere ait verilerin pG-DINA'da analizinde orta sınıflardaki bireylerin uç kesimlerde tahmin edilmesine yol açabilir. Seviye 1'deki birey negatif etkileşim etkisi nedeniyle Seviye 0'a, Seviye 2'deki birey ise pozitif etkileşim etkisi nedeniyle Seviye 3'te kestirildiği izlenimi vermektedir. Bu nedenle pG-DINA'nın diğer modellere ait verilerde kullanımına dikkat edilmelidir.

Verilerin analizinde verinin ait olduğu model diğer modellerden başarılı olmuştur. fA-M ise kendine ait olmayan her iki veride de verinin ait olduğu modele yakın performans sergilemiştir. Bu durum fA-M'ların başarılı bir model olduğunu ve başka modellere ait olan verilerde dahi kullanılabileceğini göstermektedir.

Ayrıca her modelin kendi verisinde en iyi performansı göstermesi beklenmektedir. Ancak fA-M analizlerinin pG-DINA verisinde elde ettiği DSO'lar beklenmeyen bir şekilde fA-M verisinde elde ettiğinden daha yüksektir. Benzer bir durum ÇBMTK için geçerlidir. ÇBMTK analizinin ÇBMTK verisinde performansı fA-M verisinde gösterdiğinden daha düşüktür. Bu durumun sebebinin modellerin kendi verisinden daha basit verilerde daha iyi kestirim yapması olarak düşünülmektedir. Bu durumun gerçekleşmesi için model-veri uyumunun da sağlanması gerekmektedir. Bu koşulun yerine gelmemesi nedeniyle ÇBMTK pG-DINA verisinde kötü bir kestirime sahiptir. Veriler yapılarına göre basitten karmaşığa göre pG-DINA, fA-M ve ÇBMTK olarak sıralanabilir.

Diğer verilerin analizlerin yetenekler arası korelasyon varlığının sonuçlar üzerinde kısmi de olsa pozitif etkisi görülmüştür. Ancak fA-M analizlerinde elde edilen DSO'ları etkilememiştir.

## **5.2. Öneriler**

Bu bölümde araştırmadan çıkan sonuçlara göre ileriye dönük olarak gerçekleştirilebilecek uygulamalar ve düzenlemelere değinilmiştir.

### **5.2.1. Araştırmaya Dönük Öneriler**

Çalışmada boyut sayısı bir koşul olarak ele alınmamıştır. Boyut sayısı da koşul olarak ele alınarak, üç ve daha fazla boyutlu yapılar için model sonuçlarını görmek için araştırma tekrarlanabilir.

Önerilen fA-M'lar toplama özellikli kısıtlanmış modelleri içermektedir. Aynı olasılık hesaplama tekniğine sahip fA-M'ın genelleştirilmiş model üretimi üzerinde çalışılabilir.

Araştırmada pG-DINA diğer BTM olarak ele alınmıştır. Analizler başka çok kategorili niteliklere uyumlu BTM'lerle tekrarlanabilir.

Çalışmada test dizaynı nitelik/boyutların simetrik olarak yer almasını sağlayacak şekilde biçimlendirildi. Pratikte her zaman gerçekleşmeyecek bu dizayn nitelik/boyutların birbirinden farklı ağırlık içecek şekilde asimetrik olarak gerçekleştirilebilir.

Çok kategoriye sahip niteliklerin her ikisi de 4 kategorili olarak tasarlanmıştır. Nitelik kategorileri birbirinden ve dörtten farklı olarak oluşturulabilir.

Çalışma yeni bir model ortaya koyduğundan türetilmiş veri üzerine kurulmuştur. Gerçek veri üzerinde çalışma tekrarlanarak, önerilen fA-M'lerin gerçek uygulamalarda başarısı izlenebilir.

Modellerde madde parametrelerinin farklı olması nedeniyle BTM'lerin madde parametreleri ÇBMTK verisi üzerinden oluşturulmuştur. pG-DINA ve fA-M'lerin kendine özgü madde parametreleri tespit edilip veriler bunlar üzerinden oluşturulabilir.

Çalışmada kullanılan her iki nitelik de çok kategoriliye sahiptir. Uygulamalarda ise iki kategorili nitelikler yaygın olarak kullanılmaktadır. İkili ve çoklu kategoriye sahip niteliklerin her ikisi birden kullanım durumu üzerinde yoğunlaşılabilir.

### **5.2.2. Uygulamaya Dönük Öneriler**

Araştırmada uyarlama çalışması kategorik düzene göre yapılmıştır. Bilgi kaybını da beraberinde getirebilecek bu işlem yerine BTM'lerde bulunan sürekli yapıdaki sonsal olasılıklar kullanılarak uyarlama çalışmaları yapılabilir.

Boyut/boyutların çok düzeyli niteliğe dönüştürülebilmesi halinde BTM'nin ÇBMTK'ye uyarlanması pratikte aynen tekrarlanabilir. Uygulama öncesi boyut/nitelik dönüşümünün teorik temellere dayalı olması halinde uyarlama çalışması amacına ulaşabilecektir.

ÇBMTK'nin BTM verisine uyarlanması işleminde de benzer şekilde çok düzeyli niteliğin belli bir/birkaç boyuta aitliği sağlanmalıdır. Birden fazla niteliğin bir boyutta yer alması halinde bu uyarlama işlemi belirli ve sağlıklı sonuçlara sahip olmayabilir.

## KAYNAKÇA

- 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.
- Adams, R. J., Wilson, M., & Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied psychological measurement*, 21(1), 1-23.
- Başokçu, O. T. (2011). *Geleneksel Yöntemler ve DINA Modele Göre Yapılan Sınıflamaların Geçerliliğinin Karşılaştırılması*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Yayınlanmamış Doktora Tezi.
- Başokçu, O. (2012). DINA Model Parametreleri Kullanılarak Tahminlenen Madde Ayırıcılık İndekslerinin İncelenmesi. *Eğitim ve Bilim*, 37(163), 310-321.
- Bradshaw, L., Izsák, A., Templin, J., & Jacobson, E. (2014). Diagnosing Teachers' Understandings of Rational Numbers: Building a Multidimensional Test Within the Diagnostic Classification Framework. *Educational measurement: Issues and practice*, 33(1), 2-14.
- Bradshaw, L., & Templin, J. (2014). Combining item response theory and diagnostic classification models: A psychometric model for scaling ability and diagnosing misconceptions. *Psychometrika*, 79(3), 403-425.
- 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.
- Buck, G., & Tatsuoka, K. (1998). Application of the rule-space procedure to language testing: Examining attributes of a free response listening test. *Language testing*, 15(2), 119-157.
- Chalmers, R. P., (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, 48(6), 1-29.
- Chalmers, R. P., & Flora, D. B. (2014). Maximum-likelihood estimation of noncompensatory IRT models with the MH-RM algorithm. *Applied Psychological Measurement*, 38(5), 339-358.
- Chen, J., & de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, 37(6), 419-437.
- Chen, J.S., & de la Torre, J. (2014). A Procedure for Diagnostically Modeling Extant Large-Scale Assessment Data: The Case of the Programme for International Student Assessment in Reading. *Psychology*, 5(18), 1967-1978.
- Crocker, L. M., & Algina, J. (1986). *Introduction to classical and modern test theory* (Vol. 6277). New York: Holt, Rinehart and Winston.
- de Ayala, R. J. (2009). *The theory and practice of item response theory*. Guilford Publications.

- de la Torre, J. (2009). A cognitive diagnosis model for cognitively based multiple-choice options. *Applied Psychological Measurement*, 33(3), 163-183.
- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76(2), 179-199.
- de la Torre, J. & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3), 333-353.
- de la Torre, J., & Karelitz, T. M. (2009). Impact of diagnosticity on the adequacy of models for cognitive diagnosis under a linear attribute structure: A simulation study. *Journal of Educational Measurement*, 46(4), 450-469.
- de la Torre, J. & Minchen N. (2014). Cognitively Diagnostic Assessments and the Cognitive Diagnosis Model Framework. *Psicología Educativa*, 20(2) 89-97.
- de la Torre, J., & Patz, R. J. (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.
- Demir, E. K. (2013). *Dina Model İle Geliştirilen Bir Testin Psikometrik Özelliklerinin Belirlenmesi*. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.
- DiBello, L.V. Roussos L. A. & Stout, W. (2007). *Review of Cognitively Diagnostic Assessment and a Summary of Psychometric Models*. Rao,C. Sinharay, S. (Eds.) Handbook of Statistics, Psychometrics. Vol. 26. North-Holland: Amsterdam.
- DiBello L.V. & Stout, W. (2007). Guest Editors' Introduction and Overview: IRT-Based Cognitive Diagnostic Models and Related Methods. *Journal of Educational Measurement*, Winter, Vol. 44, No. 4, 285–291.
- Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for Psychologists*. ABD: Psychology.
- Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26(4), 333-352.
- Haladyna, T. M., & Kramer, G. A. (2004). The validity of subscores for a credentialing test. *Evaluation & the health professions*, 27(4), 349-368.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications* (Vol. 7). Springer Science & Business Ortaia.
- Hartig, J., & Höhler, J. (2008). Representation of competencies in multidimensional IRT models with within-item and between-item multidimensionality. *Zeitschrift für Psychologie/Journal of Psychology*, 216(2), 89-101.
- Hartz, S. M. (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality*, Yayınlanmamış Doktora tezi, University of Illinois at Urbana-Champaign.
- Huebner, A., & Wang, C. (2011). A note on comparing examinee classification methods for cognitive diagnosis models. *Educational and Psychological Measurement*, 71(2), 407-419.




- Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika*, 74(2), 191-210.
- Hong, H., Wang, C., Lim, Y. S., & Douglas, J. (2014). Efficient models for cognitive diagnosis with continuous and mixed-type latent variables. *Applied Psychological Measurement*, 39(1), 31-43.
- Johnson, S. G., The NLOpt nonlinear-optimization package, <http://ab-initio.mit.edu/nlopt>
- Junker, B. (1999). *Some statistical models and computational methods that may be useful for cognitively relevant assessment*. Pittsburg: National Research Council.
- Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with non-parametric item response theory. *Applied Psychological Measurement*, 25(3), 258-272.
- Karelitz, T. M. (2004). *Ordered-category attribute coding framework for cognitive assessments*, University of Illinois at Urbana-Champaign, Yayınlanmamış doktora tezi.
- Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2009). A practical illustration of multidimensional diagnostic skills profiling: Comparing results from confirmatory factor analysis and diagnostic classification models. *Studies in Educational Evaluation*, 35(2), 64-70.
- Lee, Y. S., Park, Y. S., & Taylan, D. (2011). A cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the US national sample using the TIMSS 2007. *International Journal of Testing*, 11(2), 144-177.
- Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The Attribute Hierarchy Method for Cognitive Assessment: A Variation on Tatsuoka's Rule-Space Approach. *Journal of educational measurement*, 41(3), 205-237.
- Ma, W. & de la Torre, J. (2016). GDINA: The Generalized DINA model framework. R package version 0.9.2.
- Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, 64(2), 187-212.
- McKinley RL, Reckase MD (1982) *The use of the general Rasch model with multidimensional item response data* (Araştırma Raporu ONR 82-1). American College Testing, Iowa City, IA.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design. ETS Research Report Series, New Jersey.
- Pellegrino, J.W., Chudowsky, N., Glaser, R. (2001). *Knowing what Students Know: The Science and Design of Educational Assessment*. National Academy Press, Washington, DC.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement*, 9(4), 401-412.
- Reckase, M. D. (1997a). The past and future of multidimensional item response theory. *Applied Psychological Measurement*, 21(1), 25-36.
- Reckase, M. D. (1997b). A linear logistic multidimensional model for dichotomous item response data. *Handbook of modern item response theory* (s. 271-286). Springer New York.
- Reckase, M. D. (2007). Multidimensional Item Response Theory. Rao, C. Sinharay, S. (Ed.) *Handbook of Statistics, Psychometrics*. Vol. 26. North-Holland: Amsterdam.
- Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer.
- Reckase, M. D., & McKinley, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement*, 15(4), 361-373.
- Robitzsch A., Kiefer T., George A. C. & Uenlue A. (2016). CDM: Cognitive Diagnosis Modeling. R package version 4.7-0. <http://CRAN.R-project.org/package=CDM>
- Rupp, A. A., & Templin, J. (2008). Unique characteristics of diagnostic models: A review of the current state-of-the-art. *Measurement*, 6(4), 219-262.
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications*. Guilford.
- Sinharay, S., Puhan, G., & Haberman, S. (April, 2009). *Reporting diagnostic scores: Temptations, pitfalls, and some solutions*. National Council on Measurement in Education kongresinde sunulmuştur, San Diego, ABD.
- Spray, J. A., Davey T. C., Reckase M. D., Ackerman T. A., Carlson J. E. (1990). *Comparison of Two Logistic Multidimensional Item Response Theory Models*. Research Report ONR90-8.
- Sünbül, S. Ö. (2013). *Bilişsel Tanı Modellerinde Parametre Kestirimini Ve Sınıflama Tutarlılığını Etkileyen Faktörlerin İncelenmesi*. Mersin Üniversitesi Eğitim Bilimleri Enstitüsü, Yayınlanmamış Doktora Tezi.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20(4), 345-354.
- Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnosis. *Diagnostic monitoring of skill and knowledge acquisition* (453-488), Lawrence Erlbaum Associates.
- Templin, J., & Bradshaw, L. (2014). Hierarchical diagnostic classification models: A family of models for estimating and testing attribute hierarchies. *Psychometrika*, 79(2), 317-339.
- Templin, J., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, 11(3), 287-305.
- von Davier, M., Yamamoto, K. (2004). *A class of models for cognitive diagnosis*. ETS Spearman Konferansında sunulmuştur, The Inn at Penn, Philadelphia, PA, ABD.

- von Davier, M. (2008). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, 61(2), 287-307.
- Wang, W. C., & Wilson, M. (2005). The Rasch Testlet Model. *Applied Psychological Measurement*.
- Wang, Ying-Chen,(2009). *Factor Analytic Models and Cognitive Diagnostic Models: How Comparable Are They?—A Comparison of R-RUM and Compensatory MIRT Model with Respect to Cognitive Feedback*. Basılmamış Doktora Tezi, the Faculty of The Graduate School at The University of North Carolina at Greensboro).
- Whitely, S. E. (1980). Multicomponent latent trait models for ability tests. *Psychometrika*, 45(4), 479-494.

## **EKLER DİZİNİ**

# EK 1. ETİK KOMİSYONU ONAY BİLDİRİMİ

Form 40	
<b>Tez Çalışması Etik Komisyon İzin Muafiyeti Formu</b>	
03 / 02 / 2017	
Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Anabilim Dalı Başkanlığı'na	
<b>Tez Başlığı / Konusu:</b>	BİLİŞSEL TANI VE ÇOK BOYUTLU MADDE TEPKİ KURAMI MODELLERİNİN KARŞILIKLI UYUMLARININ İNCELENMESİ
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 Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre tez çalışmamın yürütülebilmesi için herhangi bir Etik Komisyondan/Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.	
Gereğini saygılarımla arz ederim.	
 Levent YAKAR (Öğrencisi Adı Soyadı, İmzası)	
<b>Öğrenci Bilgileri</b>	
<b>Adı Soyadı</b>	Levent YAKAR
<b>Öğrenci No</b>	N11248312
<b>Anabilim Dalı</b>	Eğitim Bilimleri
<b>Programı</b>	Eğitimde Ölçme ve Değerlendirme
<b>Statüsü</b>	<input type="checkbox"/> Yüksek Lisans <input checked="" type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr.
<b>Danışman Görüşü ve Onayı</b>	
Tezde simülatif veri kullanılmıştır. Etik sorun oluşturacak bir durum mevcut değildir.	
 Prof. Dr. Nuri DOĞAN (İmza) (Danışmanın Unvanı, Adı ve Soyadı)	

## EK 2. ORJİNALLİK RAPORU



### HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ YÜKSEK LİSANS/DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

#### HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLER ENSTİTÜSÜ EĞİTİM BİLİMLERİ ANA BİLİM / BİLİM DALI BAŞKANLIĞI'NA

Tarih: 23/02/2017

Tez Başlığı : BİLİŞSEL TANI VE ÇOK BOYUTLU MADDE TEPKİ KURAMI MODELLERİNİN KARŞILIKLI UYUMLARININ İNCELENMESİ

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

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Endeksi	Gönderim Numarası
23/02/2017	137	214429	03/02/2017	%4	774980871

Uygulanan filtreler:

- 1- Kaynakça hariç
- 2- Alıntılar dâhil
- 3- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

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

Gereğini saygılarımla arz ederim.

Tarih ve İmza

Adı Soyadı: LEVENT YAKAR

Öğrenci No: N11248312

Anabilim Dalı: EĞİTİM BİLİMLERİ

Programı: EĞİTİMDE ÖLÇME VE DEĞERLENDİRME

Statüsü:  Y.Lisans  Doktora  Bütünleşik Dr.

### DANIŞMAN ONAYI

UYGUNDUR.  
(Prof. Dr. NURİ DOĞAN)

## ÖZGEÇMİŞ

### Kişisel Bilgiler

<i>Adı Soyadı</i>	Levent YAKAR
<i>Doğum Yeri</i>	Kayseri
<i>Doğum Tarihi</i>	01.01.1984

### Eğitim Durumu

<i>Lise</i>	Arif Molu Anadolu Meslek Lisesi - Kayseri	2002
<i>Lisans-1</i>	Gazi Üniversitesi-End.San.Eğitim Fakültesi	2007
<i>Lisans-2</i>	Anadolu Üni.- İktisat Fak.- Kamu Yönetimi	2008
<i>Yüksek Lisans</i>	Abant İzzet Baysal Üni.-Eğitim Bil.Ens.-Eğt.Ölç.Değ.	2010
<i>Yabancı Dil</i>	İngilizce: Okuma (Çok iyi), Yazma (İyi), Konuşma (İyi)	

### İş Deneyimi

<i>Stajlar</i>	Rutgers Üniversitesi – New Jersey	10 ay
<i>Projeler</i>	Sigara Paketleri Üzerindeki Yazılı ve Birlesik Uyarıların Üniversite Gençleri Üzerindeki Etkisinin Farklı Değişkenler Bakımından İncelenmesi	2014-2015
<i>Çalıştığı Kurumlar</i>	Abant İzzet Baysal Üniversitesi (Memur) Milli Eğitim Bakanlığı (Öğretmen)	2004-2010 2010-2012

### Akademik Çalışmalar

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

- Yakar, L., Odabaş, M. , Gündeğer, C. (2016). Üniversite Öğrencilerinin Üniversiteyi Seçme Nedenlerinin İkili Karşılaştırma Yöntemiyle Ölçeklenmesi Hacettepe Siirt ve Aksaray Üniversiteleri Örneği. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 16(38), 189-200. (Ulusal Makale)
- Özdemir, A , Yakar, L. , Yavuz, S . (2017). Matematik Dersi Dönem Sonu Puanını Oluşturan Bileşenlerin İncelenmesi: Performans Görevlerinin Kaldırılması. *Uşak Üniversitesi Eğitim Araştırmaları Dergisi*, 3(1), 74-91. (Ulusal Makale)
- Asilkalkan A., Yakar L., Terzi R., (2016). *DIF and DDF Analyses on Turkish High School Entrance Exam*, Nera 2016, Boston, ABD. (Uluslararası Bildiri)
- Yakar, L. (2016). *Çok Boyutlu Madde Tepki Kuramındaki Değişik Faktörlerin Yetenek Parametresi Kestirimi Doğruluğuna Etkisi*, Epod 2016, Antalya. (Ulusal Bildiri)
- Toprak, E., Yakar L., (2016). *SBS 2011 Türkçe Alt Testindeki Maddelerin Cinsiyet Yanlılığı Açısından İncelenmesi*, Epod 2016, Antalya. (Ulusal Bildiri)

### İletişim

<i>e-Posta Adresi</i>	l_yakar@hotmail.com
<i>Jüri Tarihi</i>	03.02.2017