



**PARAMETRİK OLMAYAN BİLİŞSEL TANILAMA, YAPAY SİNİR
AĞI VE DİNO MODELİNİN SINIFLANDIRMA
PERFORMANSLARININ KARŞILAŞTIRILMASI**

Emine Yavuz

**DOKTORA TEZİ
EĞİTİM BİLİMLERİ ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

MART, 2021

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU

Bu tezin tüm hakları saklıdır. Kaynak göstermek koşuluyla tezin teslim tarihinden itibaren(....) ay sonra tezden fotokopi çekilebilir.

YAZARIN

Adı : Emine
Soyadı : Yavuz
Bölümü : Eğitimde Ölçme ve Değerlendirme
İmza :
Teslim tarihi :

TEZİN

Türkçe Adı: Parametrik Olmayan Bilişsel Tanılama, Yapay Sinir Ağı ve Dino Modelinin Sınıflandırma Performanslarının Karşılaştırılması

İngilizce Adı: Comparison of Classification Performances of Non-Parametric Cognitive Diagnosis, Artificial Neural Network And Dino Model

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik ilkelere uyduğumu, yararlandığım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttiğimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduğunu beyan ederim.

Emine YAVUZ

JÜRİ ONAY SAYFASI

Emine YAVUZ tarafından hazırlanan “Parametrik Olmayan Bilişsel Tanılama, Yapay Sinir Ağı ve DINO Modelinin Sınıflandırma Performanslarının Karşılaştırılması” adlı tez çalışması aşağıdaki jüri tarafından oy birliği ile Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme Anabilim Dalı’nda Doktora tezi olarak kabul edilmiştir.

Danışman: Prof. Dr. Hakan Yavuz ATAR

Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Gazi Üniversitesi

Başkan: Prof. Dr. Selahattin GELBAL

Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Hacettepe Üniversitesi

Üye: Prof. Dr. Şeref TAN

Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Gazi Üniversitesi

Üye: Prof. Dr. İsmail KARAKAYA

Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Gazi Üniversitesi

Üye: Prof. Dr. Nuri DOĞAN

Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Hacettepe Üniversitesi

Tez Savunma Tarihi: 10/02/2021

Bu tezin Eğitimde Ölçme ve Değerlendirme Anabilim Dalı’nda Doktora tezi olması için şartları yerine getirdiğini onaylıyorum.

Prof. Dr. Yücel GELİŞLİ

Eğitim Bilimleri Enstitüsü Müdürü

TEŞEKKÜR

Lisansüstü eğitimim süresince düşüncelerini ve yardımlarını benden esirgemeyerek her zaman yanımda olan ve kendimi geliştirmemi her zaman destekleyen, akademisyenliğe bakış açısından ilham aldığım değerli danışmanım Prof. Dr. Hakan Yavuz ATAR'a çok teşekkür ederim. Tez izleme komitemde ve tez jürimde bulunan, tezimin şekillenmesine yardımcı olup fikirleri ve bakış açılarıyla akademik hayatımda yol gösteren ve yardımlarını benden esirgemeyen saygıdeğer hocalarım Prof. Dr. Şeref TAN'a ve Prof. Dr. Selahattin GELBAL'a teşekkürlerimi sunarım. Tez jürimde bulunan, Prof. Dr. Nuri DOĞAN'a ve Prof. Dr. İsmail KARAKAYA'ya değerli katkılarından ötürü teşekkür ederim. Tez süresince akademik bilgilerini ve yardımlarını benden esirgemeyip sorularıma cevap veren Dr. Ying CUI'ye, Dr. Qi GUO'ya, Dr. Justin PAULSEN'a ve Dr. Alexander ROBİTZSCH'e teşekkür ederim.

Hayatım boyunca yanımda olup kararlarımı destekleyen babam Bayram YAVUZ'a, annem Şehriban YAVUZ'a ve sevgili teyzem Fadime DEMİRTAŞ'a teşekkürü ederim. Çalışmam sürecinde yanımda olup yardımlarını esirgemeyen ve her koşulda beni destekleyen değerli arkadaşlarım Pınar AKYILDIZ'a, Murat BORAN'a, Sema ÖZMEN'e ve Sibel ADA'ya teşekkür ederim.

Ayrıca doktora eğitimim süresince sağladığı maddi destekten ötürü Türkiye Bilimsel ve Teknolojik Araştırma Kurumu'na teşekkürlerimi sunarım. Bu aşamaya gelmemde etkisi olan, burada ismini sayamadığım bütün arkadaşlarıma ve saygı değer hocalarıma teşekkür ederim.

Emine YAVUZ

Mart, 2021

**PARAMETRİK OLMAYAN BİLİŞSEL TANILAMA, YAPAY SİNİR
AĞI VE DINO MODELİNİN SINIFLANDIRMA
PERFORMANSLARININ KARŞILAŞTIRILMASI
(Doktora Tezi)**

Emine Yavuz
GAZİ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
Mart, 2021

ÖZ

Yapılacak ölçme ve değerlendirme etkinliklerinde, amacına en iyi şekilde hizmet edecek yöntemlerin seçilmesi önemlidir. Alanyazında Bilişsel Tanı Modelleri'nin (BTM'lerin) küçük örneklem, az madde veya fazla niteliğin olduğu bazı koşullarda öğrenciler hakkında yanlış sonuçlar verebildiği görülmektedir. Bu yanlış sonuçlardan kaçınmak için kullanılan alternatif yöntemlerin sınıflandırma oranlarının farklı koşullarda birbirleriyle karşılaştırılıp hangi yöntemin hangi koşullar altında daha iyi performans gösterdiğinin belirlenmesini amaçlayan çalışmalara ihtiyaç vardır. Bu düşünceden hareketle mevcut araştırmada, Parametrik Olmayan Bilişsel Tanılamaya (POBT) ve Yapay Sinir Ağına (YSA) ait nitelik (NSO) ve örüntü (ÖSO) düzeyi sınıflandırma oranlarının madde ve nitelik sayısı, örneklem büyüklüğü ve kayıp veri oranı koşullarında Deterministic-Input, Noisy-Or Gate (DINO) modeli temelli simülasyon veri setlerinde öncelikle birbirleriyle, daha sonra DINO modeli ile karşılaştırılmaları amaçlanmaktadır. Ek olarak POBT, YSA ve DINO modelinin sınıflandırma oranlarının benzerliklerinin PISA 2015 işbirlikli problem çözme (İPÇ) veri setlerinde, nitelik sayısı ve örneklem büyüklüğü koşullarında incelenmesi amaçlanmaktadır. Simülasyon ve betimsel araştırma deseninin kullanıldığı bu araştırmada simülasyon veri setleri karmaşık Q matris yapısı ve telafisel modellerden DINO modeli temel alınarak üretilmiştir. Araştırmada gerçek veri setleri için PISA 2015 İPÇ uygulaması seçilmiştir ve örneklem büyüklüğü faktörünün koşulları PISA 2015 İPÇ uygulamasına

katılan 43 ülke ve 18170 öğrenci arasından basit seçkisiz yöntemle belirlenmiştir. Araştırmada ilk olarak simülasyon veri setleri üzerinde POBT'nin, YSA'nın ve DINO modelin nitelik (NSO) ve örüntü (ÖSO) düzeyi sınıflandırma oranları araştırma koşullarında ayrı ayrı incelenmiştir. Analizler sonucunda, nitelik sayısı (3, 5 ve 7) arttıkça POBT ve DINO modelinin NSO ve ÖSO'larının azaldığı, YSA'nın NSO'larının artarken ÖSO'larının azalıp arttığı görülmektedir. Örneklem büyüklüğünün (30, 100 ve 500) arttırılmasıyla POBT ve YSA'ya ait NSO ve ÖSO'ların değişimleri ile ilgili herhangi bir sistematikliğe rastlanmamaktadır ve örneklem büyüklüğünün artmasıyla DINO modeline ait NSO ve ÖSO'lar artmaktadır. Madde sayısının (15, 30 ve 45) arttırılmasıyla POBT ve DINO modeline ait NSO ve ÖSO'lar artarken YSA'ya ait NSO ve ÖSO'lar azalmaktadır. Kayıp veri oranının (0; 0,5 ve 0,10) arttırılmasıyla YSA'ya ait NSO ve ÖSO'ların bazı koşullarda değişmediği bazı koşullarda ise az miktarda azaldığı görülürken, POBT ve DINO modeline ait NSO ve ÖSO'lar tüm koşullarda azalmaktadır. Araştırmanın sonraki aşamasında POBT'ye, YSA'ya ve DINO modeline ait NSO ve ÖSO'ların araştırma koşullarında birlikte nasıl değiştikleri incelenmiştir. Yapılan faktöriyel ANOVA sonucunda tüm koşullarda POBT'nin, DINO modelinden biraz düşük fakat karşılaştırılabilir sınıflandırma oranlarına sahipken YSA'nın her zaman POBT ve DINO modelinden daha düşük oranlara sahip olduğu belirlenmiştir. Araştırmada son olarak PISA 2015 İPÇ veri setlerinden elde edilen POBT, YSA ve DINO modeline ait nitelik (NSOB) ve örüntü (ÖSOB) düzeyi sınıflandırma oranlarının benzerliği nitelik sayısı ve örneklem büyüklüğü koşullarında incelenmiştir. Analizler sonucunda nitelik sayısı (3, 7 ve 11) arttıkça POBT, YSA ve DINO modeline ait NSOB'ların genel olarak önce azalıp sonra artarken ÖSOB'ların genel olarak azaldığı belirlenmiştir. Örneklem büyüklüğünün (30, 100 ve 500) artmasıyla ise POBT, YSA ve DINO modeline ait NSOB ve ÖSOB'ların artması, azalması veya sabit kalmasıyla ilgili herhangi bir sistematikliğe rastlanmamıştır. Sonuç olarak, mevcut araştırmada POBT ile DINO modeli sınıflandırma oranları arasındaki benzerlik hem simülasyon hem de gerçek veri setlerinde gözlemlenmektedir. Ayrıca hem simülasyon hem de gerçek veri setlerinde örneklem büyüklüğünün artmasıyla POBT ve YSA'nın sınıflandırma oranlarının ve bu oranların benzerliklerinin artış veya azalışlarında herhangi bir sistematikliğe rastlanmamıştır. Mevcut araştırmada POBT ve YSA'nın bilişsel tanılamada yeni kullanılmaya başlanması nedeniyle çeşitli koşullarda POBT ve YSA'ya ait sınıflandırma oranlarının değişimleri hakkında genellenebilir sonuçlara ulaşılamamaktadır. Bu nedenle çeşitli koşullarda, çeşitli simülasyon ve gerçek veri setlerinde POBT ve YSA'ya ait NSO ve ÖSO'ların nasıl değiştiğinin incelenmesi önerilmektedir.

Anahtar Kelimeler : Tanılayıcı değerlendirme, biçimlendirici değerlendirme, parametrik olmayan bilişsel tanılama, yapay sinir ağı, DINO, PISA

Sayfa Adedi : xxi + 168

Danışman : Prof. Dr. Hakan Yavuz ATAR

COMPARISON OF CLASSIFICATION PERFORMANCES OF NON-PARAMETRIC COGNITIVE DIAGNOSIS, ARTIFICIAL NEURAL NETWORK AND DINO MODEL

(Ph.D Thesis)

Emine Yavuz

GAZI UNIVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

March, 2021

ABSTRACT

It is important to choose the methods that will best serve the purpose of measurement and evaluation activities. In the literature, it is seen that cognitive diagnosis models (CDMs) can give biased results about students in some conditions where there is a small sample, few items or many attributes. In order to avoid these biased results, there is a need for studies aiming to compare the classification rates of alternative methods used with each other under different conditions and to determine which method performs better under which conditions. Based on this idea, in this study, it is aimed to compare the attribute (ACR) and pattern-level (PCR) classification rates of Nonparametric Cognitive Diagnosis (NPCD) and Artificial Neural Networks (ANNs) on data sets of Deterministic-Input, Noisy-Or Gate (DINO) model-based simulation firstly with each other and then with DINO model under the conditions of the attribute number, sample size, item number and missing data rate. In addition, it is aimed to examine the similarities of the classification rates of NPCD, ANN and DINO model on the PISA 2015 collaborative problem-solving (CPS) data sets under the conditions of the attribute number and sample size. In this study, in which simulation and descriptive research design was used, simulation data sets were produced based on the complex Q matrix structure and the DINO model from compensatory models. PISA 2015 CPS competency was chosen for the real data sets and the conditions of the sample size factor were determined by simple random method among 43 countries and 18170 students participating in the PISA 2015 CPS application. In the research, firstly, the ACRs and PCRs of NPCD, ANN and DINO model on the simulation data sets were examined separately under the conditions of the present research. As a result

of the analysis, it is seen that as the number of attributes (3, 5 and 7) increases, the ACRs and PCRs of NPCD and DINO model decrease whereas the ACRs of ANN increase and the PCRs of ANN decrease and then increase. When the sample size is increased (30, 100 and 500), there is no systematicity regarding the changes of the ACRs and PCRs of NPCD and ANN whereas the ACRs and PCRs of DINO model increase as the sample size increases. As the number of items is increased (15, 30 and 45), the ACRs and PCRs of NPCD and DINO model increase while the ACRs and PCRs of ANN decrease. As the missing data rate is increased (0; 0,5 and 0,10), it is observed that the ACRs and PCRs of ANN do not change in some conditions, but decrease slightly in some conditions whereas the ACRs and PCRs of NPCD and DINO model decrease in all conditions. In the next stage of the research, how the ACRs and PCRs of NPCD, ANN and DINO model change together in these research conditions is examined. As a result of the factorial ANOVA, it is revealed that NPCD has slightly lower but comparable classification rates than DINO models in all conditions while ANN always has lower rates than NPCD and DINO models. Finally, in the study, the similarity of the attribute (SACR) and pattern-level classification rates (SPCR) of NPCD, ANN and DINO model obtained from the PISA 2015 CPS data sets is examined under the conditions of the attribute number and sample size. As a result of the analysis, it is determined that as the number of attributes (3, 7 and 11) increases, the SACRs of NPCD, ANN and DINO model generally decrease and then increase whereas the SPCRs generally decrease. As the sample size increases (30, 100 and 500), no systematicity is encountered regarding the increase, decrease or stay constant of the SACRs and SPCRs of NPCD, ANN and DINO model. As a result, the similarity between NPCD and DINO model classification rates is observed in both simulation and real data sets. In addition, regarding to the increase in sample size in both simulation and real data sets, no systematicity is found in the increase or decrease of the classification rates of NPCD and ANN and the similarities of these rates. In this study, since NPCD and ANN has been recently started to be used in cognitive diagnosis, generalizable results cannot be reached regarding the changes in the classification rates of NPCD and ANN under various conditions. For this reason, it is recommended to examine how the ACRs and PCRs of NPCD and ANN change in various simulation and real data sets under various conditions.

Key Words : Diagnostic assessment, formative assessment, nonparametric cognitive diagnosis, artificial neural network, DINO, PISA

Page Number : xxi + 168

Supervisor : Prof. Dr. Hakan Yavuz ATAR

İÇİNDEKİLER

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU	i
ETİK İLKELERE UYGUNLUK BEYANI.....	ii
JÜRİ ONAY SAYFASI.....	iii
TEŞEKKÜR.....	iv
ÖZ	v
ABSTRACT	vii
İÇİNDEKİLER.....	ix
TABLolar LİSTESİ.....	xiv
ŞEKİLLER LİSTESİ.....	xvii
SİMGELER VE KISALTMALAR LİSTESİ.....	xix
BÖLÜM I	1
GİRİŞ.....	1
1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı	8
1.3. Araştırmanın Önemi.....	9
1.4. Araştırmanın Sınırlılıkları	11
1.5. Tanımlar.....	12

BÖLÜM II.....	13
KAVRAMSAL ÇERÇEVE	13
2.1. Araştırmanın Kuramsal Temelleri.....	13
2.1.1. Bilişsel Tanı Modelleri (BTM'ler)	13
2.1.1.1. Deterministic-Input, Noisy-Or Gate (DINO)	17
2.1.2. Parametrik Olmayan Bilişsel Tanılama (POBT)	19
2.1.3. Yapay Sinir Ağları (YSA).....	22
2.1.3.1. Çok Katmanlı Algılayıcı (Multilayer perceptron) YSA	23
2.1.4. PISA 2015 İPÇ Yeterliği.....	28
2.2. İlgili Araştırmalar	31
2.2.1. Parametrik Olmayan Bilişsel Tanılama Yöntemleri ile İlgili Araştırmalar	31
2.2.2. PISA 2015 İPÇ Yeterliği ile İlgili Araştırmalar	39
BÖLÜM III	42
YÖNTEM.....	42
3.1. Araştırmanın Modeli.....	42
3.2. Çalışma Grubu	42
3.3. Veri Toplama Aracı ve Verilerin Elde Edilmesi	43
3.3.1. Simülasyon Veri Setlerinin Elde Edilmesi	44
3.3.2. Gerçek Veri Setleri İçin Veri Toplama Aracı ve Verilerin Elde Edilmesi	47
3.4. Verilerin Analizi	51
BÖLÜM IV	58
BULGULAR VE YORUM	58
4.1. Simülasyon Verisinde POBT'nin, YSA'nın ve DINO Modelinin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve	

500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması.....	58
4.1.1. POBT'nin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması	59
4.1.1.1. Değişen koşullar altında POBT'ye ait NSO'lar için faktöriyel ANOVA bulguları.....	61
4.1.1.2. Değişen koşullar altında POBT'ye ait ÖSO'lar için faktöriyel ANOVA bulguları.....	64
4.1.2. YSA'nın NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması	66
4.1.2.1. Değişen koşullar altında YSA'ya ait NSO'lar için faktöriyel ANOVA bulguları.....	68
4.1.2.2. Değişen koşullar altında YSA'ya ait ÖSO'lar için faktöriyel ANOVA bulguları.....	71
4.1.3. DINO Modelinin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması	73
4.1.3.1. Değişen koşullar altında DINO modeline ait NSO'lar için faktöriyel ANOVA bulguları	75
4.1.3.2. Değişen koşullar altında DINO modeline ait ÖSO'lar için faktöriyel ANOVA bulguları	77
4.2. Simülasyon Verisinde Nitelik Sayısı (3, 5 ve 7), Örneklem Büyüklüğü (30, 100 ve 500), Madde Sayısı (15, 30 ve 45) ve Kayıp Veri Oranı (0; 0,05 ve 0,10) Koşullarında POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması	80
4.2.1. Nitelik Sayılarına (3, 5 ve 7) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması.....	80

4.2.2. Örneklem Büyüklüklerine (30, 100 ve 500) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması	84
4.2.3. Madde Sayılarına (15, 30 ve 45) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması.....	88
4.2.4. Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması.....	92
4.3. Gerçek Veri Setinde (PISA 2015 İPÇ) POBT'nin, YSA'nın ve DINO Modeli Sınıflandırma Oranlarının Benzerliğinin Nitelik Sayılarına (3, 7 ve 11) ve Örneklem Büyüklüklerine (30, 100 ve 500) Göre Durumları	96
4.3.1. Nitelik Sayılarına (3, 7 ve 11) Göre POBT, YSA ve DINO Modeline ait Sınıflandırma Oranlarının Benzerliği	97
4.3.2. Örneklem Büyüklüklerine (30, 100 ve 500) Göre POBT, YSA ve DINO Modeline ait Sınıflandırma Oranlarının Benzerliği.....	100
BÖLÜM V	103
SONUÇ, TARTIŞMA VE ÖNERİLER	103
5.1. Sonuç ve Tartışma.....	103
5.1.1. Simülasyon Verisinde POBT'nin, YSA'nın ve DINO Modelinin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) göre Durumları	108
5.1.1.1. POBT'ye ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) göre Durumları	108
5.1.1.2. YSA'ya ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Durumları	110
5.1.1.3. DINO modeline ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Durumları	112

5.1.2. Simülasyon Verisinde POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması	113
5.1.3. Gerçek Veri Setinde (PISA 2015 İPÇ) POBT'nin, YSA'nın ve DINO Modelinin Sınıflandırma Oranlarının Benzerliğinin Nitelik Sayılarına (3, 7 ve 11) ve Örneklem Büyüklüklerine (30, 100 ve 500) Göre Durumları	115
5.2. Öneriler	117
5.2.1. Kullanıcılar için öneriler	117
5.2.2. Araştırmacılar için öneriler	117
KAYNAKLAR.....	121
EKLER.....	138
EK-1. Gerçek Veri Setlerine Ait Betimsel Bulgular	139
EK 2. Simülasyon Veri Setlerinin Üretilmesinde Kullanılan Q Matrisleri	142
EK 3. Simülasyon Veri Setlerinin Yapısı ve Koşulların Belirlenmesinde Dikkate Alınan Çalışmalar	143
EK-4. Değişen Koşullar Altında POBT, YSA ve DINO Modelinin NSO ve ÖSO'ları	146
EK-5. Değişen koşullarda POBT'ye ait NSO ve ÖSO'lar	148
EK-6. Değişen koşullarda YSA'ya ait NSO ve ÖSO'lar	154
EK-7. Değişen koşullarda DINO modeline ait NSO ve ÖSO'lar	160
EK 8. Değişen Koşullar Altında POBT ve DINO Modelinin Uyum İndeksleri.....	166

TABLÖLAR LİSTESİ

Tablo 1. Örnek <i>Q</i> Matrisi.....	15
Tablo 2. Farklı <i>Q</i> Matrislerinde Her Bir Niteliği Ölçen Madde Sayısı.....	44
Tablo 3. Araştırma Kapsamında Simülasyon Verisi için Ele Alınan Faktörler ve Bu Faktörlere İlişkin Düzeyler	46
Tablo 4. PISA 2015 İPÇ Beceri Matrisi (OECD, 2017a, s.12; 2017b, s.50)	49
Tablo 5. Gerçek Veri için İlk 20 Maddeye Ait <i>Q</i> Matrisleri	50
Tablo 6. Gerçek Veri Seti için Ele Alınan Faktörler ve Bu Faktörlere İlişkin Düzeyler	50
Tablo 7. Değişen koşullarda POBT'ye ait NSO ve ÖSO'lar	60
Tablo 8. Değişen koşullar altında POBT'ye ait NSO'lar için faktöriyel ANOVA sonuçları	62
Tablo 9. Değişen koşullar altında POBT'ye ait ÖSO'lar için faktöriyel ANOVA sonuçları	65
Tablo 10. Değişen koşullarda YSA'ya ait NSO ve ÖSO'lar	67
Tablo 11. Değişen koşullar altında YSA'ya ait NSO'lar için faktöriyel ANOVA sonuçları	69
Tablo 12. Değişen koşullar altında YSA'ya ait ÖSO'lar için faktöriyel ANOVA sonuçları	72
Tablo 13. Değişen koşullarda DINO modeline ait NSO ve ÖSO'lar	74
Tablo 14. Değişen koşullar altında DINO modeline ait NSO'lar için faktöriyel ANOVA sonuçları	76
Tablo 15. Değişen koşullar altında DINO modeline ait ÖSO'lar için faktöriyel ANOVA sonuçları	78

Tablo 16. Nitelik Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları	81
Tablo 17. Nitelik Sayısı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin NSO'larına Ait Basit Etki ANOVA Sonuçları.....	81
Tablo 18. Nitelik Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları	83
Tablo 19. Nitelik Sayısı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına Ait Basit Etki ANOVA Sonuçları	83
Tablo 20. Örneklem Büyüklüğü Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları.....	84
Tablo 21. Örneklem Büyüklüğü Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin NSO'larına Ait Basit Etki ANOVA Sonuçları.....	85
Tablo 22. Örneklem Büyüklüğü Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları.....	86
Tablo 23. Örneklem Büyüklüğü Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına Ait Basit Etki ANOVA Sonuçları.....	87
Tablo 24. Madde Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları	89
Tablo 25. Madde Sayısı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin NSO'larına Ait Basit Etki ANOVA Sonuçları.....	89
Tablo 26. Madde Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları	90
Tablo 27. Madde Sayısı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına Ait Basit Etki ANOVA Sonuçları	91
Tablo 28. Kayıp Veri Oranı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları.....	92
Tablo 29. Kayıp Veri Oranı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin NSO'larına Ait Basit Etki ANOVA Sonuçları.....	93
Tablo 30. Kayıp Veri Oranı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları.....	94

Tablo 31. <i>Kayıp Veri Oranı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına Ait Basit Etki ANOVA Sonuçları</i>	95
Tablo 32. <i>Gerçek Veride POBT, YSA ve DINO Modeline ait NSOB ve ÖSOB'ların Nitelik Sayısına Göre Değişimi</i>	98
Tablo 33. <i>Gerçek Veride POBT, YSA ve DINO modeli NSOB ve ÖSOB'ların Örneklem Büyüklüğüne Göre Değişimi</i>	101



ŞEKİLLER LİSTESİ

<i>Şekil 1.</i> Çok katmanlı algılayıcı YSA örneği	23
<i>Şekil 2.</i> Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında POBT'ye ait ortalama NSO'lar.....	63
<i>Şekil 3.</i> Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında POBT'ye ait ortalama ÖSO'lar.....	66
<i>Şekil 4.</i> Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında YSA'ya ait ortalama NSO'lar	70
<i>Şekil 5.</i> Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında YSA'ya ait ortalama ÖSO'lar	73
<i>Şekil 6.</i> Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında DINO modeline ait ortalama NSO'lar	77
<i>Şekil 7.</i> Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında DINO modeline ait ortalama ÖSO'lar	79
<i>Şekil 8.</i> Nitelik sayısının 3, 5 ve 7 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin NSO'ları.....	82
<i>Şekil 9.</i> Nitelik sayısının 3, 5 ve 7 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları.....	84
<i>Şekil 10.</i> Örneklem büyüklüğünün 30, 100 ve 500 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin NSO'ları.....	86
<i>Şekil 11.</i> Örneklem büyüklüğünün 30, 100 ve 500 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları.....	88

<i>Şekil 12.</i> Madde sayısının 15, 30 ve 45 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin NSO'ları.....	90
<i>Şekil 13.</i> Madde sayısının 15, 30 ve 45 madde olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları.....	92
<i>Şekil 14.</i> Kayıp verinin 0; 0,05 ve 0,10 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin NSO'ları.....	94
<i>Şekil 15.</i> Kayıp veri oranının 0; 0,05 ve 0,10 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları.....	96



SİMGELER VE KISALTMALAR LİSTESİ

2PL	İki Parametrelili Lojistik Model
AAR	Attribute-Wise Agreement Rate
ACR	Attribute-Level Correct Classification Rate
ACR	Attribute-Based Classification Rate
AHM	Attribute Hierarchy Method
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
ARI	Adjusted Rand Index
ATC21	Assesment and Teaching of 21st Century Skills
BIC	Bayesian Information Criterion
BTM	Bilişsel Tanı Modeli
CDM	Cognitive Diagnosis Model
CPS	Collaborative Problem-Solving
C-RUM	Compensatory Reparameterized Unified Model
ÇKA-YSA	Çok Katmanlı Algılayıcı Yapay Sinir Ağı
DINA	Deterministic-Input, Noisy-And Gate
DINA-ACDM	Deterministic-Input, Noisy-And Gate Additive Cognitive Diagnosis Model
DINO	Deterministic-Input, Noisy-Or Gate
DINO-ACDM	Deterministic-Input, Noisy-Or Gate Additive Cognitive Diagnosis Model
EAP	Expected a Posteriori
EM	Expectation–Maximization

ETS	Educational Testing Service
GBTM	Genelleştirilmiş Bilişsel Tanı Modeli
G-DINA	Generalized Deterministic-Input, Noisy-And Gate
GNIDA	Generalized Noisy Inputs, Deterministic “And” Gate
GPOBT	General Nonparametric Classification
GPCM	Generalized Partial Credit Model
HACA	Hierarchical Agglomerative Cluster Analysis
HODINA	Higher-Order Deterministic-Input, Noisy-And Gate
İ-İ	İnsan-İnsan
İ-SK	İnsan-Sanal Katılımcı
İPÇ	İşbirlikli Problem Çözme
JMLE	Joint Maximum Likelihood Expectation
JMLE-ÇKA-YSA	Çok Katmanlı Algılayıcı Yapay Sinir Ağı'nın JMLE ile birleşimi
LLTM	Lineer Lojistik Test Modeli
LSDM	The Least Square Distance Method
MAP	Maximum a Posteriori
MAR	Missing at Random
MCAR	Missing Completely at Random
MCMC	Markov Chain Monte Carlo
MCMC-ÇKA-YSA	Çok Katmanlı Algılayıcı Yapay Sinir Ağı'nın MCMC ile birleşimi
MEB	Milli Eğitim Bakanlığı
MLE	Maximum Likelihood Estimation
MNAR	Missing Not at Random
MNV	Multivariate Normal
MTK	Madde Tepki Kuramı
NIDA	Noisy Inputs, Deterministic “And” Gate
POBT	Parametri Olmayan Bilişsel Tanılama
NPCD	Nonparametric Cognitive Diagnosis

NSO	Nitelik Düzeyi Sınıflandırma Oranı
NSOB	Nitelik Düzeyi Sınıflandırma Oranlarının Benzerliği
OECD	Organisation for Economic Co-operation and Development
ÖSO	Örüntü Düzeyi Sınıflandırma Oranını
ÖSOB	Örüntü Düzeyi Sınıflandırma Oranlarının Benzerliği
Pa	Classification Accuracy
PAR	Pattern-Wise Agreement Rate
Pc	Classification Consistency
PCM	Partial Credit Model
PCR	Pattern-Based Classification Rate
PIAAC	Programme for the International Assessment of Adult Competencies
PISA	Programme for International Student Assessment
RED-RUM	Reduced Reparameterized Unified Model
RMSEA	Root Mean Square Error of Approximation
RSM	Rule-Space Methodology
RUM	Reparameterized Unified Model
SOM	Self-Organised Map
TIMSS	Trends in International Mathematics and Science Study
VCR	Vector-Level Correct Classification Rate
YSA	Yapay Sinir Ağı

BÖLÜM I

GİRİŞ

Bu bölümde araştırmanın problem durumu, amacı, önemi, sınırlılıkları ve tanımlar bölümlerine yer verilmiştir.

1.1. Problem Durumu

Bilim, teknoloji ve dünyadaki diğer gelişmelere paralel olarak ülkelerin ihtiyaç duydukları vatandaş profillerinin değişmesi, eğitimde ve eğitimin ayrılmaz bir parçası olan ölçme ve değerlendirmede birtakım değişikliklerin meydana gelmesine neden olmuştur. Örneğin, dünya üzerinde uzak mesafe kavramı kalkmış, coğrafi olarak birbirlerinden çok uzak ülkelerde bulunan kişiler gelişen iletişim ağları aracılığıyla her gün birlikte çalışır hale gelmişlerdir. Özellikle Covid-19 salgını nedeniyle pek çok çalışan, ekip arkadaşlarıyla işlerini internet platformu üzerinden beraber çalışarak sürdürmek durumunda kalmışlardır. Meydana gelen bu değişimler, bireylerin etkili iletişim kurma, olaylara farklı bir perspektiften bakma, eleştirel düşünme ve problem çözme gibi karmaşık becerilere sahip olmalarını gerekli kılmıştır. 21. yüzyıl becerileri olarak adlandırılan bu beceriler, çeşitli uluslararası kuruluşlarca (bkz. Amerikan Yönetim Derneği, 2010; Organisation for Economic Co-operation and Development [OECD], 2013) ele alınmış, tanımları ve ölçme ve değerlendirme yöntemleri üzerine araştırmalar yapılmıştır. Bu araştırmalar sonucu bu beceriler, ülkelerin eğitim programlarında yer almaya, uluslararası öğrenci değerlendirme programı (PISA, Programme for International Student Assessment) ve Uluslararası matematik ve fen eğilimleri çalışması (TIMSS, Trends in International Mathematics and Science Study) gibi uluslararası sınavlarda ölçülmeye ve değerlendirilmeye başlanmıştır.

Yirmibirinci yüzyıl becerileri, bilginin yanı sıra karmaşık bilişsel süreçlerini de beraberinde getirmektedir. Bu durum, yaklaşık 30 yıl öncesine kadar test tasarımı üzerinde çok az etkisi olan bilişsel teorinin (Embretson, 1998) giderek önem kazanmasına sebep olmuştur. Bununla beraber, 21. yüzyıl becerilerinin karmaşık yapıları eğitim ve öğretimde kullanılan ölçme ve değerlendirme yöntemleri üzerinde de etkisini göstermiştir. Önceleri basit yapılar görece sade ve basit analizlerle değerlendirilirken, sonraları karmaşık yapılar karmaşık analizlerle değerlendirilmeye başlanmıştır. Buna paralel olarak, Amerika'daki Hiçbir Çocuk Geri Kalmasın (No Child Left Behind, 2001) kanunu ile birlikte öğrencilerin neyi ne kadar öğrendikleri konusunda bilişsel tanı raporları hazırlanmaya başlanmıştır (Kato, 2009). Bu raporların hazırlanması değerlendirmelerin geleneksel yani düzey belirleyici (summative) olmaktan çıkarak daha çok tanılayıcı (diagnostic) ve biçimlendirici (formative) olarak kullanılmasında ve yaygınlık kazanmasında etkili olmuştur. Tanılayıcı ve biçimlendirici değerlendirmelerin yaygınlık kazanması ve 21. yüzyıl becerilerinin ölçülmeleri ve değerlendirilmeleri öğrencilere donut vermeyi amaçlayan bilişsel modellerin gelişimini hızlandırmıştır.

Geleneksel değerlendirmelerde öğrencilerin öğrenmeleri hakkında tek bir puan elde edilmekte ve bu puana dayanarak yargılara ulaşılmaktadır. Tanılayıcı ve biçimlendirici değerlendirmelerde ise önce bir işin başarılı bir şekilde tamamlanması için gerekli bilişsel süreçler incelenmekte, sonra öğrencilerin tek bir puanı yerine bilişsel süreci oluşturan ve öğrencilerin uzmanlaşmalarını gerektiren her bir beceri hakkında detaylı bilgiler sağlanmaktadır (McGlohen, 2004). Böylece öğrencilerin neyi, ne kadar ve nasıl öğrendikleri ve belirlenen hedeflere ulaşmak için hangi becerilere ihtiyaç duydukları belirlenmektedir. Öte yandan tanılayıcı ve biçimlendirici değerlendirmeler, öğretmenlerin kullandıkları öğretim yöntemlerinin etkililiğini detaylı olarak değerlendirebilmelerine ve öğretim yöntemlerinde gerekli değişiklikleri yapmalarına da olanak sunmaktadır. Buna ek olarak öğretmenler, sınıflarında bulunan öğrenci profillerine göre öğretim etkinliklerini düzenleyebilmekte ve öğrencilere güçlü ve zayıf becerileri hakkında detaylı dönütler verebilmektedirler. Böylece öğrenciler, eksik oldukları becerilerine yönelik çalışmalar yaparak bilgi ve becerilerini kısa sürede geliştirebilmektedirler. Dolayısıyla öğrenciler, kendi öğrenmeleri üzerinde söz sahibi olup eğitim hayatlarında daha aktif rol alabilmektedirler. Black ve Wiliam (1998) çalışmalarında, etkili bir tanılayıcı ve biçimlendirici değerlendirmenin öğrenci öğrenmelerinin artmasını sağladığını ve yüksek öğrenci performansını teşvik ettiğini ifade etmişlerdir. Bunun nedenlerinden biri olarak

tanılayıcı ve biçimlendirici değerlendirmelerde kullanılan ölçme araçları düşünülebilir. Çünkü düzey belirleyici değerlendirmelerde kullanılan testlerin geçerlik ve güvenirlikleri ne kadar yüksek olsa da bu testlerle öğrencilerin güçlü ve zayıf yönleri ile zihinsel süreçleri hakkında detaylı bilgilere ulaşılamamaktadır. Bu nedenle test maddeleri ile bilişsel süreçlerin birleştirilme ihtiyacı ortaya çıkmıştır. Bu ihtiyacın karşılanması amacıyla da tanılayıcı ve biçimlendirici değerlendirme kapsamında bilişsel modeller kullanılmaya başlanmıştır (Leighton, Gierl, & Hunka, 2004). Öğrenciler geleneksel değerlendirmede yapının geneline ait tek bir puan ile başarılı-başarısız şeklinde sınıflandırılırken, bilişsel modellerde yapıyı oluşturan her bir niteliğe (beceriye) ait farklı puanlarla başarılı-başarısız şeklinde sınıflandırılırlar. Böylece öğrencilerin güçlü ve zayıf yönleri, yapıyı oluşturan niteliklere sahip olma düzeyleriyle belirlenir (Rupp & Templin, 2008b). Bu tür değerlendirmeler alanyazında bilişsel tanı değerlendirmesi (cognitive diagnostic assessment) olarak yer almaktadır. Bilişsel tanı değerlendirmesinde (bilişsel tanılamada) öğrenci performansı öncelikle parametrik ve yarı-parametrik modeller ile değerlendirilmeye başlanmıştır ve bu kapsamdaki alanyazın gelişmeye devam etmektedir. Daha sonra parametrik modellerin bazı özellikleriyle ortak paydaları olan parametrik olmayan model ve tekniklerin bilişsel tanılamada kullanıldığı görülmektedir. Bu nedenle mevcut araştırma kapsamında öncelikle parametrik modeller üzerinde durulmuş daha sonra parametrik olmayan modellerden bahsedilmiştir.

Bilişsel tanılamada, yapıyı oluşturan nitelikler ile öğrencilerin gözlenen puanları arasındaki ilişkiyi kurmak için farklı istatistiksel modeller kurulabilmektedir (Cui, Gierl, & Guo, 2016; de la Torre & Douglas, 2004; Gierl, Cui, & Hunka, 2007; Junker & Sijtsma, 2001). Bu modeller, tanısal sınıflandırma modelleri (diagnostic classification models) veya bilişsel tanı modelleri (BTM'ler; cognitive diagnosis models) çatısı altında toplanmaktadır. Ayrıca BTM'leri öğrencilerin niteliklere sahip oluş düzeylerini test eden olasılıksal, doğrulayıcı ve çok boyutlu sınırlandırılmış gizil sınıf modelleri grubu olarak da tanımlamak mümkündür (Rupp, Templin, & Henson, 2010).

BTM'lerde öğrenci performansı değerlendirilirken, kullanılan testin her bir maddesi en az bir nitelik ile ilişkilendirilir. Bu, maddenin doğru cevaplanması halinde öğrencinin o niteliğe sahip olduğu; yanlış cevaplanması halinde ise o niteliğe sahip olmadığı şeklinde yorumlanmaktadır. Böylece bilişsel tanılamada öğrencilerin güçlü veya zayıf oldukları nitelikler kolayca belirlenebilmekte (Junker & Sijtsma, 2001; de la Torre & Douglas, 2004; Henson & Templin, 2007; von Davier & Yamamoto, 2008; Rupp vd., 2010) ve öğrencilere

detaylı dönütler verilebilmektedir. Ayrıca BTM’lerde bir madde birden fazla nitelik ile ilişkilendirilebilmektedir. Böylece eşit sayıda maddeye sahip testlerden elde edilen bilgiler kıyaslandığında bilişsel tanılamayla elde edilen bilgiler, geleneksel değerlendirmelerden daha fazla olabilmektedir. BTM’ler bu yönüyle de testlerin kullanılabilirliklerinin artmasına yardımcı olmaktadır. BTM’lerde madde parametre kestirimi için yaygın olarak beklenti maksimizasyonu (Expectation–maximization: EM) veya Markov Zinciri Monte Carlo (Markov Chain Monte Carlo: MCMC) prosedürleri kullanılmaktadır. Ancak bu prosedürlerin parametre kestirimlerinin ve nitelik sınıflandırmalarının doğru olması ve iyi model uyum indekslerinin elde edilmesi için genellikle büyük örneklemelere ihtiyaç duyulmaktadır (Chiu, Gird, & Guo, 2016, Shu, Henson, & Willse, 2013). BTM’lerin büyük örneklem ihtiyacı onların uygulanmalarını sınırlandırırken (Shuying, 2016) parametrik olmayan bilişsel tanılamanın ortaya çıkmasına sebep olmuştur (bkz. Chiu, & Douglas, 2013; Shu, vd. 2013).

BTM’lerde izlenen ilk adım geçerli ve güvenilir bir Q matrisinin belirlenmesi ve uygun bir BTM’nin seçilmesidir. BTM’lerde yapıyı oluşturan niteliklerin kendi aralarındaki ilişkileri ile bu niteliklerin maddelerle olan ilişkileri Q matrisleri aracılığıyla kurulmaktadır. Bu nedenle Q matrisinin doğru bir şekilde tanımlanması BTM’ler için oldukça önemlidir (de la Torre & Chiu, 2015). Q matrislerinde nitelikler arasında önkoşul ilişkileri sıklıkla görülmektedir. Önkoşul ilişkilerin bulunduğu durumlarda, temel niteliklere sahip olma daha karmaşık niteliklere ulaşmak için gereklidir (de la Torre, Hong, & Deng, 2010; Leighton vd., 2004; Templin & Bradshaw, 2014). Nitelikler bu tür bir hiyerarşik yapıya sahip olduğunda Q matrisleri, nitelikler arasındaki hiyerarşi göz önüne alınarak tasarlanmaktadır. Q matrisinin geçerli ve güvenilir bir şekilde belirlenmemesi, model-veri uyumsuzluğuna ve öğrencilerin sınıflandırmalarında yanlılığa yol açabilmektedir (Kunina-Habenicht, Rupp, & Wilhelm, 2012; Rupp & Templin, 2008a, 2008b). Bu da öğrenciler hakkında yanlış değerlendirmelerin yapılmasına neden olabilmektedir. İdeal olarak, nihai Q matrisinin belirlenmesi için hem alanyazındaki teorik bilgiler hem de deneysel çalışmalara dayalı madde cevapları dikkate alınmalıdır (Liu, Huggins-Manley, & Bulut, 2018). Alanyazında Q matrisinin belirlenmesi için öznel (uzman kanıları) ve nesnel (çeşitli istatistiksel yöntemler) yöntemler kullanılmaktadır (bkz. Ma, 2017; Templin & Bradshaw, 2014).

BTM’de izlenen ilk adımda belirlenen Q matrisinden yararlanılarak uygun bir BTM seçilir. BTM’nin seçilmesi öğrencilerin doğru nitelik sınıflarına yerleştirilmesinde önemli bir yere

sahiptir (de la Torre & Minchen, 2014). Ölçülen nitelikler ile öğrencilerin cevapları arasındaki uyumu açıklamak için çok sayıda BTM geliştirilmiş ve çeşitli şekillerde sınıflandırılmıştır. Bu sınıflandırmalarda modellerin genel olarak telafisel ve telafisel olmayan şeklinde ayrıldıkları görülmektedir. Modellerin bu şekilde ayrılması nitelikler arasındaki ilişkiye bağlıdır. Telafisel olmayan modellerde, bir maddenin doğru cevaplanması için gerekli olan tüm niteliklere sahip olunması gerekir. Telafisel modellerde ise maddenin doğru cevaplanması için gerekli niteliklerden sadece birine sahip olunması yeterlidir. Bununla birlikte, telafisel ve telafisel olmayan modellerin seçiminde, niteliklerin (gizil yordayıcı değişkenler) ve maddelerin puanlanma şekilleri (ikili veya çoklu) de dikkate alınmaktadır. Sonuç olarak, bir araştırmada kullanılacak BTM gerek Q matrislerinin yapısı gerekse nitelikler arasındaki ilişkiler dikkate alınarak seçilmelidir, aksi halde seçilen modeller uygun ve kullanışlı olmayabilir (Templin & Bradshaw, 2014).

Q matrisinin tasarlanması ve uygun bir BTM'nin seçilmesinin ardından, bilişsel tanılamadaki ikinci adım, öğrencilerin nitelik profillerine göre sınıflandırılması ve analiz sonuçlarının değerlendirilmesidir. BTM'lerde sınıflandırma, öğrencilerin incelenen yapıyı oluşturan niteliklere sahip olma derecelerini ortaya koymaktadır. Bu doğrultuda doğru bir sınıflandırma, öğrenciler hakkındaki doğru tanısal bilgileri elde edebilmek için oldukça önemlidir.

Bilişsel tanı sınıflandırması alanyazınında öğrencileri nitelik profillerine göre sınıflandırmak için yapılan çalışmaların iki gruba ayrıldığı görülmektedir. Birinci gruptaki çalışmalar (bkz. Chen & de la Torre, 2013; Liu vd., 2018; Liu, Qian, Luo, & Woo, 2018; Templin & Bradshaw, 2014), güçlendirilmiş BTM'leri (retrofitting) incelemektedir. Güçlendirilmiş modeller, BTM'lere uygun olarak geliştirilmemiş testlerden elde edilen bilgilerin en fazla olmasını sağlamaktadır. Bu çalışmalar, yeterli sayıda ve çeşitlilikte tanı testleri geliştirilememe sorunu nedeniyle ortaya çıkmıştır. BTM'lere uygun testlerin tasarlanması, üretilmesi ve analiz edilmesi için önemli miktarda zaman ve kaynak gerekmektedir. Bu nedenle bilişsel tanı testlerindeki gelişmeler, analizleri için gerekli bilişsel tanı modellerindeki gelişmelerden daha yavaştır (de la Torre & Minchen, 2014). Bilişsel tanı modelleri ile tanı testlerinin geliştirilme hızındaki dengesizliğin üstesinden gelmek için BTM'leri uyarılama/güçlendirme çalışmaları başlamıştır (Chen & de la Torre, 2013). Mevcut testler için BTM'lerin güçlendirilmesi, var olan testlerden elde edilen tanısal bilgileri artırsa da (de la Torre & Minchen, 2014) değerlendirme tasarımının veya istatistiksel kalitenin bir sonucu olarak ortaya çıkabilecek bazı zorlukları da beraberinde

getirmektedir (Liu vd., 2018). Bu zorluklardan bazıları; (a) tek boyutlu değerlendirmelerden çok boyutlu niteliklerin tanımlanması, (b) niteliklerin yüksek düzeyde ilişkilendirilmiş olması, (c) istatistiksel kalite düşüklüğü ve (d) zayıf model uyumudur. Bununla beraber, Liu vd. (2018) çalışmalarında belirli koşullarda uyarlama/güçlendirme modellerinin kullanılmasını tavsiye etmektedirler. Bu koşullar; (a) tanısal testlerin geliştirilmesinin mümkün olmaması, (b) öğrenmeyi destekleyen ya da bilişsel teori hakkında tanılayıcı ve biçimlendirici bilgiye ulaşmanın amaçlanması, (c) maddelerin geliştirilmesi sırasında cevaplanmaları için gereken niteliklerin açıkça belirtilmiş ve boyutsal değerlendirme yoluyla desteklenmiş olmasıdır.

İkinci gruptaki çalışmalar (bkz. Ayers, Nugent, & Dean, 2008; Gierl, Wang, & Zhou, 2008; Willse, Henson, & Templin, 2007) ise BTM'nin uygun olmadığı durumlarda uygulanabilecek alternatif model ve teknikleri incelemektedir. Bu model ve tekniklere örnek olarak Kural-uzayı Metodolojisi verilebilir (KUM; Rule-space Methodology). KUM, MTK'dan (Madde tepki kuramı) parametrik tek boyutlu bir model ile hata temelli kişi uyum indeksini birleştirerek yarı parametrik bir bakış açısı sunmaktadır (Tatsuoka, 2009). Benzer şekilde, Nitelik Hiyerarşi Yöntemi (Attribute Hierarchy Method; AHM), maksimum olasılık kestirim (Leighton, Gierl, & Hunka, 2004) ve sinir ağı (Gierl vd., 2007) sınıflandırma yöntemlerini kullanmaktadır. Bu yöntemlerin her ikisi de matematiksel küme teorisi tarafından yönlendirilen ve bir sınıflandırma yaklaşımı olan Bilgi-uzay Teorisine (Knowledge-space Theory) dayanmaktadır (Schrepp, 2005). Ayrıca parametrik olmayan yaklaşımlar da (parametrik olmayan bilişsel tanılama, yapay sinir ağı, k-ortalama ve hiyerarşik kümeleme teknikleri, vb.) uzaklık ölçütleri kullanarak bilişsel tanılama kapsamında sınıflandırma amaçlı kullanılmaktadır (Rupp, 2013).

BTM'lerin kestirimlerde kullandıkları tahmin algoritmalarının ilk zamanlar kamuya açık olmaması (Chiu, Sun, & Bian, 2017), genellikle büyük örneklem gerektirmeleri ve BTM uygulamasının genişliğini sınırlayan (daha maliyetli) özel yazılımlarla ağır hesaplama prosedürlerini içermeleri (Chiu & Douglas, 2013; Wang & Douglas, 2015; Chiu & Köhn, 2019) araştırmacıları bilişsel tanılamada kullanılabilecek parametrik olmayan yöntem arayışına yöneltmiştir. Bu arayış bilişsel tanılama alanında yapılan ikinci grup araştırmaların temelini oluşturmaktadır. Bu çalışmaların odak noktası ise bilişsel tanı sınıflandırmasında kullanılabilecek alternatif model ve tekniklerin çeşitli koşullar altında sınıflandırma doğruluklarının karşılaştırılmasıdır.

Sınıf veya kurs düzeyi öğretme ve öğrenme sürecini izlemek için sahip olunan örneklem büyüklüğü, bazen BTM kestirimlerinde kullanılan algoritmaların güvenilir sonuçlar vermesi için yeterli olmamaktadır (bkz. Akbay, 2016). Ayrıca bu tür değerlendirmelerde odak genellikle madde parametrelerinin kestirilmesinden ziyade öğretimin ve öğrenci öğrenmelerinin değerlendirilmesinde olmaktadır (Chiu & Köhn, 2019). Bu nedenle madde parametrelerinin kestirilmesinden çok öğrencilerin nitelik sınıflarına yerleştirilmesine odaklanan parametrik olmayan tekniklerin, sınıf düzeyi tanılayıcı ve biçimlendirici değerlendirmelerde güvenilir ve kullanışlı bir yöntem olabileceği düşünülmektedir. Küçük örneklem sorunu nedeniyle benzer durumlar, bilgisayar ortamında bilişsel tanı temelli bireye uyarlanmış testler (cognitive diagnosis based computerized adaptive tests) için de geçerlidir. Sınava girenlerin nitelik sınıflarına sahip oluş düzeylerinin güvenilir bir şekilde değerlendirilmesi için etkili bir hesaplama yönteminin bulunamaması, şimdiye kadar sınıflarda bilgisayar ortamında bilişsel tanı temelli bireye uyarlanmış testlerin kullanımını önleyen ciddi bir engel olmuştur (Chiu & Köhn, 2019). Parametrik olmayan bilişsel tanılama ile birlikte bu alanda yapılan çalışmalar da hız kazanmıştır. Küçük örneklem sorununa ek olarak BTM'ler, az madde veya fazla niteliğin olduğu bazı koşullarda öğrenciler hakkında yanlış sonuçlar verebilmektedir (Shuying, 2016; Shu, Henson, & Willse, 2013). Tanılayıcı ve biçimlendirici değerlendirmelerin önem kazandığı günümüzde 21. yüzyıl becerilerinin karmaşık yapıya sahip olmaları (fazla nitelik sayısı içermeleri) da parametrik olmayan bilişsel tanılama tekniklerinin geliştirilmesine olan ihtiyacı ortaya çıkarmıştır. Bu bağlamda mevcut araştırma, bilişsel tanı sınıflandırmasında kullanılan parametrik olmayan bilişsel tanılama (POBT) ve yapay sinir ağlarına (YSA) odaklanmaktadır. BTM'lerle analiz için veri yapısının (telafisel veya telafisel olmayan) ve Q matrisinin bilinmesi hususunda benzerlik gösteren bu yöntemler, öğrencileri nitelik profillerine sınıflandırmada farklı bakış açıları kullanmaları yönüyle birbirinden farklıdır. POBT, öğrencilerin ideal cevap örüntüleri ile gözlenen cevap örüntülerini Hamming uzaklığıyla karşılaştırarak öğrencileri nitelik sınıflarına yerleştirirken, YSA ideal cevap örüntüleriyle eğitildikten sonra öğrencileri sınıflandırmaktadır.

Alanyazında YSA (bkz. Cui vd., 2016; Guo, Cutumisu, & Cui, 2017; Shu vd. 2013) ve POBT (bkz. Akbay, 2016; Chiu & Douglas, 2013; Lim & Drasgow, 2017; Wang & Douglas, 2015) ile ilgili yapılan araştırmalarda yöntemlerin hangi koşullarda daha iyi sınıflandırma yaptıklarının belirlenmesi için daha çok araştırmaya ihtiyaç olduğu belirtilmiştir. Bununla beraber parametrik olmayan tekniklerin hangi şartlarda etkili

olabileceğinin belirlenmesi de kritik bir öneme sahiptir (Chiu & Douglas, 2013). Bu nedenle geliştirilen alternatif model ve tekniklerin performanslarını birbirleriyle ve BTM'ler ile çeşitli koşullarda ve veriler üzerinde karşılaştıran çalışmalara ihtiyaç duyulmasından hareketle mevcut çalışma gerçekleştirilmiştir.

YSA ve POBT ile ilgili çalışmalar (bkz. Akbay, 2016; Chiu & Douglas, 2013; Cui vd., 2016; Guo vd., 2017) incelendiğinde araştırmacıların parametrik olmayan yöntem kullandıkları, fakat parametrik olmayan yöntemlerin küçük veri gruplarının analizlerinde sahip oldukları avantajı test etmedikleri görülmektedir. Benzer şekilde bu çalışmalarda en fazla sekiz niteliğin sınıflandırmaya olan etkisinin (bkz. Shu vd., 2013) incelendiği kayıp veri etkisinin ise hiç incelenmediği görülmektedir. Ek olarak alanyazında çeşitli koşullarda YSA ve POBT'lerin sınıflandırma oranlarının birlikte incelendiği çalışmalar (bkz. McCoy & Willse, 2014; Paulsen, 2019) da bulunmaktadır. Bu çalışmalardaki veriler Deterministic-Input, Noisy-And Gate (DINA) modeli temel alınarak üretildiğinden çalışmalardan elde edilen bulgular telafisel olmayan veri yapılarına genellenebilmektedir. McCoy ve Willse (2014) ve Paulsen'nin (2019) elde ettikleri bulguların gerçek ve telafisel modeller temel alınarak üretilen simülasyon veri setlerinden elde edilen bulgularla karşılaştırılması, parametrik olmayan bilişsel tanılamada çeşitli faktörlerin sınıflandırmaya olan etkileri üzerinde bütüncül bir bakış açısı sağlayacağı düşünülmektedir.

Bu düşüncelerden hareketle, mevcut araştırmada POBT ve YSA sınıflandırma oranları öncelikle simülasyon veri setleri üzerinde birbirleriyle ve alanyazında sıkça kullanılan Deterministic-Input, Noisy-Or Gate modeli (DINO, Templin & Bradshaw, 2014) ile karşılaştırılmıştır. Alanyazından hareketle karşılaştırma faktörleri nitelik sayısı, örneklem büyüklüğü, madde sayısı ve kayıp veri oranı olarak belirlenmiştir. Ardından POBT, YSA ve DINO modelinin sınıflandırma oranlarının benzerlikleri PISA 2015 İşbirlikli Problem Çözme (İPÇ; Collaborative Problem Solving) veri seti üzerinde incelenmiştir.

1.2. Araştırmanın Amacı

Bu araştırmada, parametrik olmayan bilişsel tanılama ve yapay sinir ağına ait nitelik ve örüntü düzeyi sınıflandırma oranlarının çeşitli koşullarda DINO temelli simülasyon veri setlerinde öncelikle birbirleriyle, sonra DINO model ile karşılaştırılmaları, ardından parametrik olmayan bilişsel tanılama, yapay sinir ağı ve DINO modelinin sınıflandırma

oranlarının benzerliklerinin PISA 2015 işbirlikli problem çözme veri seti üzerinde incelenmesi amaçlanmaktadır. Bu amaç doğrultusunda aşağıdaki sorulara cevap aranmıştır:

1. Simülasyon verisinde nitelik sayısına (3, 5 ve 7), örneklem büyüklüğüne (30, 100 ve 500), madde sayısına (15, 30 ve 45) ve kayıp veri oranına (0; 0,05 ve 0,10) göre,

1.1. Parametrik olmayan bilişsel tanılamanın nitelik ve örüntü düzeyi sınıflandırma oranları,

1.2. Yapay sinir ağının nitelik ve örüntü düzeyi sınıflandırma oranları,

1.3. DINO modelinin nitelik ve örüntü düzeyi sınıflandırma oranları farklılaşmakta mıdır?

2. Simülasyon verisinde parametrik olmayan bilişsel tanılamanın, yapay sinir ağının ve DINO modelinin nitelik ve örüntü düzeyi sınıflandırma oranları,

2.1. nitelik sayılarına (3, 5 ve 7),

2.2. örneklem büyüklüklerine (30, 100 ve 500),

2.3. madde sayılarına (15, 30 ve 45) ve

2.4. kayıp veri oranlarına (0; 0,05 ve 0,10) göre farklılaşmakta mıdır?

3. Gerçek veride (PISA 2015 İşbirlikli Problem Çözme) parametrik olmayan bilişsel tanılamanın, yapay sinir ağının ve DINO modeli sınıflandırma oranlarının benzerliği,

3.1. nitelik sayılarına (3, 7 ve 11) ve

3.2. örneklem büyüklüklerine (30, 100 ve 500) göre nasıldır?

1.3. Araştırmanın Önemi

Ülkelerin eğitim-öğretim programlarındaki hedef ve kazanımlara ulaşma düzeyleri, ülkelerin coğrafî, siyasi vb. faktörlerine göre farklılaşabilmektedir. Bu farklılıklar, öğrencilerin öğrenme düzeylerinin belirlenmesini ve öğretmenlerin eksiklerini görüp kendilerini geliştirmelerini ve bu doğrultuda öğretim etkinliklerini düzenlemelerini sağlamak amacıyla yapılan tanılayıcı ve biçimlendirici değerlendirmelerle daha net ortaya konulabilmektedir. Nitekim 2023 Eğitim Vizyonu'nun belirlenmesiyle birlikte Türk Eğitim Sistemi'nde ölçme ve değerlendirmede süreç ve sonuç odaklı bütünsel bir anlayış ortaya konulmasına karar verilmiştir (Millî Eğitim Bakanlığı [MEB], 2018). Bu kararla birlikte

öğrencilere, öğretmenlere ve okul yöneticilerine 21. yüzyıl becerilerinin kazandırılması, değerlendirilmesi ve izlenmesi de hedeflenmektedir. Öğrencilerin akademik çıktılarını değerlendirmek üzere Öğrenci Başarı İzleme Araştırması'nın yapılması planlanırken, öğretmen ve okul yöneticilerine çeşitli yeterliliklere dayalı yazılı sınavların uygulaması planlanmaktadır. Bu hedeflerin gerçekleştirilmesi adına Türkiye'nin her ilinde ölçme ve değerlendirme birimleri kurulmuştur (MEB, 2018).

Yapılacak ölçme ve değerlendirme etkinliklerinin amacına en iyi şekilde hizmet edecek tanılayıcı ve biçimlendirici değerlendirme yöntemlerinin seçilmesi önemli olduğundan var olan yöntemlerin özelliklerinin farklı koşullarda birbirleriyle karşılaştırılıp hangi yöntemin hangi koşullar altında daha iyi performans gösterdiğinin belirlenmesini amaçlayan çalışmalara ihtiyaç vardır (Cui vd., 2016). Bu düşünceden hareketle bu çalışmada POBT ve YSA'lar çeşitli koşullarda karşılaştırılmış ve maksimum tanısal bilginin elde edilmesinde hangi yöntemin daha uygun olduğunun belirlenmesi amaçlanmıştır. Bu araştırma sonucunda elde edilecek detaylı bulgular ile öğretmenlerden politika yapıcılara kadar uzanan eğitim sistemindeki paydaşlar, daha geçerli ve güvenilir bir tanısal değerlendirme yapabilmek için amaçlarına en uygun yöntemi seçebileceklerdir. Bu kapsamda Türkiye'de her ilde kurulan ölçme ve değerlendirme birimleri, izleme araştırmalarının değerlendirilmesinde kendi verilerine özgü en uygun yöntemi belirleyebilecek ve kullanabileceklerdir. Ayrıca küçük örneklerde çalışacak araştırmacıların, mevcut araştırma bulgularıyla çalışmalarında kullanabilecekleri yöntemler hakkında daha detaylı bilgilere ulaşabilecekleri düşünülmektedir. Bu nedenle bu çalışmanın ileride yapılacak araştırmalara yol göstermesi bakımından önemli olduğu düşünülmektedir.

Teknolojideki gelişmelerin eğitimle birleştirilmesi ile birlikte ölçme ve değerlendirme etkinlikleri ve yöntemleri de değişime uğramıştır. Bu duruma, teknolojide yapay zekânın geliştirilmesiyle birlikte internet ortamında interaktif öğrenme ortamları ve etkileşimin odak tutulduğu değerlendirme yöntemlerinin uygulanmaya başlanması örnek olarak verilebilir. Yapay zekâ; uzman sistemler, bulanık mantık, genetik algoritmalar ve YSA gibi tekniklerden oluşmaktadır (Matzavela, Chrysafiadi & Alepis, 2017). Yapay zekânın tekniklerinden olan YSA'nın AHM dışında küçük örneklerde bilişsel tanılama için kullanılabileceği 2013 yılında Shu vd. tarafından önerilmiştir. Fakat YSA'nın tanısal bilgi elde etme performansını diğer modellerin performansları ile karşılaştıran çok az çalışma bulunmaktadır (bkz. Cui vd., 2016; Guo vd., 2017; McCoy & Willse, 2014; Paulsen, 2019;

Shu vd., 2013). Bu araştırma, YSA'ların gerek tanılayıcı değerlendirmelerde kullanılmasının yaygınlaştırılması gerekse farklı koşullarda YSA performanslarının karşılaştırılması açısından önem arz etmektedir.

Alanyazında mevcut araştırma konusuna benzer iki çalışma bulunsa da (bkz. McCoy & Willse, 2014; Paulsen, 2019) mevcut araştırma, bu çalışmalardan bazı farklı yönlerle sahiptir. Mevcut araştırmada hem gerçek hem de DINO temelli simülasyon veri setlerinin kullanılması ve veri yapısına bağlı olarak telafisel modellerin sınıflandırma performanslarının incelenmesi bu farklara bir örnektir. Ayrıca McCoy ve Willse (2014) ve Paulsen'in (2019) çalışmalarında en fazla nitelik sayısı sekiz iken mevcut araştırmanın gerçek veri analizi koşullarında nitelik sayısı 11'dir. McCoy ve Willse (2014) ve Paulsen'in (2019) çalışmalarında madde ayırt edicilikleri yüksek veya düşük olarak ele alınıp incelenirken mevcut araştırmada önceki çalışmaların önerileri üzerine madde ayırt edicilikleri karışık (orta) düzeyde ele alınmıştır. Ayrıca mevcut araştırmada POBT, YSA ve DINO modellerinin performansları bir arada incelenerek bahsedilen iki çalışma sonucuna katkı sağlamak hedeflenmiştir.

Son olarak bu çalışma, öğrencilerin 21. yüzyıl becerilerinin değerlendirilmesinde parametrik olmayan bilişsel tanılama ve YSA'nın kullanılması ve bu analizlerin sınıflandırma performanslarının karşılaştırılması konusunda bu alanda yapılacak çalışmalara bir temel oluşturması nedeniyle önemli görülmektedir.

1.4. Araştırmanın Sınırlılıkları

Mevcut araştırmada kullanılan gerçek veri setiyle (PISA 2015 İPÇ) ilgili teknik raporlarda her madde paylaşılmamaktadır. Bu nedenle doğru cevaplanması için birden fazla niteliğe sahip olunması gereken maddelerde, hangi niteliğin maddenin doğru cevaplanmasına daha çok katkı sağladığı belirlenememektedir. Bu kapsamda, maddenin doğru cevaplanmasında her bir niteliğin eşit katkı yaptığı varsayılan DINO modeli seçilmiştir.

Gerçek veri setine ait ilk Q matrisleri, PISA uygulamasına ait teknik raporlarda verilen bilgiler doğrultusunda oluşturulmuştur. Bu nedenle gerçek veri setine ait ilk Q matrislerinin doğruluğu bu raporlarda sunulan bilgilerin doğruluğu ile sınırlıdır. Mevcut araştırma kapsamında teknik raporlar (OECD, 2017a; 2017b) ve Tablo 5 temel alınarak PISA 2015 İPÇ yapısını temsil eden üç farklı Q matrisi oluşturulabilmektedir. PISA 2015

İPÇ uygulamasında A3 hücrelerini ölçmeye yönelik herhangi bir madde olmadığı için 12 nitelikli olarak oluşturulması gereken Q matrisi 11 nitelikli olarak oluşturulmuştur.

Son olarak, mevcut araştırmada bazı koşullar model kurulmasını engellediği için kayıp veri atamada model temelli veri atama yöntemleri kullanılamamış, bunun yerine iki yönlü veri atama yöntemi kullanılmıştır.

1.5. Tanımlar

Karmaşık Q matris yapısı: Mevcut araştırma sadece nitelik sayısının fazlalığını ifade etmektedir.

Nitelik: Bir öğrencinin bir test maddesini çözmek için sahip olması gereken prosedürleri, becerileri veya işlemleri belirtir. Mevcut araştırma kapsamında İPÇ yapısını oluşturan beceriler birer nitelik olarak ifade edilmiştir.

Sınıflandırma performansı: Sınıflandırma oranını ifade etmektedir.

Sınıflandırma oranı: Tahmin edilen ve bilinen gerçek sınıflandırma arasındaki uyum oranını yansıtır.

BÖLÜM II

KAVRAMSAL ÇERÇEVE

Bu bölümde, BTM, POBT, YSA ve PISA 2015 İPÇ yeterliği hakkında detaylı bilgiler yer almaktadır. Bilişsel tanılama alanyazınında ilk olarak parametrik modeller (BTM'ler) kullanılmaya başlanmış, Q matrisine ve analiz için veri yapısına (telafisel veya telafisel olmayan) ait özellikler bu çerçevede belirlenmiştir. Sonraları BTM'lerle analiz için veri yapısının ve Q matrisinin bilinmesi hususunda benzerlik gösteren parametrik olmayan model ve teknikler kullanılmaya başlanmıştır. Bu nedenle bu bölümde, öncelikle parametrik Bilişsel Tanı Modelleri (BTM'ler), ardından Parametrik olmayan Bilişsel Tanılama (POBT) ve Yapay Sinir Ağları (YSA'lar) hakkında bilgiler bulunmaktadır. Ayrıca araştırmanın ikinci bölümünde PISA 2015 İPÇ yeterliğine ait veriler kullanıldığı için bu bölümde PISA 2015 İPÇ hakkında bilgilere yer verilmiştir. Son olarak, parametrik olmayan bilişsel tanılama ve PISA 2015 İPÇ yeterliğine ilişkin alanyazında yer alan çalışmalara yer verilmiştir.

2.1. Araştırmanın Kuramsal Temelleri

Bu bölümde Bilişsel Tanı Modelleri (BTM'ler), Parametrik olmayan Bilişsel Tanılama (POBT) ve Yapay Sinir Ağları (YSA'lar) hakkındaki bilgilere yer verilmiştir.

2.1.1. Bilişsel Tanı Modelleri (BTM'ler)

Test puanının geleneksel karşılaştırmaları, akademik alanlardaki grup farklılıklarını tanımlamak anlamına gelir. Ancak grupların akademik alanda başarılı olmak için gerekli olan bilişsel niteliklere göre nasıl farklılaştığını açıklayamaz. Bu nedenle bilişsel tanı

modelleri ortaya çıkmıştır. BTM'ler, istatistiksel modeller arasındaki sınırlandırılmış/kısıtlanmış gizil sınıf modelleridir (Templin & Henson, 2006). İlk bilişsel tanı modeli, 1973 yılında oluşturulan Fischer'in Lineer Lojistik Test Modeli'dir (LLTM; Linear Logistic Test Model). LLTM, maddenin doğru cevaplanmasında gereken bilişsel adımları dikkate alan temel Rasch modelinin bir uzantısıdır. LLTM'de bir Rasch maddesinin gücü, kategorik yapıya sahip niteliklere dayalı güçlüklerin toplamı olarak hesaplanır. Tanısal değerlendirme için uygun hale getirilen LLTM'de bir maddenin gücü, bir maddenin doğru bir şekilde çözülmesi için kritik olan temel bilişsel düzeylerin veya faktörlerin etkilerinin bileşimidir (Fischer, 1973).

BTM'lerin ortaya çıkışlarından günümüze kadar geçen süreçte çeşitli şekillerde isimlendirildikleri görülmektedir. Sünbül, 2013'teki çalışmasında bu isimlendirmeleri şu şekilde listelemiştir:

bilişsel psikolojik modeller (cognitive psychometric models) (bkz. Rupp, 2007); bilişsel tanı modelleri (cognitive diagnostic models) (bkz., Nichols, Chipman, & Brennan, 1995); örtük tepki modelleri (latent response models) (bkz., Maris, 1995); sınırlandırılmış örtük sınıf modelleri (restricted latent class models) (bkz., Haertel, 1989; Macready & Dayton, 1976); çoklu sınıflandırma örtük sınıf modelleri (multiple classification latent class models) (bkz., Maris, 1995, 1999) ve yapısal madde tepki kuramı modelleri (structured item response theory models) (bkz., Mislevy, 2007; Rupp & Mislevy, 2007)”

BTM'lerde niteliklerle madde cevaplarının ilişkilendirilmesi Q matrisleri aracılığıyla olmaktadır. Bir Q matrisi, satırlarını maddelere ait bilgilerin, sütunlarını ise yapısal niteliklere ait bilgilerin oluşturduğu iki boyutlu bir yapıya sahiptir. Tablo 1'de üç madde ve üç nitelikten oluşan bir Q matrisi görülmektedir. Bu matris, bir öğrencinin birinci maddeyi doğru olarak cevaplayabilmesi için nitelik bir ve ikiye (toplama ve çıkarma işlemi bilgisine) sahip olması gerektiğini göstermektedir. Benzer şekilde öğrencinin madde ikiyi doğru cevaplayabilmesi için nitelik üçe (çarpma işlemi bilgisine); madde üçü doğru cevaplayabilmesi için ise nitelik bir ve üçe (toplama ve çarpma işlemi bilgisine) sahip olması gerekmektedir.

Q matrislerini çeşitli şekillerde sınıflandırmak mümkündür. Bu sınıflandırmalardan biri, matris yapısının basitliği veya karmaşıklığıdır. Basit yapılı bir Q matrisinde bir maddenin cevaplanması sadece bir niteliğe sahip olunmasına bağlı iken; karmaşık yapılı bir matriste bir maddenin doğru olarak cevaplanması en az bir niteliğe sahip olunmasına bağlıdır. Q matrislerinin sınıflandırılmasında bir diğer ölçüt nitelik yapılarıdır. Bazen bir niteliğe sahip olabilmek için başka niteliklere de sahip olmak gerekir. Nitelikler arasında bu tür ilişkileri barındıran matrisler ise hiyerarşik ve hiyerarşik yapıda olmayan matrisler şeklinde ayrılmaktadır. Hiyerarşik yapıda olmayan matrislerde nitelikler arası herhangi bir ilişki

beklenmez ve bir sorunun doğru cevaplanması için gereken nitelik ek olarak başka bir niteliğe sahip olmayı gerektirmez. Hiyerarşik yapıda olan matrislerde ise en az iki nitelik birbirini ile ilişkilidir ve bir nitelik diğerinin önkoşuludur. Böyle bir durumda öğrencinin maddeyi doğru cevaplayabilmesi için her iki niteliğe de sahip olması gerekmektedir. Q matrisi sınıflandırmasında son grupta ise niteliklerin matrislerde ifade edilmiş şekilleridir. Tablo 1’de ikili kodlanan (iki kategoriden oluşan) niteliklerin olduğu bir matris örneği görülmektedir. Burada “1” niteliğin varlığını; “0” ise yokluğunu ifade etmektedir. Bu tür bir kodlama niteliğin kişide bulunma miktarı hakkında bir bilgi sağlamamaktadır. Çok kategorili matrisler ile niteliğin öğrencide ve maddede bulunma miktarı hakkında bilgi sahibi olunabilmektedir.

Tablo 1

Örnek Q Matrisi

Maddeler	Nitelikler		
	Nitelik 1	Nitelik 2	Nitelik 3
	(Toplama)	(Çıkarma)	(Çarpma)
Madde 1 (3+4-5)	1	1	0
Madde 2 (2x3)	0	0	1
Madde 3 (1+3x4)	1	0	1

Bu çalışmada, PISA 2015 İPÇ verisine ait ilk Q matrisi PISA raporları doğrultusunda oluşturulmuştur ve maddelerin çözümü için niteliklerin yaptıkları bireysel katkılar raporlanmadığı için gerçek veriye ait Q matrisi iki kategoriden oluşan niteliklerden oluşturulmuştur. Ayrıca İPÇ yeterliğinin yapısını oluşturan beceriler arasında önkoşul ilişkisi bulunmadığından çalışma kapsamında hiyerarşik olmayan Q matris yapısı dikkate alınmıştır. Son olarak, gerçek veriye (PISA 2015 İPÇ) ait ilk Q matrisleri basit ve karmaşık yapıda temsil edilebilmektedir. PVAF yöntemine (bkz. de la Torre & Chiu, 2015) dayalı Q matris geçerlemesi yapıldıktan sonra elde edilen Q matrisleri ise karmaşık yapıdadır. Simülasyon verisi için gerçek verinin Q matris yapısına benzer karmaşık yapıda Q matrisleri kullanılmıştır. Simülasyon veri setlerine ait Q matrisleri EK-2’de raporlanmıştır.

Q matrisi iki kategoriden oluşan BTM’lerin matematiksel formülasyonlarında bireylerin bu iki kategoride bulunma olasılıkları v_1 ve v_2 ile gösterilmektedir. Bu olasılıklara “karışık oranlar (mixed ratio)” denilmektedir (Rupp vd., 2010). Genel olarak BTM’lerde c gizil

sınıfında bulunan r bireyinin i maddesini doğru cevaplama olasılığı aşağıdaki gibi hesaplanmaktadır (Rupp vd., 2010):

$$P(X_{ri} = 1|c) = \pi_{ic} \quad (1)$$

Eşitlikte X_{ri} , r bireyinin i maddesindeki gözlenen puanını ifade etmektedir. Bu eşitlik c sınıfındaki tüm katılımcılar için madde i 'ye doğru cevap verme olasılığı veya c sınıfındaki katılımcılar için madde i 'nin gücülüğü olarak da yorumlanabilmektedir. Her madde için, iki gizil sınıfın her biri için bir madde gücülüğü parametresi olacaktır. Genel olarak iki gizil sınıftan birinde olma olasılığının tahmin edilmesi yeterlidir. Diğer sınıfta bulunma olasılığı, tahmin edilen birinci sınıfta olma olasılığının birden çıkarılmasıyla hesaplanabilmektedir. BTM'lerde bir sınıfta bulunma olasılığının matematiksel gösterimi genel olarak aşağıdaki gibidir (Rupp vd., 2010):

$$P(X_r = x_r) = \underbrace{\sum_{c=1}^C v_c}_{\text{Yapısal bileşen}} \underbrace{\prod_{i=1}^I \pi_{ic}^{x_{ir}} (1 - \pi_{ic})^{1-x_{ir}}}_{\text{Ölçme bileşeni}} \quad (2)$$

Burada $P(\cdot)$ olasılığı, X_r ve x_r , tüm bireylere ait gözlenen cevapları; x_{ir} , r kişinin madde i 'deki gözlenen cevabı; v_c , c sınıfında bulunma olasılığını; π_{ic} , c gizil sınıfındaki bir kişinin madde i 'ye doğru cevap verme olasılığını ve $\sum(\cdot)$, $c=1$ 'den C 'ye kadar tüm gizil sınıfların toplandığını gösterir. $\prod(\cdot)$ ise $i=1$ 'den I 'ya kadar tüm maddelere ait olasılıkların çarpılacağını ifade etmektedir. Formül 2'de, I 'nın temsil ettiği maddeler yalnızca 0 ve 1 değerlerini alabilmektedir ve Formül 2, I maddeleri için rastgele bir Bernoulli değişkenleri kümesini ifade etmektedir. Bu durum her bir gizil sınıf için, belirli bir cevap örüntüsünün ortak olasılığını (joint probability) veya olasılığını (likelihood) sağlamaktadır. Formül 2'deki çarpma işlemi, temel olasılık teoremindeki bağımsız olayların birlikte gerçekleşme olasılıklarının hesaplanmasına uygun olarak kullanılmıştır ve belirli bir gizil sınıftaki bireylerin maddelere verdikleri cevapların bağımsızlığını ifade etmektedir. Formül 2'nin çarpma kısmı gözlemlenemeyen gizil değişkenleri gözlemlenebilir verilerle doğrudan bağlayan bileşen olduğu için modelin ölçüm bileşeni olarak; toplama kısmı ise gizil nitelik değişkenleri arasındaki ilişkiyi tanımlamada kullanıldığı için modelin yapısal bileşeni olarak adlandırılmıştır.

Alanyazında, Q matrisi iki kategorili niteliklerden oluşturulmuş farklı BTM modelleri bulunmaktadır (Örn. DINO, Reparameterized Unified Model-RUM, vb.). Bu modeller yukarıdaki formülün ölçme bileşeni kısmında farklılıklara sahiptir. Ayrıca BTM'lerin

genel olarak çeşitli ölçütlere göre sınıflandırıldığı görülmektedir. Bu ölçütleri (Rupp vd., 2010) şu şekilde özetlemek mümkündür:

1. Maddelerin (gözlenen değişkenlerin) ölçek tipi (ikili/çoklu puanlanan)
2. Niteliklerin (gizil özelliklerin) ölçek tipi (ikili/çoklu puanlanan)
3. Veri yapısının telafisel olup olmaması.

BTM'ler, bir maddenin doğru cevaplanması için bireyin tüm niteliklere sahip olması koşuluna bağlı olarak telafisel ve telafisel olmayan modeller olarak ikiye ayrılmaktadırlar. Telafisel modellerde bireyin bir maddenin doğru cevaplanması için gerektirdiği niteliklerden en az birine sahip olması yeterlidir. Telafisel olmayan modellerde ise birey maddenin doğru cevaplanması için gerektirdiği tüm niteliklere sahip olmalıdır. BTM'ler telafisel olup olmamaları durumuna benzer şekilde bağlayıcı (conjunctive) ve bağlayıcı olmayan (disjunctive) modeller olarak da ayrılmaktadır. Bağlayıcı olmayan modellerde bireyin gerekli niteliklerden birine sahip olması, maddenin doğru cevaplanma olasılığını en üst düzeye çıkarmaktadır. Bu nedenle bağlayıcı olmayan modeller telafisel modellerin, bağlayıcı modeller ise telafisel olmayan modellerin özel durumlarıdır.

Araştırma kapsamında, gerçek veri seti ile uyumlu olması için simülasyon veri seti, telafisel modellerden DINO model temel alınarak üretilmiştir. Telafisel modellerden DINO modelin seçilmesinin sebebi, bu modelin alanyazında sıklıkla kullanılmasıdır. Araştırmada simülasyon verisinin DINO temelli olması nedeniyle POBT ve YSA'nın sınıflandırma oranları DINO model sınıflandırma oranları ile karşılaştırılmıştır. Böylece POBT ve YSA'nın ne kadar iyi sınıflandırma yaptıkları tespit edilmek istenmiştir.

2.1.1.1. Deterministic-Input, Noisy-Or Gate (DINO)

Telafisel modellerin basit bir örneği olarak Templin ve Henson (2006) tarafından ortaya atılmış bir modeldir. DINO modelinde, bir maddenin doğru cevaplanması için gerekli niteliklerden sadece birine sahip olunması maddenin doğru cevaplanma olasılığını arttırmaktadır. DINO modelinin her bir madde düzeyinde, maddenin doğru cevaplanma olasılıklarını etkileyen sadece iki parametre vardır: kayma (s) ve tahmin etme (g). DINO modelinde madde parametrelerinin tahmini EM algoritması veya Markov zinciri Monte Carlo kullanılarak yapılabilmektedir. Ençok Olabilirlik Kestirimi (maximum likelihood estimation, MLE), Maksimum Sonsal (maximum a posteriori, MAP) ve Beklenen Sonsal (expected a posteriori, EAP) yaklaşımları da yaygın olarak bireylerin sınıflandırılmasında

kullanılmaktadır (Huebner ve Wang, 2011). DINO modeli matematiksel olarak aşağıdaki eşitlikte belirtilmiştir (Rupp vd., 2010):

$$\omega_{ic} = 1 - \prod_{a=1}^A (1 - \alpha_{ca})^{q_{ia}} \quad (3)$$

Formül 3'teki ilk bileşen ω_{ic} 'dir ve modelin gizil cevap değişkenidir, c sınıfındaki bireylerin madde i'ye verdikleri cevapları ifade etmektedir. Bu değişken bireyin ölçülen niteliklerden en az birine sahip olup olmadığını belirten bağlayıcı olmayan kümeleme kuralı kullandığı için modelin belirleyici girdi (deterministic input) kısmını oluşturmaktadır. Bu değişkeni oluşturan yapısal bileşen ise bireyleri bir parka belirli bir şekilde girmeye veya çıkmaya zorlayan kapı gibi bir işlev gördüğü için veya-kapısı (or-gate) olarak adlandırılmaktadır.

DINO modelde, a niteliği madde i'yi ölçmüyor ise $q_{ia}=0$ olur. Eğer madde i, a niteliği ile ölçülüyorsa $\alpha_{ca}=0$ veya $\alpha_{ca}=1$ olmasına bakılmaksızın $q_{ia}=1$ değerini almaktadır. $\alpha_{ca}=1$ olması c gizil sınıfındaki bireylerin en az bir niteliğe sahip oldukları anlamına gelmektedir. Bu nedenle tek bir niteliğin varlığı, diğer niteliklerin eksikliklerini telafi edebilmektedir. DINO modelde c gizil sınıfındaki bir bireyin i maddesine doğru cevap verme olasılığı aşağıdaki eşitlikle hesaplanır:

$$\pi_{ic} = P(X_{ic} = 1 | \omega_{ic}) = (1 - s_i)^{\omega_{ic}} g_i^{1 - \omega_{ic}} \quad (4)$$

Formül 4'te görüldüğü üzere bireyin maddeyi doğru cevaplama olasılığı gizil cevap değişkeni (ω_{ic}), kaydırma (s) ve tahmin (g) parametrelerine bağlıdır. Burada P olasılık; π_{ic} , c gizil sınıfındaki bir kişinin madde i'ye doğru cevap verme olasılığı; X_{ic} , c sınıfındaki öğrencilerin i maddesine ait gözlenen cevapları; ω_{ic} , gizil c sınıfında bulunan bir öğrencinin i maddesine verdiği gizil tepki değişkeni; $(1 - s_i)$, i maddesinde kaydırma yapmama olasılığı ve g, i maddesini tahmin etme olasılığıdır. DINO modelin kaydırma (s) ve tahmin (g) parametrelerine ait formüller aşağıdaki gibidir:

$$g_i = P(X_{ic} = 1 | \omega_{ic} = 0) \quad (5)$$

$$s_i = P(X_{ic} = 0 | \omega_{ic} = 1) \quad (6)$$

DINO modelde gürültüye (noisy) sebep olan bu parametreler, madde düzeyinde hesaplanırlar. Başka bir ifade ile model nitelikler arasında eşit sınırlamayla her bir madde için birer kaydırma ve tahmin parametresi üretir. DINO model, bağlayıcı olmayan kümeleme kuralını kullandığı için tahmin parametresi (g_i), bireyin ölçülen tüm niteliklere sahip olmadığında 1 puan alma olasılığını gösterirken kaydırma parametresi (s_i), birey

ölçülen niteliklerden en az birine sahip olduğunda 0 puan alma olasılığını temsil etmektedir.

2.1.2. Parametrik Olmayan Bilişsel Tanılama (POBT)

Parametrik olmayan Bilişsel Tanılama’da (POBT; Nonparametric Cognitive Diagnosis), gözlenen cevap vektörlerinin ideal cevap örüntülerine yakınlığı incelenerek öğrenciler sınıflandırılır (Chiu & Douglas, 2013). Gözlenen cevap vektörleri öğrencilerin teste verdikleri iki puanlı cevaplarının vektörleridir ve genellikle bir i öğrencisine ait gözlenen cevap y_i ile temsil edilir. İdeal cevap örüntüleri ise öğrencilerin sahip oldukları niteliklerle maddelerin doğru cevaplanmaları için gereken niteliklerin karşılaştırılması sonucu ortaya çıkması beklenen madde cevap örüntüleridir. Bu cevap örüntülerinin oluşturulmasında öğrencilerin cevapları tahmin etmesinden veya maddeleri cevaplarken kaydırma yapmalarından kaynaklanan hatalar bulunmamaktadır. Bu nedenle bu cevap örüntüleri “ideal” olarak isimlendirilmektedirler. İdeal cevap örüntülerinin kestirilmesi için veri yapısının telafisellik durumu (telafi edici veya telafi edici olmayan) bilinmelidir. Telafisel olan modellerden DINO modeli için i öğrencisinin j maddesine vermesi beklenen (ideal) cevabı Formül 7 ile hesaplanır:

$$\omega_{ij} = 1 - \prod_{k=1}^K (1 - \alpha_{ik})^{q_{jk}} \quad (7)$$

α_{ik} : i öğrencisinin k . niteliğe sahip oluş durumu

q_{jk} : j maddesinin k . niteliği gerektirme durumu

Bu formülден öğrencilerin ideal cevaplarının Q matrisine dayalı ve α ’nın bir fonksiyonu olarak üretildiği görülmektedir.

k tane niteliğin bulunduğu bir test için $m=1,...,2^k$ tane nitelik profili bulunduğundan α_i için olası tüm ideal cevaplar $\omega_m = \omega_1, ..., \omega_{2^k}$ şeklinde ifade edilebilir. ω_i, α_i tarafından belirlendiği için bir i öğrencisinin α_m nitelik profili altında gözlemlenen (y_i) ve ideal cevap örüntüleri arasındaki uzaklık $d(y_i, \alpha_m)$ şeklinde ifade edilebilir. POBT’de kestirilen nitelik örüntüsü ($\hat{\alpha}_i$), tüm ideal madde cevap örüntüleri ile i kişinin gözlemlenen cevap örüntüleri arasındaki uzaklığı en aza indiren nitelik örüntüsüdür (Chiu & Douglas, 2013). Farklı bir şekilde ifade edildiğinde, $\hat{\alpha}_i$, tüm ideal madde cevap örüntüleri arasında gözlemlenen madde cevap örüntülerine en yakın veya en benzer olan ideal madde cevap örüntüsüne ait nitelik örüntüsü olarak tanımlanabilir (Chiu vd., 2017).

$$\hat{\alpha}_i = \arg \min D(y_i, \alpha_m) \quad m \in (1, 2, \dots, 2^k) \quad (8)$$

Alanyazında telafisel olmayan modellerden DINA modeli için $\hat{\alpha}$ 'nın Wang ve Douglas (2015) tarafından belirli örneklem büyüklüğü ve kabul edilebilir madde sayısı koşulları altında tutarlı olduğu kanıtlanmıştır. Fakat benzer bir çalışma telafisel modellerden DINO model için yapılmamıştır. DINA ve DINO modelin ikili (duality) ilişkisine (bkz. Köhn & Chiu, 2016) dayalı olarak Wang ve Douglas'ın (2015) çalışmalarında buldukları yargı DINO modeline genişletilebilir (Chiu & Köhn, 2019). Buradan $\hat{\alpha}_i$ 'nin DINO model için tutarlı olduğu söylenebilir.

POBT'ler gözlemlenen madde cevap örüntülerini 2^k olası nitelik sınıflarına ait ideal cevap profillerinin her biri ile karşılaştırarak nitelik sınıfı üyeliğini tahmin etmektedir (Chiu vd., 2017). İki vektörün ne kadar benzer veya farklı olduğunu ölçmek için ölçüt olarak çeşitli uzaklık ölçütleri kullanılabilir. Bu nedenle, bu sınıflandırma yaklaşımında uzaklık (veya benzerlik) ölçüsünün seçimi çok önemlidir (Akbay, 2016). Bu uzaklıklardan bazıları Hamming, ağırlıklandırılmış Hamming ve cezalandırılmış (penalized) Hamming uzaklığıdır (Chiu & Douglas, 2013). Hamming uzaklığı, ikili veriler için sıklıkla kullanılır ve iki vektördeki girdilerin kaç kez birbirine benzemediğini sayar. Örneğin, $A = (1, 1, 0, 1, 0)$ ve $B = (1, 0, 1, 1, 1)$ dizelerinde farklı üç sayı bulunmaktadır. Burada A ve B arasındaki Hamming uzaklığı üçtür. Ağırlıklandırılmış Hamming uzaklığının, testteki bazı maddelerin cevaplarının diğerlerinden daha fazla varyansa sahip olması durumlarında kullanılması olasıdır. Bu durumlarda, daha küçük varyansları olan maddelerin ideal ve gözlenen cevapları arasındaki farkları, diğer maddelerin sahip oldukları farklardan daha az olduğu için daha fazla ağırlığa sahip olmaktadır (Paulsen, 2019). Benzer düşünceyle, testteki farklı madde türlerine ait tahmin etme ve kaydırma parametreleriyle ilişkili terimlere daha fazla ağırlık veren cezalandırılmış Hamming uzaklığı hesaplanmaktadır. Ağırlıklandırılmış Hamming uzaklığının diğer uzaklık ölçülerine göre daha iyi performans göstermesi (Chiu & Douglas, 2013) nedeniyle bu araştırmada uzaklık ölçüsü olarak ağırlıklandırılmış Hamming uzaklığı kullanılmıştır. Ayrıca, ağırlıklandırılmış Hamming uzaklığının a) daha verimli olduğu, b) olası ideal cevap örüntüleri arasındaki bağların sayısını azalttığı ve c) daha yüksek sınıflandırma doğruluğu sonuçları ürettiği belirlenmiştir (Paulsen, 2019). Telafisel yapıya sahip bir veri için Hamming uzaklığı Formül 9 ile hesaplanmaktadır:

$$d_h(y_i, \alpha_m) = \sum_{j=1}^J |y_{ij} - \omega_{mj}| \quad (9)$$

Telifisel yapıya sahip bir veri için Ağırlıklı Hamming uzaklığı ise Formül 10 ile hesaplanır:

$$d_{wh}(y_i, \alpha_m) = \sum_{j=1}^J \frac{1}{p_j(1-p_j)} |y_{ij} - \omega_{mj}| \quad (10)$$

POBT'nin kullanımını sınırlandırabilecek bazı durumlar aşağıda özetlenmiştir:

- POBT'de iki vektörün elemanları arasındaki uyumsuzluk üzerinden uzaklık hesapladığı için niteliklerin ikili (var-yok) olarak incelendiği durumlarda kullanılabilir (Akbay, 2016; Shuying, 2016).
- BTM'de olduğu gibi bir Q matrisi gerektirir (Akbay, 2016; Chiu & Douglas, 2013; Chiu & Köhn, 2019).
- POBT'nin uygulanması için veri yapısının (telafi edici ve telafi edici olmayan) bilinmesine gerek olmamasına rağmen ideal cevap örüntüsünün kestirilmesi için veri yapısının bilinmesi gerekmektedir (Chiu & Douglas, 2013; Chiu vd., 2017).
- Doğru sınıflandırma oranının hesaplanmasında kullanılabilecek, niteliklerin sonsal olasılık dağılımları hesaplanamaz (Chiu & Douglas, 2013; Wang & Douglas, 2015).

POBT'nin avantajlarını ise aşağıdaki şekilde özetlemek mümkündür:

- Parametrik modellerin sınıf ortamı gibi küçük örneklemelerde genel olarak karşılaştığı istikrarsız ve güvenilir olmayan tahminlerden etkilenmezler (Chiu vd., 2017). Bu nedenle parametrik bir modelin madde kalibrasyonu mümkün olmadığında bilişsel tanı koymayı mümkün kılarak (Akbay, 2016; Chiu & Douglas, 2013; Wang & Douglas, 2015) potansiyel uygulamaları büyük ölçüde artırabilirler.
- Çok sayıda niteliğin ölçüldüğü testlerde rahatlıkla kullanılabilirler (Chiu & Douglas, 2013). Sınıflandırma analizi kısa sürede gerçekleşir (Chiu & Douglas, 2013) ve daha ekonomiktir (Chiu & Köhn, 2019). Tüm bu özellikler POBT'nin, bilgisayar ortamında bireylere uyarlanmış testler için bir hesaplama yöntemi olarak kullanılmasını destekler (Chiu & Köhn, 2019).
- Veri yapısının biraz yanlış belirtilmesi halinde, s (kayma) ve g (tahmin) parametre değerlerinin çok yüksek olmaması koşuluyla parametrik modellerden daha iyi performans gösterebilir (Chiu & Douglas, 2013).

2.1.3. Yapay Sinir Ağları (YSA)

İnsan beynindeki nöron yapılarından ve işlevlerinden esinlenerek geliştirilen modellerdir. Genellikle üç şekilde kullanılırlar: (1) biyolojik sinir sistemleri ve zekâ modeli olarak, (2) gerçek zamanlı, uyarlanabilir, sinyal işlemcileri olarak ve (3) veri analitik yöntemleri olarak (Lamb, Annetta, Vallett & Sadler, 2014). Bu çalışmada YSA modeli, bir veri analitik yöntemi olarak kullanılmıştır. Veri analitik yöntemi olarak YSA'lar, insanın biyolojik nöral sistemlerini anlamak ya da girdi ve çıktılar arasındaki karmaşık ilişkileri modellemek için doğrusal olmayan istatistiksel bir yöntem olarak kullanılmaktadırlar. Garson'un 1998 yılındaki çalışmasında YSA analizinin çok sayıda girdiyi az sayıda çıktı birimine eşlemeyi içeren sınıflandırma problemleri için kullanılabileceğini belirtmesiyle birlikte YSA'lar sınıflandırma alanında da kullanılmaya başlanmıştır.

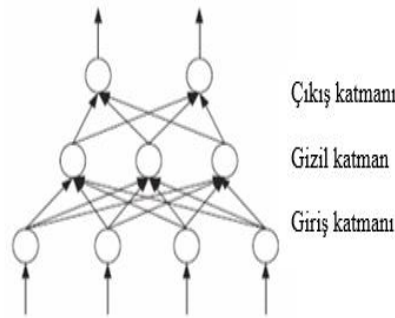
YSA'lar sınıflandırma alanında, parametrik olmayan istatistiksel modeller sınıfında yer almaktadır (Cui vd., 2016). İstatistiksel modellerle YSA modelleri arasında önemli benzerlikler olmasına rağmen YSA'nın terminolojisi ve dili, istatistiksel modellerden farklıdır (Lamb vd., 2014). Gierl ve diğerlerinin 2007 yılında YSA'ların AHM ile birlikte kullanılmalarını önermelerinin ardından YSA'lar BTM kapsamında da kullanılmaya başlanmıştır. Böylece BTM alanında incelenen cevap modelleri, YSA analizi ile kalibre edilen beklenen cevap modeli bağlantı fonksiyonları aracılığıyla nitelik modellerine dönüştürülebilmektedir (Shu vd., 2013). Başka bir ifadeyle bireylerin gözlemlenen cevap örüntülerine dayanarak, onların belirli nitelik kombinasyonlarına sahip olma olasılıkları tahmin edilebilmektedir (Gierl vd., 2007; 2008). YSA'lar önceden belirlenmiş veri setleriyle eğitildikten sonra analiz için belirli örneklem büyüklüğü gerektirmemektedirler (Shu vd., 2013). Bu özellikleri, onların sınıflandırma alanında küçük örneklem etkisini inceleyen çalışmalara konu olmalarına sebep olmuştur. Nitekim YSA'ların BTM kapsamında küçük örneklem üzerinde farklı yöntemlerle sınıflandırma performanslarının karşılaştırıldığı çalışmalar mevcuttur (bkz. Cui vd., 2016; Paulsen, 2019). Bu araştırmada da YSA'ların BTM kapsamında, farklı koşullar altında sınıflandırma performansları incelenmiştir.

BTM kapsamında denetlenen (supervised) ve denetlenmeyen (unsupervised) şeklinde iki farklı öğrenme paradigmasına sahip YSA'larla çalışıldığı görülmektedir. Bu iki paradigma arasındaki temel fark, denetlenen YSA'ların öğrenmeyi mümkün kılmak için çıktı değerlerini gerektirirken denetlenmeyen YSA'ların gerektirmemesidir (Cui vd., 2016). Denetlenen YSA'ların amacı, gerçek ve tahmin edilen çıktı değerleri arasındaki farkı en

aza indirerek, istenen hassasiyet seviyesinde giriş ve çıkış birimleri arasındaki ilişkiyi tahmin etmektir. Denetlenmeyen YSA'larda ise gerçek çıktı değerleri bilinmemektedir. Bu nedenle, denetlenmeyen YSA'lar, çıktı sınıflarını oluşturmak için girdi vektörleri arasındaki benzerlikten yararlanırlar (Cui vd., 2016). Bu araştırmada BTM'lerin doğrulayıcı bakış açısına uyumlu olmaları nedeniyle denetlenen öğrenme paradigması kullanılmıştır. Aşağıda denetlenen öğrenme paradigmasına sahip, alanyazında sıklıkla kullanılan çok katmanlı algılayıcı YSA modeline ait detaylı bilgilere yer verilmiştir.

2.1.3.1. Çok Katmanlı Algılayıcı (Multilayer perceptron) YSA

Denetimli YSA'larda en sık kullanılan modeldir ve nöronları/düğümüleri katmanlı bir mimaride düzenlendiği için çok katmanlı algılayıcı (ÇKA; Multilayer perceptron) model olarak adlandırılmaktadır (Cunningham & Ogilvie, 1972). Çok katmanlı algılayıcı yapay sinir ağları (ÇKA-YSA), girdi birimleriyle uyarıcıları alıp gizil birimlerle bu uyarıları çıktı birimine taşıyan paralel işleme mimarisi (Gierl 2007; Garson 1998) olarak tanımlanabilirler. Ağın amacı, giriş ve çıkış birimleri arasında fonksiyonel ilişkiyi kurmaktır (Gierl vd., 2007).



Şekil 1. Çok katmanlı algılayıcı YSA örneği. Cui, Y., Gierl, M. & Guo, Q. (2016). Statistical classification for cognitive diagnostic assessment: An artificial neural network approach. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 1-24 kaynağından erişilmiştir.

ÇKA-YSA mimarisinin bir örneği Cui ve diğerlerinin (2016) çalışmasında Şekil 1'deki gibi gösterilmiştir. Genel olarak bu YSA'lar üç veya daha fazla katmandan oluşurlar: bir giriş katmanı, en az bir gizil katman ve bir çıkış katmanı. Her bir katman kendine özgü yorumu olan nöronlardan (Şekil 1'de daire olarak gösterilmiştir) oluşur. Ağdaki gizil katman/katmanlar giriş nöronlarının etkileşimlerinin çıkış nöronları üzerindeki etkilerini

modellemede yardımcı olur. Eğer gizil katman ortadan kaldırılırsa sadece giriş katmanı nöronlarının çıkış katmanı nöronları üzerindeki doğrudan etkileri modellenir.

ÇKA-YSA'da, belirli bir katmandaki nöronlar birbiriyle doğrudan etkileşemez, sadece bitişik katmanlarda bulunan nöronlarla bağlantı kurarlar. Örneğin, Şekil 1'de, giriş katmanındaki tüm nöronlar, gizil katmandaki nöronlara, gizil katmandaki nöronlar da çıkış katmanındaki nöronlara bağlanır. Şekil 1'deki tek yönlü oklar bu bağlantıların ağırlıklarını gösterir. Bir YSA'daki ağırlıkları tahmin etme işlemi veya giriş katmanının çıkış katmanı ile ilişkilendirilmesi süreci "YSA eğitimi veya öğrenimi" olarak adlandırılır. Ağırlık eğitimi, kestirilen bir cevap örüntüsü için beklenen ve gözlenen nitelik değerleri arasındaki farkı (hata) her defasında en aza indirecek şekilde tekrarlanarak yapılır. Bu nedenle, sinir ağı güçlü bir örüntü tanıma tekniği olarak işlev görür ve yeterli eğitim verisi verildiğinde, giriş ile çıkış katmanı arasındaki herhangi bir ilişkiyi hassasiyetle kurabilir (Gierl, 2007). Tekrarlamalı (iteratif) bu süreçteki ağırlık tahminleri, standart normal dağılımdan alınan rastgele değerlerle başlatılır (R paketi neuralnet in varsayılan yaklaşımı).

Ağırlık tahminleri geri yayılım (backpropagation, bkz. Rumelhart, Hinton, & Williams, 1986) ve esnek geri yayılım (resilient backpropagation) algoritmaları ile yapılabilmektedir. Geleneksel geri yayılım algoritmasına dayanan esnek geri yayılım algoritmasında ondan farklı olarak, eğitim süreci sırasında her bir ağırlık için değiştirilebilen ayrı bir öğrenme oranı kullanılır. Bu, tüm eğitim süreci ve tüm ağı için uygun olan genel bir öğrenme oranı tanımlama sorununu çözmektedir. Ayrıca esnek geri yayılım algoritmasında ağırlıkları güncellemek için kısmi türevlerin büyüklüğü yerine sadece onların işaretini kullanılarak tüm ağı üzerinde öğrenme hızının eşit etmesi sağlanmaktadır (Günther & Fritsch, 2010). Esnek geri yayılım algoritmasının ağırlık geri takibi olan (bkz. Riedmiller, 1994), ağırlık geri takibi olmayan (bkz. Riedmiller & Braun, 1993) ve modifiye edilmiş küresel yakınsak (bkz. Anastasiadis, Magoulas, & Vrahatis, 2005) algoritması olmak üzere üç türü bulunmaktadır. Ağırlık geri takibi, son tekrarlamayı geri alma ve bir sonraki adımda ağırlığa daha küçük bir değer ekleme tekniğidir. Geriye dönük ağırlık takibi kullanılmadığında algoritmanın işlem süreci sınırlanmaktadır. Anastasiadis ve diğerleri (2005) tarafından modifiye edilmiş küresel yakınsak algoritmasında ise en küçük mutlak kısmi türevle ilişkili öğrenme oranı ya da en küçük öğrenme oranı eklenerek ağırlık öğrenme hızında değişiklik yapılmaktadır. Bu çalışma kapsamında ağırlık geri takibi ile esnek geri yayılım yaklaşımı kullanılmıştır. Yapay sinir ağının eğitimi tamamlandıktan sonra, girdilerin bilindiği, ancak çıktılarının bilinmediği yeni bir veri setinin analizi yapılır.

BTM kapsamında, YSA'ların giriş katmanı nöronlarından her biri, testi oluşturan bir maddeyi temsil eder ve girdi olarak o maddelere verilen cevaplar kullanılır. Çıkış katmanının nöronları ise önceden belirlenmiş olan testin ölçmeye çalıştığı nitelikler şeklinde yorumlanır. Örneğin, toplam dört niteliği ölçen ve 16 maddeden oluşan bir testte ÇKA-YSA, 16 giriş nöronundan ve dört çıkış nöronundan oluşmaktadır. YSA'yı eğitmek için kullanılan girdiler Q matrisinden türetilen beklenen cevap örüntüleri, yani Q matrisinden üretilen ideal cevap vektörleri kümesidir. Çıktılar ise Q matrisinin doğru olduğu varsayımıyla birlikte ilgili nitelik profilleridir (Cui vd., 2016; Cui, Gierl, & Guo 2017). Ağın eğitimi, bu ideal cevap vektörleri ile ilgili nitelik profilleri arasındaki ilişkinin kurulmasını ifade etmektedir. Eğitim süreci, sinir ağı ilişkileri öğrenene kadar devam eder. Ağ öğrenimini tamamladıktan sonra analiz için öğrencilerin test maddelerine verdikleri cevaplar ağa girdi olarak sunulur.

İstatistiksel olarak ÇKA-YSA'yı, doğrusal olmayan çok değişkenli bir regresyon modeline (Cui vd., 2016), bir yapısal eşitlik modeline (Dijkstra & Henseler, 2011), lojistik veya log-doğrusal modele (Briggs & Cerci, 2017) benzetmek mümkündür. Bu benzetmede giriş katmanı nöronları bağımsız değişkenler kümesine benzerken, çıkış katmanı nöronları bağımlı değişkenlere benzer. Giriş ve çıkış katmanları arasındaki ilişki, regresyon modelindeki matematiksel formulla benzerdir ve nöronlar arasındaki bağlantıların gücü, modeldeki beta katsayılarına benzer ağırlıklar tarafından kontrol edilir (Cui vd., 2010). Gizil katmandaki nöron sayısı, genel olarak kullanıcı tarafından bağımsız ve bağımlı değişkenler arasındaki doğrusallık derecesinin kontrol edilmesi için seçilir. Gizil katmandaki nöron sayısındaki artış, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi daha çok doğrusal olmayan yapar (Cui vd., 2016).

Bir YSA eğitiminin matematiksel formunu aşağıdaki adımlarla özetlemek mümkündür:

I giriş nöronlarını, J gizil nöronlarını ve K çıkış nöronlarını temsil ettiği üç katmanlı ÇKA-YSA olsun. İlk adımda, tüm giriş nöronlarının ağırlıklı bir toplamı elde edilir.

$$a_j = \sum_{i=1}^I W_{ji} X_i \quad (11)$$

a_j : j gizil nöronu için ağırlıklı toplamı,

W_{ji} : i giriş nöronundan j gizil nörona olan bağlantı ağırlığı ve

X_i : i giriş nöronunun değeridir.

Gizil nöron değerinin hesaplanması için bu toplam, etkinleştirme fonksiyonu (activation function; $f(\cdot)$) tarafından aşağıdaki forma dönüştürülür:

$$h_j = f(a_i) = f(\sum_{i=1}^l W_{ji}X_i) \quad (12)$$

Bilişsel tanılama kapsamında YSA’larda aktivasyon fonksiyonu olarak lojistik/sigmoid ve hiperbolik tanjant fonksiyonlarının kullanıldığı görülmektedir. Alanyazında aktivasyon fonksiyonu olarak sıklıkla lojistik fonksiyon tercih edildiği (Guo vd., 2017) için bu araştırmada da her iki bağlantı ağırlığının belirlenmesinde lojistik fonksiyon kullanılmıştır. Lojistik fonksiyonunu kullanmanın bir yararı, değerlerin 0 ile 1 arasında değişmesidir. Bu, nöronlar tarafından alınan değerlerin olasılıksal yorumlanmasına izin vermektedir (McClelland’dan aktaran Cui vd., 2010). Buradan hareketle bu araştırmada, daha yüksek bir değer alan öğrencinin belirli bir niteliğe sahip olma olasılığının daha yüksek olduğu şeklinde yorum yapılmıştır.

Gizil nöronların değerleri elde edildikten sonra, çıkış nöronlarının değerleri hesaplanmalıdır. Çıkış nöronları, gizil nöronların hesaplanmasına benzer şekilde hesaplanmaktadır. Yani, öncelikle tüm gizil nöronların ağırlıklı bir toplamı elde edilir. Daha sonra bu toplam lojistik fonksiyon ile aşağıdaki forma dönüştürülür:

$$\hat{Y}_k = f(b_k) = f(\sum_{j=1}^l V_{kj}h_j) \quad (13)$$

$\hat{Y}_k$: k . çıkış nöronunun tahmini değeri,

$f(\cdot)$: aktivasyon fonksiyonu,

b_k : k . çıkış nöronu için gizil nöronlarının ağırlıklı toplamı,

V_{kj} : j gizil nöronundan k çıkış nöronuna olan bağlantı ağırlığı ve

h_j : j gizil nöronunun değeridir.

Giriş nöronlarının ardışık iki dönüşümle çıkış nöronlarıyla ilişkilendirilmesi, hata fonksiyonu en az değere veya önceden belirlenen değere indirgenene veya hata fonksiyonu sabitlenene kadar devam eder. Başka bir ifadeyle, hata fonksiyonunu en aza indiren bağlantı ağırlıklarının değerleri, model parametrelerinin tahmini için seçilir. Araştırma kapsamında bir hata hesaplama fonksiyonu olan çapraz entropi (cross-entropy) değeri incelenerek gizil katman ve nöron sayısı belirlenmiştir.

ÇKA-YSA eğitilirken altta yatan modelin doğrusal olmama derecesini kontrol etmek için gizil katman sayısı ve bu katmanların nöronlarının sayısı belirtilmelidir. Pratikte, en uygun

sayıda gizil katman ve nöronlarının seçilmesine yardımcı olmak için denemeler yapılabilir. Örneğin, bir ÇKA-YSA ilk olarak sıfır gizil nöronla eğitilir ve hata fonksiyonunun değeri kaydedilir. Gizil nöronların sayısı daha sonra bir artırılır ve hata kararlı hale gelinceye veya yükselmeye başlayıncaya veya maksimum tekrarlama sayısına ulaşana kadar süreç tekrarlanır.

ÇKA-YSA'nın eğitimiyle ilgili olası birkaç problem bulunmaktadır. Bunlardan birincisi ağırlık parametrelerinin doğrudan tanımlanamadıkları ve rastgele bir değerle kestirilmeye başlandıkları için yerel minimum değer (local minima) üzerinde yakınlaşma eğilimi (Günther & Fritsch, 2010) göstermeleridir. Bu tür problemler için önlem almak ve sonuçların yorumlanabilirliğini arttırmak için çoklu YSA eğitimlerinde olasılık tahminlerinin ortalaması alınabilir, tüm ağırlıklar için sabit bir başlangıç değeri kullanılabilir, nihai ağırlık tahminlerinin değişkenliği araştırılabilir veya birkaç küresel arama tekniği (örn., Parçacık sürü optimizasyonu; Al-Shareef & Abbod, 2010) kullanılabilir (Briggs & Circi, 2017). Cui vd. (2016) ise bu problemle baş etmenin en basit yolunun tüm eğitim sürecini farklı bir rastgele ağırlık seti ile tekrarlamak olduğunu belirtmişlerdir.

ÇKA-YSA'nın eğitim süreciyle ilgili olası bir diğer problem, ağırlık ideal cevap örüntüleri ve ilgili nitelik profillerine dayalı olarak eğitilebilmesidir. Bu özellik, ÇKA-YSA'ların küçük örneklerde kolaylıkla sınıflandırma yapmalarına yardımcı olurken ampirik verilerde doğal olarak meydana gelen ideal cevap modelinden sapmaları hesaba katmadığı anlamına da gelir. İdeal verilere dayanan iyi kalibre edilmiş bir ağırlık, gözlemlenen verilerle ideal cevap modellerinden sık sık saparak kötü performans gösterebilir (Cui vd., 2016; Shu vd., 2013).

ÇKA-YSA'nın eğitiminde ideal cevap verilerinin kullanılmasıyla ilgili başka bir problem ise bir test ile ölçülen nitelik sayısının az olması durumunda, sadece birkaç ideal cevap örüntüsünün (örn. iki niteliğin ölçüldüğü bir test için dört) olmasıdır. Çok az sayıda ideal cevap örüntüsü, ÇKA-YSA'daki bağlantı ağırlıklarının tahminlerinde kararsızlığa yol açabilir. Cui vd. (2016), nitelik sayısının ÇKA-YSA'nın performansını nasıl etkilediğinin araştırılması gerektiğini vurgulamışlardır. Buradan hareketle çalışmada ele alınan nitelik koşullarında (3, 5 ve 7), düşük nitelik etkisinin oluşup oluşmayacağı incelenmiştir.

YSA'larda eğitim tamamlandıktan sonra örneklem büyüklüğü çok önemli değildir ve bu, onun güçlü yönünü göstermektedir. Fakat YSA analizinin doğruluğu, temel olarak eğitim sürecinin doğruluğuna, yani eğitim sürecinin girdi ve çıktı arasındaki ilişkiyi ne kadar doğru temsil edebileceğine bağlıdır. Ayrıca analizin doğruluğu, niteliklerin birbirleriyle

ilişkisinden ve Q matris yapısından etkilendiği gibi ağ yapısı, ağın öğrenme hızı ve önceden tanımlanmış fonksiyonlardan da çok etkilenir (Shu vd., 2013). Sinir ağı eğitildikten sonra, gözlemlenen her cevap örüntüsü için nitelik olasılıkları hesaplanabilir. Böylece her bir öğrenciye kendisinin genel yeterlik seviyesi veya her bir niteliğe ne seviyede sahip olduğu hakkında özel bilgiler verilebilir (Gierl, 2007).

Mevcut araştırma kapsamında YSA'lerden sadece ÇKA-YSA ile çalışılmıştır. Bu nedenle çalışmanın kalan bölümlerinde ÇKA-YSA yerine YSA kullanılmıştır. Ayrıca araştırmanın ikinci bölümünde POBT, YSA ve DINO modeline ait nitelik ve örüntü düzeyi sınıflandırma oranlarının benzerliği gerçek veri setlerinde incelenmiştir. Mevcut araştırma kapsamında gerçek veri setleri için PISA 2015 İPÇ yeterliği verileri kullanıldığı için bu bölümde PISA 2015 İPÇ yeterliği hakkında bilgilere yer vermenin yararlı olduğu düşünülmektedir.

2.1.4. PISA 2015 İPÇ Yeterliği

Hızlı değişen iş yaşamında önceleri bireylerin çalışma alanlarında çözümü açık olmayan problemleri bireysel olarak çözmeleri beklenirken günümüzde bu tür problemleri işbirliği içerisinde (İPÇ) çözmeleri beklenmektedir (OECD, 2017b). Bireylere ait İPÇ yeterliklerinin yüksek olması beklentisi, bu yeterliğin ölçülmesi ve değerlendirilmesine sebep olmuştur (OECD, 2017a). İPÇ yeterliğinin güvenilir bir şekilde değerlendirilmesi, uluslararası değerlendirmelerde ilk defa PISA 2015'te ele alınmıştır (Herborn, Stadler, Mustafić, & Greiff, 2020; OECD, 2017a). Ayrıca PISA 2015, İPÇ konusundaki uluslararası karşılaştırılabilir verilerin eksikliğine de değinerek farklı eğitim sistemlerinde yetişen öğrenciler arasındaki ilişkilerin belirlenmesini sağlamaktadır. PISA 2015'e göre işbirlikçi problem çözme yeterliği bir bireyin, bir çözüme ulaşmak için gerekli olan anlayış ve çabayı paylaşarak ve bu çözüme ulaşmak için bilgi, beceri ve çabalarını bir araya getirerek, iki veya daha fazla kişinin bir sorunu çözmeye çalıştığı bir süreçte etkin bir şekilde katılma kapasitesidir.

İPÇ bireysel problem çözmeye göre etkili bir iş bölümü, bilginin çoklu bakış açıları, deneyimler ve bilgi kaynaklarından elde edilmesi, son olarak diğer grup üyelerinin fikirleri tarafından teşvik edilen yaratıcılık ve çözümlerin kalitesi konularında bazı avantajlara sahiptir (OECD, 2017a; 2017b). Bununla birlikte, işbirliği ekip üyelerine potansiyel zorluklar da getirmektedir. Örneğin iş yükü, belki de uygun olmadıkları veya sevmedikleri

görevler üzerinde çalışan ekip üyeleri ile eşit ya da verimli bir şekilde bölünmeyebilir. Bazı grup üyeleri takıma adil bir şekilde katkıda bulunmazken, diğerleri de takım hedeflerine göre kendi hedeflerine öncelik verebilir. Yaratıcı çözümler geliştirmeyi engelleyen ekip üyeleri arasında çatışma ortaya çıkabilir. Son olarak, ekip üyeleri görevleri etkin bir şekilde koordine etmeyebilir, bu da zaman kaybına ve verimlilikte düşüşe neden olabilir. Bu nedenle, başarılı bir işbirliği, tüm taraflardan uyumlu ve yapıcı bir çaba gerektirir ve kendi başına bir beceridir (OECD, 2017a). Bu noktadan hareketle bir bireyin birlikte çalıştığı grup üyelerinin özelliklerini kontrollü bir şekilde değiştirerek, bireyin ortaklaşa problem çözme yeterliğinin genel bir değerlendirmesi yapılabileceği düşünülmüş ve PISA 2015 İPÇ değerlendirmesinde insan-sanal katılımcı rolündeki bilgisayarların (agent) etkileşimi (İ-SK) kullanılmıştır (OECD, 2017a).

İ-SK yaklaşımları, insan-insan (İ-İ) doğal işbirliğinin yerine geçebilecek bir alternatif olarak kabul edilir (Herborn vd., 2020). Herborn ve diğerlerinin (2020) PISA 2015 uygulamasındaki bazı maddelerle İ-SK ve İ-İ yaklaşımlarının etkililiğini inceledikleri çalışmalarında öğrencilerin işbirliği performanslarında önemli bir farklılık bulmadıklarını belirtmişlerdir. Benzer şekilde Stadler, Herborn, Mustafa ve Greiff (2020) ve Rosen ve Tager (2013) çalışmalarında iki yaklaşım arasındaki İPÇ puan farkının önemli olmadığı, İ-SK yaklaşımının öğrencilere İPÇ becerilerini göstermeleri için daha fazla fırsat sunduğunu belirlemişlerdir. PISA 2015 uygulamasında sanal katılımcılar iletişim kurma, başkalarının eylem ve ifadelerine tepki verme, eldeki görevi iletme ve grup organizasyonunu yönetme gibi işbirlikli problem çözme ile ilgili tüm eylemleri teorik olarak yapabilmektedirler (OECD, 2017a). Sanal katılımcılar, öğrencilerin sohbet ve problem çözme eylemlerini öğrencilerin seçimleriyle çeşitlenen senaryoları kullanarak yönlendirmektedirler (Borysewicz & Steinhauer, 2014). Bununla birlikte bazen öğrencilerin yanlış başlangıç veya problem çözme tercihleri, onların ileriki maddelerdeki tercihlerini olumsuz etkileyebilmektedir. Böyle durumlarda sanal katılımcılar öğrencilere kurtarma seçenekleri oluşturmaktadırlar. Bu seçenekler öğrenciye kritik bilgiyi sormak için fırsat verme veya doğrudan gerekli bilgiyi sağlama şeklindedir (Borysewicz & Steinhauer, 2014). Farklı kültürlerde öğrenci gruplarında öğretmenlerin bulunması, yani grup üyeleri arasında statü farkının olması, grup kararlarını farklı etkileyeceğinden sanal katılımcılar arkadaş grubu olarak tanımlanmışlardır (Borysewicz & Steinhauer, 2014). Ayrıca sanal katılımcılar lojistik olarak, öğrencilerin sonuçlarının hızlı bir şekilde puanlanmasına yardımcı olurlar (OECD, 2017a).

İPÇ çerçevesi, İPÇ ile ilgili becerilerin değerlendirildiği çeşitli alanlardaki araştırma ve en iyi uygulamalara dayanan tanımları ve teorik yapıları içermektedir. Bu alanlarda bilgisayar destekli işbirlikçi çalışma, grup söylem analizi, bilgi paylaşımı, bireysel problem çözme ve örgütsel psikoloji bulunmaktadır. Çerçeve ayrıca 21. Yüzyıl Becerilerinin Değerlendirilmesi ve Öğretilmesi (ATC21: Assessment and Teaching of 21st century skills), Yetişkin Yeterliliklerinin Uluslararası Değerlendirilmesi (PIAAC; Programme for the International Assessment of Adult Competencies), 21. Yüzyıl Becerileri için Ortaklık ve PISA 2012 bireysel problem çözme değerlendirmesi için problem çözmeden elde edilen bilgileri de içermektedir (OECD, 2017b). Ayrıca İPÇ becerileri, Ulusal Değerlendirme, Standartlar ve Öğrenci Testi Araştırma Merkezi ekip çalışması işleme modeli, Salas ve meslektaşlarının ekip çalışması modeli ve ATC21'ler gibi diğer İPÇ çerçevelerinin incelenmesine dayanmaktadır (OECD, 2017a).

OECD'de (2017a, 2017b) açıklanan PISA 2015 İPÇ yeterliği bileşenleri aşağıdaki gibi özetlenmiştir:

1. Ortak anlayışın oluşturulması ve sürdürülmesi (Establishing and maintaining shared understanding): Diğer grup üyelerinin problem durumları ve bu durumlara bağlı faaliyetler hakkında bakış açılarının ve takım vizyonunun belirlenmesi ve sürdürülmesidir. Öğrenciler süreç içerisinde birbirleriyle iletişime geçerek görevler hakkında bilgilerini ve fikirlerini paylaşır, durum hakkında tartışır ve ortak görüş oluştururlar, oluşturulan ortak görüşü izleyip sürdürebilirler. Bu beceriler, öğrencinin görevi yerine getirme konusundaki kendi yeterliliklerini bilme, görevle ilişkili olarak kendi güçlü ve zayıf yönlerini tanıma (meta-hafıza) ve takımdaki sanal katılımcıların güçlü ve zayıf yönlerini (işlemsel hafıza) tanıma becerisini içermektedir.

2. Problemi çözmek için uygun eylemlerin belirlenmesi (taking appropriate action to solve the problem): Problemi çözmek için gerekli olan işbirlikli problem çözme ile ilgili eylemlerin türünün belirlenmesi ve çözüme ulaşmak için bu faaliyetlerin yürütülmesidir. Bu eylemler, problem sınırlılıklarını anlama, çözüme yönelik takım hedefleri oluşturma, görevlerde harekete geçme ve sonuçları grup ve problem hedefleriyle ilgili olarak izleme çabalarını içermektedir. Bu eylemler, karmaşık bilgilerin ve perspektiflerin aktarılması ve daha yaratıcı veya en uygun çözümler elde edilmesi için açıklamak, haklı çıkarmak, müzakere etmek ve tartışmak gibi iletişim eylemlerini de içerebilir.

3. Takım organizasyonunun oluşturulması ve sürdürülmesi (establishing and maintaining team organisation): Bireyin kendi rolünü ve diğer sanal katılımcıların rollerini anlamak,

grup üyelerinin rollerine uyup uymadıklarını takip etmek, grup organizasyonunu izlemek ve performansı en iyi hale getirmek veya problemi çözme konusundaki engelleri ortadan kaldırmak için gereken değişiklikleri kolaylaştırmak anlamına gelmektedir. Bu, geri bildirimde bulunmayı ve grup organizasyonunun problem çözmedeki başarısını yansıtmayı içerir.

2.2. İlgili Araştırmalar

Bu bölümde parametrik olmayan bilişsel tanılama yöntemleri ve PISA 2015 İPÇ yeterliğiyle ilgili çalışmalar farklı alt başlıklarda incelenmiştir.

2.2.1. Parametrik Olmayan Bilişsel Tanılama Yöntemleri ile İlgili Araştırmalar

Bu başlık altında parametrik olmayan yöntemlerin sınıflandırma performanslarının karşılaştırıldığı çalışmalara yer verilmiştir.

BTM'ler küçük örneklemin, az sayıda maddenin veya çok fazla niteliğin bulunduğu bazı durumlarda yanlış sonuçlar verebilmektedir (Shuying, 2016). BTM'nin bu gibi sınırlamalarından dolayı bilişsel tanılama alanyazınında parametrik olmayan teknikler kullanılmaktadır. Bilişsel tanılamayla ilgili alanyazın incelendiğinde çeşitli kümeleme yöntemlerinin öğrencileri nitelik profillerine sınıflandırmak için parametrik olmayan yöntemler olarak kullanıldıkları görülmüştür. Açıklayıcı bakış açısına sahip kümeleme yöntemleri niteliklere ait toplam puanlar elde edildikten sonra öğrencileri sınıflandırmada çeşitli uzaklık ölçülerinin ve bakış açılarının kullanılmasıyla birbirlerinden farklılaşmaktadırlar. Örneğin Willse vd. (2007) azaltılmış RUM (RED-RUM; Reduced RUM) temelli bilişsel tanı verilerine k-ortalama kümeleme (k-means clustering) uygulamışlardır. Ayers vd. (2008) ise öğrencilerin sınıflandırılmasında çeşitli yaygın kümeleme (capability vectors) yöntemlerinin performansını test etmişlerdir. Chiu, Douglas ve Li (2009) ve Köhn, Chiu ve Brusco (2015) ise sınıflandırma için Asimtotik Teori'nin DINA temelli veride k-ortalama ve hiyerarşik aglomeratif küme analizi (hierarchical agglomerative cluster analysis; HACA) için uygunluğunu ispat etmişlerdir. Chiu ve Köhn (2015), Asimtotik Teori'nin DINO temelli veride HACA için uygunluğunu kanıtladıktan sonra araştırmacılar, 2016 yılında Asimtotik Teori'nin RUM ve Genelleştirilmiş Bilişsel Tanı Modeli (GBTM) için de uygun olduğunu göstermişlerdir.

Alanyazında öğrencileri nitelik profillerine sınıflandırmada kullanılan parametrik olmayan bir diğer yöntem Denetimli Yapay Sinir Ağları'dır. İlgili çalışmalarda araştırmacıların Denetimli Yapay Sinir Ağları'ndan ÇKA-YSA'lar ile çalıştıkları görülmektedir. ÇKA-YSA'lar BTM kapsamında ilk olarak Gierl ve diğerlerinin 2007 yılında sundukları bildiride Nitelik Hiyerarşisi Modeli (Attribute Hierarchy Model; AHM) altında küçük gruplu öğrencileri sınıflandırmak için önerilmiştir. AHM, Kural Uzayı Modeli'nin (Rule Space Model; RSM) niteliklerin hiyerarşik olarak ifade edildiği durumlara özelleştirilmiş halidir (Leighton vd., 2004). AHM, MTK temelli olması nedeniyle parametrik bir yöntem olarak kullanılırken, ÇKA-YSA'larla parametrik olmayan bir yöntem olarak da kullanılmaya başlanmıştır. Gierl (2007), ÇKA-YSA'ların AHM kapsamında kullanılmasıyla birlikte AHM ve RSM sonuçları arasında nasıl bir farklılaşma meydana gelebileceğini her iki model varsayımlarını temel alarak tartışmıştır. Cui vd. (2017) ise gerçek veri üzerinde bu modellerin sınıflandırma performanslarını karşılaştırmışlardır. Araştırmacılar çalışmalarının sonucunda RSM ve AHM çıktıları arasında %41 uyum bulmuşlardır. Buna bağlı olarak RSM ve AHM performanslarının farklı çalışmalarda da karşılaştırılmasını önermişlerdir. Gierl vd. (2008) ve Wang ve Gierl (2011) çalışmalarında, YSA'ların AHM kapsamında kullanılmalarının anlaşılması ve yaygınlaştırılması için gerçek veri üzerinde bir uygulama gösterirken, Gierl, Cui ve Zhou (2009) AHM'de nitelik güvenilirliği için yeni bir kavram sunmuşlardır. Brigg ve Circi (2017) ÇKA-YSA'ların daha güvenilir sınıflandırma yapmalarını sağlamak için ÇKA-YSA'ların eğitimlerinde karşılaşılan yerel minimum sorununa odaklanmışlardır. Araştırmacılar bu sorunun çözümü için ağırlık hesaplamalarında çoklu YSA eğitimlerinden elde edilen olasılık tahminlerinin ortalamalarının kullanılmasını önermişlerdir. BTM kapsamında ÇKA-YSA'ların analizi R gibi çeşitli yazılımlar aracılığıyla yapılmaktadır. Bu konuda Cui vd., 2010 yılında ÇKA-YSA'nın AHM analizinde kullanılması için Mathematica'yı önermişlerdir.

Shu vd. (2013), ÇKA-YSA'ları öğrencileri nitelik profillerine sınıflandırmada AHM olmadan bir sınıflandırma yöntemi olarak kullanan ilk araştırmacılarıdır. Bu araştırmacılar, DINA temelli simülasyon verisinde öğrencilerin nitelik profilini tahmin etmek için ilk olarak ÇKA-YSA'yı, daha sonra onun Ortak Maksimum Olabilirlik Kestirimi (JMLE; Joint maximum likelihood expectation) ve MCMC algoritmaları ile kombinasyonu yoluyla genişletilmiş halini uygulayıp bulguları karşılaştırmışlardır. Çalışmada ÇKA-YSA'nın MCMC ve JMLE ile birleştirilmesiyle (sırasıyla MCMC_ÇKA-YSA ve JMLE_ÇKA-YSA) ÇKA-YSA'nın sınıflandırma performansının genel olarak arttığı belirlenmiştir.

MCMC ve JMLE'nin ise küçük örnekleme sınıflandırma performanslarını ÇKA-YSA analizinin doğruluk derecesiyle orantılı olarak arttırdıkları görülmüştür. Ayrıca çalışmada ÇKA-YSA'nın MCMC ve JMLE ile birleştirilmesiyle verilere ait madde parametreleri de hesaplanabilmiştir. Çalışma bu yönüyle diğer araştırmalardan ayrılmaktadır. Alanyazında gerek Shu ve diğerlerine (2013) benzer şekilde DINA modeli üzerinde gerekse başka modeller üzerinde ÇKA-YSA'ların MCMC ve JMLE ile birleşme koşulları veya yöntemlerin performansları karşılaştırılmamış, madde parametreleri hesaplatılmamıştır. MCMC ve JMLE'nin ÇKA-YSA'larla birleştirilerek kestirilen madde parametrelerinin, öğrencileri nitelik profillerine sınıflandırmada ÇKA-YSA'ların performanslarına etkilerinin detaylı incelenmesinin önemli olduğu düşünülmektedir. Shu vd. (2013) bu bilgilere ek olarak ÇKA-YSA'ların sınıflandırma performanslarının, örneklem büyüklüğünden etkilenmediklerini, Q matris yapısı basitleştikçe ve test ile ölçülen nitelik sayısı azaldıkça da arttığını belirlemişlerdir. Ancak ÇKA-YSA'ların yüksek performans sergileyebilmeleri için bir testte en fazla kaç nitelik ölçülebileceği sorusu cevapsız kalmaktadır. Mevcut araştırma kapsamında araştırma sorularından biri bu sorunun cevaplanmasına yöneliktir.

ÇKA-YSA'ların BTM kapsamında bir sınıflandırma yöntemi olarak kullanıldığı ikinci çalışma Cui ve diğerlerine (2016) aittir. Cui vd. (2016) çalışmalarında, ÇKA-YSA'ların sınıflandırma performanslarını, Kohonen'in Öz Örgütlemeli Harita (Self-organised Map; SOM)-YSA, DINA ve DINO modelleriyle karşılaştırmışlardır. SOM-YSA'lar, giriş vektörleri arasındaki benzerliği kullanarak öğrencileri ilgili nitelik profillerine sınıflandırır (Cui vd., 2016). Açıklayıcı bakış açısına sahip SOM'ların BTM kapsamında birer sınıflandırma yöntemi olarak kullanılması bu çalışmayı alan yazıdaki diğer çalışmalardan ayıran ikinci bir özelliktir. Araştırma sonucunda, DINA temelli simülasyon verisi üzerinde birçok koşulda modellerin performansları iyiden kötüye sırasıyla DINA, SOM-YSA, ÇKA-YSA ve DINO şeklinde olduğu belirlenmiştir. DINA'ya uygun üretilen veri üzerinde DINO modelin sınıflandırma performansının düşük olması beklenen bir durumdur. Fakat parametrik olmayan yöntemlerin sınıflandırma performanslarının DINO modelden daha iyi olması, onları kullanmanın gerekliliğini ortaya koymaktadır (Cui vd., 2016). Ayrıca, Cui vd. (2016) çalışmalarında Shu vd. (2013) çalışmalarına benzer nitelik sayısı ile çalışarak benzer bulgular elde etmişler ve yine benzer önerilerde bulunmuşlardır. Son olarak araştırmacılar, ÇKA-YSA'ların telafi edici niteliklerin ölçülmesine yönelik testlerde bir sınıflandırma yöntemi olarak kullanılıp

kullanılamayacağını incelenmesi gerektiğini belirtmişlerdir. Mevcut araştırmada da bu öneri dikkate alınarak telafi edici niteliklerin ölçülmesine odaklanılmıştır.

ÇKA-YSA'ları BTM kapsamında bir sınıflandırma yöntemi olarak kullanan üçüncü çalışma ise Guo ve diğerlerine (2017) aittir. Araştırmacılar niteliklerin önkoşul ilişkisi içinde olduğu bir simülasyon durumu oluşturarak ÇKA-YSA'ların nasıl analiz edileceğini göstermişlerdir. ÇKA-YSA'yı DINA ile karşılaştıran araştırmacılar, ÇKA-YSA'nın beceri profillerine sınıflandırma performanslarının DINA modeliyle benzer, ancak beceri önkoşul ilişkilerini içeren sınıflandırmada DINA modelinden daha iyi olduğunu göstermişlerdir. Araştırmacılar ÇKA-YSA'larla ilgili daha fazla araştırma yapılması gerektiğini belirtmişlerdir.

ÇKA-YSA'ları bilişsel tanılama kapsamında bir sınıflandırma yöntemi olarak kullanan yukarıdaki çalışmalar incelendiğinde araştırmacıların tümünün parametrik olmayan yöntem kullandıkları, fakat parametrik olmayan yöntemlerin küçük veri gruplarının analizlerinde sahip oldukları avantajları sadece Shu ve diğerlerinin (2013) test ettikleri görülmektedir. Mevcut araştırma kapsamında örneklem koşulları 30, 100 ve 500 olarak seçilip ÇKA-YSA'nın küçük örneklemlerdeki sınıflandırma performansına yönelik avantajı test edilmesi amaçlanmıştır.

Alanyazında öğrencileri nitelik profillerine sınıflandırmada kullanılan son parametrik olmayan yöntem POBT'dir. Chiu ve Douglas (2013) ilk kez Hamming uzaklığını öğrencileri nitelik profillerine sınıflandırmada kullanarak POBT'yi geliştirmişlerdir. Araştırmacılar öncelikle DINA, DINO ve Generalized NIDA (GNIDA) modellerini temel alarak farklı koşullar altında simülasyon verileri üretmişlerdir. Simülasyon verilerinin analizi sonucunda, s ve g parametrelerinin 0,4 veya 0,5'ten küçük olmaları koşuluyla, geliştirilen parametrik olmayan yöntemlerin çeşitli durumlarda parametrik yöntemlerden daha iyi sınıflandırma performansı gösterdiklerini belirlemişlerdir. Araştırmacılar daha sonra Tatsuoaka'nın (1990) tanıttığı bölme-çıkarma testi verilerini analiz etmişlerdir. Gerçek verilerin analizi sonucunda ise Yüksek Düzey DINA (Higher-Order DINA; HODINA) ile parametrik olmayan yöntem arasında nitelik düzeyinde yüksek (0,87), örüntü düzeyinde ise düşük (0,46) benzeşme olduğunu belirlemişlerdir. Ayrıca araştırmacılar Hamming, ağırlıklandırılmış Hamming ve cezalandırılmış Hamming uzaklıklarının POBT yönteminde kullanılmalarını önermişler, analizler sonucunda ağırlıklı Hamming uzaklığının daha hassas ve tutarlı sonuçlar sunduğunu belirlemişlerdir. Chiu ve Douglas (2013), POBT yönteminde farklı uzaklık ölçülerinin de kullanılmasını önerse de incelenen alanyazında

farklı uzaklık ölçüleri deneyen araştırmacılarla karşılaşmamıştır. Alanyazında Wang ve Douglas (2015), Lim ve Drasgow (2017) ve Chiu ve diğerlerinin (2017) çalışmalarında Hamming uzaklığını, Akbay (2016) ve Paulsen'ın (2019) ağırlıklandırılmış Hamming uzaklığını kullandıkları görülmektedir. Chiu ve Douglas (2013) çalışmalarının sonucunda POBT sayesinde küçük ve orta ölçekli test programlarında bilişsel tanının yapılabileceğini öne sürerken, öne sürdükleri yöntemin hangi koşullarda daha etkili olduğunun belirlenmesi için daha fazla araştırmanın yapılması gerektiğini vurgulamışlardır.

Wang ve Douglas (2015) çalışmalarında, Chiu ve Douglas'ın (2013) geliştirdikleri yönteme ait tutarlık teoremini açıkladıktan sonra, yöntemin DINA, Noisy Inputs, Deterministic “And” Gate (NIDA) ve RED-RUM temelli simülasyon verilerinde hangi koşullar altında tutarlı sonuçlar verdiğini araştırmışlardır. Ayrıca araştırmacılar tutarlığın sadece bu modellerle sınırlı kalmadığını, diğer modellerle de çalışılması gerektiğini belirtmişlerdir. Wang ve Douglas (2015), genel olarak niteliğe sahip öğrencilerin maddeye doğru cevabı verme olasılıklarının 0,5'ten yüksek, niteliğe sahip olmayan öğrencilerin ise maddeye doğru cevabı verme olasılıklarının 0,5'ten düşük olması şartı sağlanıyorsa madde sayısı arttıkça parametrik olmayan sınıflandırmanın tutarlı olacağını belirlemişlerdir. Wang ve Douglas (2015) ile benzer şekilde Lim ve Drasgow (2017) da Hamming uzaklığına dayalı parametrik olmayan yöntemi açıklamışlar, daha sonra yöntemin hangi koşullar altında tutarlı sonuçlar verdiğini incelemişlerdir. Ayrıca araştırmacılar daha önce kalibre edilmiş bir dizi maddenin Q matrisi oluşturulduktan sonra, yeni maddelerin Q matrislerini kalibre etmek için doğru ve otomatik bir istatistiksel yöntem sunmuşlardır. Araştırmacılar POBT tutarlığının belirlenmesi için öncelikle DINA, RUM, DINO ve Compensatory Reparameterized Unified Model (C-RUM), temelli simülasyon verilerini, daha sonra Tatsuoka'nın (1990) bölme-çıkarma testinden elde edilen verileri analiz etmişlerdir. Analizler sonucunda POBT'nin çok genel bir bilişsel tanı modeli sınıfı altında tutarlı olduğunu göstermişlerdir. Ayrıca yöntemin sınıflandırma doğruluğunun asimtotik teori ile tutarlı ve örneklem büyüklüğü ile orantılı olarak arttığını fakat nitelik başına düşen madde sayısının az olduğu durumlarda azaldığını belirlemişlerdir. Bunlara ek olarak nitelikler arası ilişki daha az olduğunda, sınıflandırma oranlarının daha tutarlı ve daha yüksek olduğunu belirlemişlerdir.

Akbay (2016), parametrik olmayan durumlarda tanısallık değerlendirme amaçlı yöntem seçiminde yardımcı olması için DINA, DINA-additive CDM (DINA-ACDM), DINO-additive CDM (DINO-ACDM) ve DINO temelli simülasyon verilerini analiz etmiştir.

Araştırmacı çeşitli koşullarda BTM'lere kıyasla parametrik olmayan yaklaşımın görece verimliliğini incelemiştir. Ayrıca, araştırmacı bu karşılaştırmalarda, parametrik olmayan yaklaşımdan ziyade bir BTM'nin rahatça kullanılması için örneklem büyüklüğünün ne kadar olması gerektiği sorusuna da cevap aramıştır. Karşılaştırmalar sonucunda, örneğin, 20 ayırt ediciliği yüksek madde altı niteliği ölçtüğünde Deneysel Bayes yönteminin POBT kadar doğru sınıflandırma oranına sahip olması için örneklem büyüklüğünün en az 500 olması gerektiğini belirlemiştir. Maddeler nispeten daha düşük ayırt edicilikte olduğunda bu sayı 1000'e yükselmektedir. Sonuç olarak, örneklem büyüklüğü 1000'den küçük olduğunda, parametrik olmayan yaklaşım genel olarak daha yüksek doğru sınıflandırma oranları sağlamaktadır. Son olarak, çalışmada madde başına belirtilen nitelik sayısının etkisi ortaya çıkarılmaya çalışılmıştır. Bunun için oluşturulan basit ve karmaşık yapıları iki Q matrisinin analizi sonucunda, POBT'nin basit yapıları Q matrisi verisinde daha yüksek sınıflandırma doğruluğuna sahip olduğu belirlenmiştir. Bu sonuç, öğrencilerin güçlü ve zayıf yönlerinin belirlenmesi için her bir niteliğin mümkün olduğu kadar çok ölçülmesinin önemli olmasına rağmen, birçok niteliğin aynı anda ölçülmesinin istenmeyen sonuçlara yol açabileceğini düşündürmektedir (Akbaş, 2016).

Alanyazında Chiu ve Douglas'ın (2013) önerdiği POBT'nin bir uzantısı olarak farklı yöntemlerin de geliştirildiği görülmektedir. Örneğin Chiu vd. (2017) genel parametrik olmayan bir sınıflandırma (general nonparametric classification-GPOBT) yöntemi önermişlerdir. GPOBT yöntemi karmaşık modellere uygun verileri analiz etmek için POBT yönteminden daha pratiktir ve herhangi bir verinin analizi için veri yapısının (telafisel veya telafisel olmayan) bilinmesine gerek yoktur (Chiu vd., 2017). Generalized DINA (G-DINA) ve DINA temelli simülasyon verilerinin ve gerçek verinin (Tatsuoka'nın (1990) bölme-çıkarma testinden elde edilen) analizi sonucunda Chiu vd. (2017), GPOBT yönteminin örneklem küçük olduğunda Genelleştirilmiş BTM'lerden (örn. G-DINA) daha iyi performansa sahip olduklarını belirlemişlerdir. Araştırmacılar GPOBT yönteminin bir sınırlılığı olarak yöntemin sınıflandırma tutarlılığının bilinmemesini göstermişlerdir. Chiu ve Köhn (2019) çalışmalarında GPOBT'nin istatistiksel olarak tutarlı olduğunu kanıtlayarak bu sınırlılığı ortadan kaldırmışlardır. Alanyazında POBT'nin farklı bir uzantısını kullanan Shuying (2016) çalışmada testlet etkisini değerlendirmek için bir testlet Hamming uzaklığı yöntemini ve koşullu korelasyon tabanlı parametrik olmayan bir yaklaşımı önermiştir. Buradan hareketle Lim ve Drasgow (2019) niteliklerin yanlış

tanımlanmasını ve POBT yöntemlerindeki testlet etkilerini tespit etmek için Mantel-Haenszel ki-kare istatistiğini bir indeks olarak önermişlerdir.

POBT'nin kullanıldığı çalışmalar (Akday, 2016; Chiu & Douglas, 2013; Lim & Drasgow, 2017; Wang & Douglas, 2015) incelendiğinde ÇKA-YSA'lara benzer bir durumla karşılaşmaktadır. Yani, çalışmalarında POBT'yi kullanan araştırmacıların tümünün parametrik olmayan yöntem kullandıkları, fakat parametrik olmayan yöntemlerin küçük veri gruplarının analizlerinde sahip oldukları avantajı test etmedikleri görülmektedir. Araştırma kapsamında örneklem koşulları 30, 100 ve 500 olarak seçilip POBT'nin küçük örneklemdeki sınıflandırma performansına yönelik avantajı test edilmiştir. Son olarak, Akday (2016), Chiu ve Douglas (2013) ve Lim ve Drasgow (2017) çalışmalarında POBT'nin sınıflandırma performansını nitelik ve örüntü düzeyi uyum indeksi ile incelerken, Wang ve Douglas (2015) sadece örüntü düzeyi uyum indeksini kullanmışlardır. Mevcut araştırma, POBT'nin sınıflandırma performansını nitelik ve örüntü düzeyinde incelemesi yönüyle diğer çalışmalarla benzerlik göstermektedir.

Mevcut araştırmada temel amaç, bilişsel tanılamada parametrik olmayan yöntemlerin çeşitli koşullar altında sınıflandırma performanslarını incelemektir. Bu amaç doğrultusunda nitelikler arası önkoşul ilişkilerinin olmadığı durumlarda kullanılan ve doğrulayıcı bakış açısına sahip yöntemlere, yani ÇKA-YSA ve POBT'ye yer verilmiştir. BTM'lerle analiz için veri yapısının (telafisel veya telafisel olmayan) ve Q matrisinin bilinmesi hususunda benzerlik gösteren bu yöntemler, öğrencileri nitelik profillerine sınıflandırmada farklı bakış açıları kullanmaları yönüyle birbirlerinden farklıdır. Alanyazında ÇKA-YSA'lar ile POBT'lerin sınıflandırma performanslarını bir arada inceleyen iki çalışmanın olduğu görülmüştür. Bu çalışmalardan ilki McCoy ve Willse'e (2014) aittir. Araştırmacılar DINA temelli simülasyon verisi üzerinde Hamming ve ağırlıklandırılmış Hamming gibi iki farklı uzaklığın kullanıldığı POBT, ÇKA-YSA ve MLE kestirime dayalı DINA'nın sınıflandırma performanslarını çeşitli koşullar altında (madde ve nitelik sayısı, madde ayırteediciliği, örneklem büyüklüğü, Q matris yapısı, nitelikler arasındaki korelasyon, vb.) karşılaştırmışlardır. Araştırmacılar madde ayırt ediciliğinin her koşulunda POBT'lerin ÇKA-YSA'dan daha iyi sınıflandırma performansı gösterdiklerini belirlemişlerdir. Ayrıca POBT'ler madde ayırtediciliğinin düşük olduğu durumlarda DINA'dan az bir farkla daha iyi, ayırtediciliğinin yüksek olduğu durumlarda ise MLE ile kıyaslanabilecek bir performans gösterdiğini belirlemişlerdir. Son olarak, ÇKA-YSA'ların performanslarının POBT'lerle ancak kısa test, birbirleri arasında düşük korelasyon olan çok sayıda nitelik ve

basit bir Q matrisi yapısı söz konusu olduğunda karşılaştırılabilir sonuçlar verdiğini belirlemişlerdir.

POBT'lerle ÇKA-YSA'ların sınıflandırma performanslarını karşılaştıran ikinci çalışma ise Paulsen'e (2019) aittir. BTM'leri küçük örneklem koşulunda ele alan Paulsen (2019), DINA temelli simülasyon verisinde örneklem büyüklüğünü, Q matris yapısını (normal ve hiyerarşik), Q matrisin yanlış belirlenme oranını ve madde ayırt ediciliği koşullarında yöntemlerin performanslarını karşılaştırmıştır. Araştırmacı, her durumda DINA ve POBT modellerinin birbirleriyle karşılaştırılabilir ve ÇKA-YSA modelinden daha yüksek sınıflandırma oranlarına sahip olduklarını belirlemiştir. Bu nedenle Paulsen (2019), eğer koşullar POBT yapmak için uygunsa ÇKA-YSA'ların kullanılmamasını önermiştir. Yapılan ÇKA-YSA çalışmalarında DINO model temelli simülasyon ve gerçek veri setlerinin kullanılmadığı belirlenmiştir. Paulsen'in (2019) ÇKA-YSA'ların kullanımıyla ilgili önerisinin bu veri setlerinde geçerli olup olmadığının incelenmesinin gerekli ve önemli olduğu düşünülmektedir. Ayrıca, Paulsen (2019) sınıf içi değerlendirmelerde kullanılabilecek yöntemlerin hangi koşullarda daha yüksek sınıflandırma oranlarına sahip olduklarının belirlenebilmesi için daha fazla araştırmaya ihtiyaç olduğunu vurgulamıştır.

Mevcut araştırma, McCoy ve Willse (2014) ve Paulsen (2019) çalışmalarıyla bazı benzer ve farklı yönleri sahiptir. Mevcut araştırmanın diğer çalışmalarla benzer yönlerini niteliklerin ikili kodlanması (0-1), simülasyon verilerinde niteliklerin çok değişkenli normal dağılması ve sınıflandırma doğruluğu indeksleri olarak nitelik ve örüntü düzeyi sınıflandırma oranlarının raporlanması şeklinde özetlemek mümkündür. Hem gerçek veri hem de DINO temelli simülasyon verisinin kullanılması, simülasyon ve gerçek veri yapısına bağlı olarak telafisel modellerin sınıflandırma performanslarının incelenmesi ise mevcut araştırmanın diğer çalışmalardan farklı yönlerine örnektir. McCoy ve Willse (2014) ve Paulsen'in (2019) çalışmalarında en fazla nitelik sayısı sekiz iken mevcut araştırmanın gerçek veri analizi koşullarında nitelik sayısı 11'dir. McCoy ve Willse (2014) ve Paulsen'in (2019) çalışmalarında madde ayırt edicilikleri yüksek veya düşük olarak ele alınıp incelenmiştir. Mevcut araştırmada ise önceki çalışmaların tavsiyesi üzerine madde ayırt edicilikleri karışık (orta) düzeyde ele alınmıştır.

Son olarak, bilişsel tanılama alanında yapılan çalışmaların çok azının kayıp veri örüntü türleri ve kayıp veri ile baş etme yollarının öğrencileri nitelik sınıflarına yerleştirmedeki etkileri üzerine odaklandığı görülmektedir. Kayıp veri ile ilgili yapılan çalışmalar incelendiğinde ise çoğunun (Başokçu, Kalkan & Öğretmen, 2016; Dai, 2017; Sünbül,

2018; Kalkan & Başokçu, 2019; Xu & von Davier, 2006) kayıp veri etkisini parametrik modeller üzerinde incelediği, sadece Ayers ve diğerlerinin (2009) bu etkiyi kümeleme (capability vectors) yöntemi üzerinde incelediği görülmektedir. Mevcut araştırma kayıp veri etkisini parametrik olmayan bilişsel tanılama alanında ve POBT ve YSA'nın sınıflandırma oranlarına etkisini incelemesi yönüyle de diğer çalışmalardan farklılık göstermektedir.

2.2.2. PISA 2015 İPÇ Yeterliği ile İlgili Araştırmalar

Yapılan alanyazın taramasında PISA 2015 İPÇ yeterliğini oluşturan becerilerin incelendiği az sayıda çalışmaya rastlanmıştır. Ek olarak PISA 2015 İPÇ ile ilgili çalışmaların genele olarak üç gruba ayrıldığı görülmektedir. Birinci grup çalışmalar PISA 2015 İPÇ değerlendirmesindeki ölçeklerin analizi ve görevlerin geçerlik kanıtlarının elde edilmesiyle ilgili iken ikinci grup çalışmalar PISA 2015 İPÇ çerçevesini temel alarak öğrencilerin İPÇ yeterliklerini belirlemekte, üçüncü gruptaki çalışmalar ise PISA 2015 İPÇ verilerini kullanmaktadırlar.

PISA 2015 İPÇ değerlendirmesindeki yeni ölçeklerin analizi için çoklu grup Rasch modeli, kısmi kredi modeli (PCM, partial credit model), çok gruplu iki parametrelili lojistik model (2PL) ve genelleştirilmiş kısmi kredi modeli (GPCM, generalized partial credit model) gibi ölçekleme modelleri kullanılmıştır (He, von Davier, Greiff, Steinhauer, & Borysewicz, 2017). Uygulamadaki her bir madde ya İPÇ becerilerinin göstergesi olarak ya da değerlendirme birimlerinin geliştiricileri tarafından sağlanan puanlama rehberine dayanılarak puanlanmıştır. Nouri, Åkerfeldt, Fors ve Selander (2017) ise yaptıkları çalışmada PISA 2015 İPÇ değerlendirme çerçevesinin yordama geçerliğini ve sanal katılımcılı bilgisayar temelli görevlerin 24 öğrencinin İPÇ performanslarına etkisini incelemişlerdir. Elde edilen bulgular PISA 2015 İPÇ değerlendirme çerçevesinin zayıf bir yordama geçerliğine sahip olduğunu, iletişimde kalite veya üretkenliği hesaba katmadığını ve sanal katılımcıların gerçekten CPS süreçlerini ve dolayısıyla neyin değerlendirilebileceğini etkilediğini göstermektedir. Stadler, Herborn, Mustafić ve Greiff (2020) de sanal katılımcılarla bilgisayar temelli yapılan PISA 2015 İPÇ uygulamasındaki görevlerin yapı ve yordama geçerliklerini incelemişlerdir. Bunun için orijinal PISA 2015 İPÇ görevlerini OECD'den alarak görevlerin öz, öğretmen ve yardımcı oyuncu tarafından değerlendirilen işbirliği ile ne ölçüde ilişkili olduğunu belirlemişlerdir. Analizler sonucunda, öğrencilerin kendilerini değerlendirdikleri işbirliği puanlarının ve öğretmen

tarafından derecelendirilen işbirliği puanlarının PISA 2015 İPÇ performansı ile orta düzeyde ilişkili olduğunu belirlemişlerdir. Bu nedenle Stadler vd. (2020), sanal katılımcılı bilgisayar destekli PISA 2015 İPÇ görevlerinin yapı geçerliğine sahip olduklarını ifade etmişlerdir. Sonraki aşamada Stadler vd. (2020), yardımcı oyuncu işbirliği puanının grubun İPÇ toplam puanıyla orta düzeyde ilişkili olduğunu, bu durumda grup puanlarının İPÇ yeterliği en yüksek öğrenciyi temsil etmesinin dışında grubun tüm üyelerinin yeterliklerini temsil ettiğini, buradan PISA 2015 İPÇ görevlerinin yordama geçerliğine sahip olduklarını ifade etmişlerdir. Ek olarak, araştırmacılar yapı ve yordama geçerliği kanıtları elde edilmiş görevlerin sanal katılımcılı bilgisayar destekli PISA 2015 İPÇ uygulamasının geçerliğine ait birer kanıt olduğu belirtmişlerdir. Nouri vd. (2017) ve Stadler ve diğerlerinin (2020) elde ettikleri sonuçlar çelişkili gibi görünse de bu durum PISA 2015 İPÇ görevlerinin ve diğer İPÇ değerlendirmesinde kullanılan görevlerin farklı türden geçerliklerini farklı yöntemlerle inceleyen daha fazla çalışmaya ihtiyaç olduğunu göstermektedir.

PISA 2015 İPÇ ile ilgili ikinci grup çalışmalar ise PISA 2015 İPÇ çerçevesini temel alarak çeşitli alanlarda (fizik, geometri, vb.) öğrencilerin İPÇ yeterliklerini belirlemeye çalışmaktadır. Bu gruptaki çalışmaları Chang vd. (2017), Haung (2017), Herborn, Mustafic ve Greiff (2017), Kuo vd. (2019), Nordin ve Osman (2018), Özkan ve Öner (2019) ve Wang'ın (2015) çalışmaları oluşturmaktadır. Öğrencilerin PISA 2015 İPÇ yeterliklerini belirlemede Kuo ve diğerlerinde (2019) çok boyutlu MTK, Herborn ve diğerlerinde (2017) kümeleme analizi kullanılırken diğer çalışmalarda öğrencilerin davranışları gözlemlenerek veya internet ortamında çeşitli platformlarda oluşturulan sohbet kutularındaki yazışmaların nitel analizi yapılarak değerlendirilmektedir.

Son olarak PISA 2015 İPÇ ile ilgili üçüncü gruptaki çalışmaları de Boeck ve Scalise (2019), Elkonca, Karakaya, Şata ve Ceyhan (2018), Şata, Karaduman ve Karakaya (2018), Leong (2020), Li ve Liu (2017), Lim ve Lee (2018), Stadler, Herborn, Mustafic ve Greiff (2018), Şata, Çorbacı ve Karakaya (2018) ve Yavuz ve Atar'ın (2020) çalışmaları oluşturmaktadır. Leong (2020), PISA 2015 İPÇ uygulamasının Singapur için ölçme değişmezliğini sağladığını belirlemiştir. Fakat Şata, Çorbacı ve Karakaya (2018) çalışmalarında İPÇ uygulamasından “Xandar” isimli birimin bazı maddelerinin Türkiye örnekleme için cinsiyete göre yanlı olduğunu belirtmişlerdir. Stadler vd. (2018), beş büyük kişilik özelliklerinin PISA 2015 İPÇ yeterliği üzerindeki etkilerini incelemişlerdir. Yapısal eşitlik modellemesinden elde edilen sonuçlar, deneyime açıklığın ve uyumlu olmanın İPÇ performansının iyi birer yordayıcıları olduğunu göstermiştir. Benzer şekilde Şata,

Karaduman ve Karakaya (2018) Türk öğrencilerin cinsiyetlerinin, sosyo ekonomik ve kültürel durumlarının, motivasyonlarının ve işbirliğine karşı tutumlarının onların PISA 2015 İPÇ başarılarına etkilerini incelemişlerdir. Araştırmacılar tüm değişkenlerin öğrencilerin başarıları üzerinde etkili olduğunu, en yüksek etkiye ise sosyo ekonomik ve kültürel durum indeksinin sahip olduğunu belirtmişlerdir. de Boeck ve Scalise (2019), PISA 2015 İPÇ uygulamasındaki Xandar görevinin dört bölümünü kullanarak problem çözme performansının harcanan zaman ve işbirliği bölümlerindeki eylem sayısı ile ilişkisini araştırmışlardır. Görevin her bölümünde gerçekleştirilen eylemler PISA tarafından belirlenmiş ve sayısal olarak veri setine kaydedilmiştir. Doğrulayıcı faktör analizi sonucunda daha fazla eyleme sahip katılımcıların daha az başarılı ve daha hızlı oldukları, başarılı katılımcıların ise görevleri yapmak için daha fazla zaman harcadıkları belirlenmiştir. Lim ve Lee (2018) ise PISA 2015 İPÇ puanları üzerinde okul ve öğrenci düzeyinden oluşan iki düzeyli HLM analizi gerçekleştirmişlerdir. HLM analizi sonucunda öğrenci düzeyinde cinsiyetin, sosyoekonomik/kültürel geçmişin, işbirliği düzeyinin İPÇ puanlarını pozitif yönde, öğretmenin adaletsizlik algısının ise olumsuz yönde yordadığını göstermiştir. Okul düzeyinde ise öğretmenlerin olumlu bir çalışma ortamı algılamaları ve öğrenci başına düşen bilgisayarın az olmasının İPÇ puanlarını pozitif yönde yordadığını belirlenmişlerdir.

PISA 2015 İPÇ verileriyle yapılan çalışmalardan sadece Li ve Liu (2017) ve Yavuz ve Atar (2020)'ın çalışmalarında PISA 2015 İPÇ yeterliğini oluşturan becerilerin ayrı ayrı incelendiği görülmektedir. Li ve Liu (2017), yaptıkları çalışmada İPÇ yeterliğini 11 beceri matrisi şeklinde ele alıp incelemişlerdir. Bunun için her bir becerinin ortalamaları elde edilerek karşılaştırılmıştır. Yavuz ve Atar (2020) ise öncelikle İPÇ yeterliğini 3, 7 ve 11 alt beceri olarak telafisel modellerden DINO model ile incelemişlerdir. Araştırmacılar model uyumu indekslerini inceleyerek PISA İPÇ yeterliğinin üç alt yeterlik olarak incelenmesinin daha uygun olduğuna karar vermiş ardından Türkiye örnekleminde öğrencilerin PISA 2015 İPÇ yeterliklerini yorumlamışlardır. Mevcut araştırma, PISA 2015 İPÇ verilerinin POBT ve YSA ile analiz edilmesi yönüyle diğer araştırmalardan farklılaşmaktadır.

BÖLÜM III

YÖNTEM

Bu bölümde araştırmanın modeli, çalışma grubu, veri toplama aracı ve verilerin elde edilmesi, son olarak verilerin analizi hakkında bilgilere yer verilmiştir.

3.1. Araştırmanın Modeli

Bu araştırmada POBT, YSA ve DINO modelinin farklı koşullar altında sınıflandırma performanslarının belirlenmesi amaçladığından araştırma koşullarının oluşturulması için araştırmanın ilk bölümünde simülasyon veri setleri üretilmiştir. Bu nedenle araştırmanın ilk bölümü simülasyon çalışmasıdır. Erkuş (2013), simülasyon çalışmalarının korelasyonel, deneysel ve de betimsel olabileceğini belirtmiştir. Ek olarak, bu araştırmanın simülasyon ve gerçek veri setlerinin kullanıldığı durumlarda POBT, YSA ve DINO modelinin sınıflandırma performansları çeşitli koşullarda incelenip ardından birbirleriyle karşılaştırıldığından bu çalışma betimsel araştırma desenindedir. Betimsel araştırmalarda, üzerinde çalışılan durum olabildiğince tüm yönleriyle incelenir ve raporlanır (Fraenkel, Wallen, & Hyun, 2012, s.15).

3.2. Çalışma Grubu

Araştırmada, POBT, YSA ve DINO modelin sınıflandırma oranları öncelikle simülasyon yolu ile elde edilmiş veri setlerinde incelenmiştir. Simülasyon için gerekli bilgilere verilerin toplanması ve elde edilmesi başlığı altında yer verilmiştir. Araştırmanın ikinci bölümünde POBT, YSA ve DINO modelin sınıflandırma oranlarının benzerliği gerçek veri

setlerinde incelenmiştir. Q matrisi elde edilmesi kolaylığı nedeniyle bu karşılaştırma için PISA 2015 İPÇ uygulaması seçilmiştir. İPÇ değerlendirmesine 52 ülkeden (32 OECD ülkesi ve 20 ortak ülke) 414498 öğrenci katılmıştır.

PISA 2015 İPÇ uygulamasında 67 form kullanılmıştır. Formların madde-nitelik dağılımları ve temsil ettikleri İPÇ becerilerinin sayısı incelenmiştir. Form 93 ve Form 96 İPÇ yapısını temsil eden daha fazla maddeye sahip olduğundan, Form 93 ve Form 96 ile çalışılmasına karar verilmiştir. Bu formlar aynı maddeleri içermektedir ve sadece maddelerin formlardaki sırası değişmektedir. Bu nedenle formlardaki tüm veriler birleştirilmiştir. Ayrıca, kapsam geçerliliğinin azalmaması için çok kategorili olarak puanlanan maddeler, araştırma kapsamından çıkarılmadan ikili olarak tekrar puanlanmıştır. Maddeleri ikili olarak puanlamak için seçenek dağılımları incelenmiştir. Dağılımın ilk yarısında kalan seçenekler "0", diğer seçenekler "1" olarak puanlanmıştır. Form93 ve Form 96'yı toplamda 43 ülkeden 30053 öğrenci cevaplamıştır. Kayıp verilerin ayıklanması ile 18170 öğrenciye ait veri kalmıştır. Sonuç olarak, araştırmanın ikinci bölümü için çalışma grubu, PISA 2015 İPÇ uygulamasına katılan 18170 öğrenciden seçilmiştir.

Çalışma grubunun belirlenmesi için basit seçkisiz örnekleme (simple random sampling) yöntemi kullanılmıştır. Basit seçkisiz örneklemede evreni oluşturan her bir üye eşit ve bağımsız seçilme şansına sahiptir (Fraenkel, Wallen, & Hyun, 2012). IBM SPSS 26.0 kullanılarak 18170 öğrenciden 30, 100 ve 500 kişilik örneklem büyüklüğü koşulları belirlenmiştir. Çalışma grubunu oluşturan ülkelere ve bu ülkelerin öğrencilerine ait bilgiler EK-1'de verilmiştir.

3.3. Veri Toplama Aracı ve Verilerin Elde Edilmesi

Araştırma iki bölümden oluşmaktadır. İlk bölümde simülasyon veri setlerinde POBT, YSA ve DINO modelinin sınıflandırma oranları incelenmiştir. İkinci bölümde ise gerçek veri (PISA 2015 İPÇ) setlerinde POBT, YSA ve DINO modelin sınıflandırma oranlarının benzerliği incelenmiştir. Bu nedenle simülasyon ve gerçek veri setleriyle ilgili bilgiler ayrı alt başlıklarda açıklanmıştır.

3.3.1. Simülasyon Veri Setlerinin Elde Edilmesi

Simülasyon veri setleri, gerçek veri setlerinin yapısıyla uyumlu olması için telafisel modellerden sık kullanılan DINO modeline göre üretilmiştir. Veri üretimine Q matris yapılarının belirlenmesiyle başlanmıştır.

Q matrislerin oluşturulmasında her bir niteliği ölçen aynı sayıda madde olmasına ve birden fazla niteliği ölçen maddelerin eşit sayıda olmasına dikkat edilmiştir (bkz. de la Torre, 2008; de la Torre & Douglas, 2008; de la Torre vd., 2010; Rupp & Templin, 2008a). Nitelik sayısı yedi olduğunda $2^7 = 128$ olası nitelik kombinasyonu vardır ve hepsini tek bir teste dâhil etmek pratik değildir. Bu nedenle her bir Q matrisinde en fazla üç niteliğin bir madde ile temsil edilmesine izin verilmiştir. Araştırma kapsamında 15 madde ile oluşturulan tüm Q matrisleri, yedi madde bir niteliği, altı madde iki niteliği ve iki madde üç niteliği temsil edecek şekilde tasarlanmıştır. Ayrıca araştırma kapsamında kullanılan tüm Q matrislerinde maddelerin temsil ettikleri nitelik sayıları sabittir. Örneğin üç, beş ve yedi nitelikli Q matrislerinde altı maddenin her biri iki niteliği temsil etmektedir. Buradan hareketle Q matrislerinde sadece nitelik sayısındaki artışın, modellerin sınıflandırma performansına etkisi incelenmek istenmiştir. Bu bağlamda araştırma kapsamında Q matris yapısının karmaşıklığının artması, nitelik sayısındaki artışı ifade etmektedir. Son olarak, tüm Q matrisleri tamamlanmıştır (Chiu vd., 2009; Köhn & Chiu, 2017). Başka bir ifadeyle tüm Q matrisleri olası tüm gizil sınıfları ayırt edebilmektedir. Araştırma kapsamında 15 maddelik, nitelik sayısı (3, 5 ve 7) koşulları için kullanılan Q matrisleri EK-2’de verilmiştir. 15 maddelik Q matrisine ait madde-nitelik dağılımları ise Tablo 2’de verilmiştir.

Tablo 2

Farklı Q Matrislerinde Her Bir Niteliği Ölçen Madde Sayısı

	1.Nitelik	2.Nitelik	3.Nitelik	4.Nitelik	5.Nitelik	6.Nitelik	7.Nitelik
3 Nitelikli Q matrisi	8	9	8				
5 Nitelikli Q matrisi	5	5	5	5	5		
7 Nitelikli Q matrisi	4	4	4	4	3	3	3

Tablo 2’de görüleceği üzere farklı nitelik sayılarında oluşturulan Q matrislerinde her bir niteliğin en az üç madde ile testte yer almasına ve her bir niteliğin testte dengeli dağılmasına dikkat edilmiştir. Araştırma kapsamında ilk olarak 15 maddelik Q matrisi oluşturulmuş daha sonra bu yapı 30 maddelik Q matrisini oluşturmak için iki kez, 45 maddelik Q matrisini oluşturmak için üç kez kullanılmıştır. Q matris yapısı belirlendikten sonra Q matrisinin sütunlarını oluşturan niteliklere ait özellikler (dağılımları ve aralarındaki ilişki düzeyi) belirlenmiştir.

Parametrik olmayan bilişsel tanılama alanyazınındaki çalışmalar incelendiğinde simülasyon verilerine ait niteliklerin tekdüze (uniform) veya çok değişkenli normal (multivariate normal, MVN) dağılıma sahip oldukları görülmektedir. Tekdüze dağılım, nitelik örüntülerinin veri içerisinde eşit olması istenildiğinde tercih edilmektedir. Araştırma kapsamında simülasyon verilerine ait niteliklerin üretilmesinde nitelik örüntülerinin eşit dağılmadığı ve nitelikler arası ilişkilerin olduğu gerçekçi durumları betimlemek için çok değişkenli normal dağılım kullanılmıştır. Nitelikler arası korelasyon değeri alanyazında sık tercih edilen (bkz. Chiu & Douglas, 2013; Chiu vd., 2017; McCoy & Willse, 2014; Wang & Douglas, 2015) $\rho = 0,5$ olarak belirlenmiştir.

Nitelik dağılımları ve nitelikler arası ilişki düzeyi belirlendikten sonra madde parametreleri (s ve g parametreleri) belirlenmiştir. Bilişsel tanılama alanında yapılan simülasyon ve ampirik çalışmalar s ve g parametrelerinin 0,5’i aştığında yeterli sınıflandırma doğruluğunun gerçekleşemeyeceğini göstermektedir (Chiu & Douglas, 2013). Ayrıca araştırma kapsamında kullanılan gerçek veri setlerinin (PISA 2015 İPÇ yeterliği) nitelik sayı 3, 7 ve 11 olduğu durumlarda s ve g parametre değerleri incelenmiştir. İncelemeler sonucu ortalama s ve g parametre değerlerinin 0,33 ile 0,35 arasında olduğu belirlenmiştir. Araştırmada hem gerçek veri koşuluna yaklaşmak hem de güvenilir sınıflandırma doğruluğunu aşmamak için tüm koşullarda s ve g parametre değerleri tekdüze dağılımda $U[0; 0,4]$ olarak belirlenmiştir.

Araştırmada ele alınan faktörler nitelik sayısı, örneklem büyüklüğü, madde sayısı ve kayıp veri oranıdır. Faktörler ve koşulları, alanyazın taraması sonucunda simülasyon ve gerçek veri araştırmalarında sık kullanılmasına, konuyla ilgili uygulayıcıya ve okuyucuya zengin bilgi vermesine dikkat edilerek seçilmiştir. Böylece koşulların gerçeği daha çok yansıtması beklenmektedir. EK-3’te verilen simülasyon verilerinin yapısı ve koşulların belirlenmesinde dikkate alınan çalışmalar özetlenmiştir. Araştırma kapsamında ele alınan faktörler ve bu faktörlere ilişkin düzeyler Tablo 3’te verilmiştir.

Tablo 3

Araştırma Kapsamında Simülasyon Verisi için Ele Alınan Faktörler ve Bu Faktörlere İlişkin Düzeyler

Faktörler	Karşılaştırma Koşulları		
Nitelik sayısı (Ni)	3	5	7
Örneklem büyüklüğü (N)	30	100	500
Madde sayısı (M)	15	30	45
Kayıp veri oranı (%)	0	5	10
Modeller	POBT	YSA	DINO

Bu araştırma kapsamında simülasyon veri setlerinde $3 \times 3 \times 3 \times 3 \times 3 = 243$ karşılaştırma durumu oluşturulmuştur. Simülasyon verilerinin her bir karşılaştırma durumu için 100 replikasyon yapılmıştır. Tablo 3'te yer alan faktörlerin açıklanmasının faydalı olacağı düşünüldüğünden bu faktörler aşağıda sırasıyla bahsedilmiştir.

Nitelik sayısı faktörüne ait koşulların belirlenmesi için alanyazın incelendiğinde üç ve beş nitelik boyutlarının eğitimsel ve psikolojik değerlendirmelerde sıklıkla kullanıldığı (Kunina-Habenicht vd., 2012) görülmektedir. Bu nedenle araştırma kapsamında nitelik sayısı faktörünün düzeyleri olarak üç ve beş nitelik ele alınmıştır. Ayrıca araştırma kapsamında gerçek veri yapısı ve koşulları ile simülasyon veri yapısı ve koşulları arasında benzerlik oluşturulmak istenmektedir. Gerçek veri yapısı yedi niteliğe uygun olduğundan nitelik sayısı faktörünün düzeylerine yedi nitelik de eklenmiştir.

Örneklem büyüklüğü faktörüne ait düzeylerin belirlenmesi için EK-3 incelendiğinde parametrik olmayan bilişsel tanılama alanyazınında örneklem büyüklüğü olarak koşulların genellikle 25 ile 1000 arasında alındığı görülmektedir. Alanyazın dikkate alınarak araştırma bulgularının sınıf ve kurum içi bilişsel tanılamada kullanılabilmesi için örneklem büyüklüğü faktörünün düzeyleri 30, 100 ve 500 olarak belirlenmiştir.

Madde sayısı faktörüne ait koşulların belirlenmesi için EK-3 incelendiğinde yapılan çalışmalarda madde sayısı/nitelik sayısı oranının 1,5 ile 16 arasında değiştiği görülmektedir. Bu orandan hareketle araştırma kapsamında nitelik sayıları 3, 5 ve 7 olarak belirlendiğinden madde sayısı faktörünün düzeyleri de 15, 30 ve 45 madde olarak belirlenmiştir.

Kayıp veri oranı faktörüne ait koşulların belirlenmesi için öncelikle kayıp veri türü belirlenmiştir. Üç kayıp veri türünden bahsetmek mümkündür. Bunlar, rastgele (missing at random, MAR), tamamıyla rastgele (completely at random, MCAR) ve rastgele olmayandır (missing not at random, MNAR). Araştırma kapsamında kayıp veri türü tamamıyla rastgele olarak belirlenmiştir. Bilişsel tanılama alanyazınında kayıp veri ile baş etme yöntemlerinin sınıflandırma doğruluğu üzerindeki etkilerinin araştırıldığı çalışmalarda model temelli ve iki yönlü (two way) veri atama yöntemlerin kayıp veri ile baş etmede daha güçlü yöntemler olduğu (Başokçu vd., 2016; Dai, 2017; Sünbül, 2018) görülmektedir. Mevcut araştırmada, bazı koşullarda gözlem sayısı, parametre sayısını tahmin etmek için çok düşük olduğundan model kurulamamaktadır. Bu nedenle model temelli veri atama yöntemleri kullanılamamış, onların yerine iki yönlü veri atama yöntemi kullanılmıştır. İki yönlü (two way) veri atama, Bernaards ve Sijtsma (2000) tarafından ortaya atılmıştır. Bu yöntem aşağıdaki şekilde ifade edilir:

$$TW_{ij} = IM_j + PM_i - OM \quad (14)$$

TW_{ij} : j maddesini cevaplayan i öğrencisi için atanan değer

IM_j : j maddesine verilen cevapların ortalaması

PM_i : i öğrencisinin tüm maddelere ait ortalaması

OM : verilerin genel ortalaması

Formülден görüldüğü üzere iki yönlü atama yöntemi kişi ve madde ortalaması atama yöntemlerini birlikte kullanmaktadır.

3.3.2. Gerçek Veri Setleri İçin Veri Toplama Aracı ve Verilerin Elde Edilmesi

Bu araştırmanın ikinci bölümünde gerçek veri seti olarak PISA 2015 İPÇ yeterliği verilerinden yararlanılmıştır. Daha önce bahsedildiği üzere bilişsel tanılamada parametrik olmayan teknikler Q matrisinin elde edilmesinde kullanılamamaktadır ve karşılaştırma çalışmasına Q matrisinin oluşturulmasından kaynaklanabilecek en az hata ile devam edilmelidir. Bu nedenle mevcut araştırmada Q matrisi hakkındaki bilgilere kolay ulaşılabilindiğinden PISA 2015 İPÇ uygulaması verileri kullanılmıştır. Bu duruma ek olarak PVAF yöntemi ile Q matrisi geçerlemesi yapılarak gerçek veri setleri üzerinde analizler tekrar edilmiştir. Her iki tür Q matrislerinden elde edilen bulgular beraber yorumlanmış ve tartışılmıştır.

PISA 2015 ana uygulama öncesi yapılan alan denemesinde toplam altı birim (farklı temalara/konulara sahip madde kümeleri) oluşturulmuştur. Bu altı birimden beşi ana uygulamada kullanılmıştır (Educational Testing Service [ETS], 2016). Alan denemesinde oluşturulan bir birim (Visit başlıklı birim) ise daha sonra örnek birim olarak kişilerin erişimine açılmıştır. Bir grup uluslararası öğrencinin bir okulu ziyarete gelmesini konu alan bu örnek birim 44 maddeden oluşmaktadır ve cevaplanması ortalama 17 dakika sürmektedir. Öğrencilerin bu birimdeki maddeleri cevaplayabilmesi için okul ziyaretini planlamaları, ziyaretçileri rehberlere atamaları ve ortaya çıkan beklenmedik bir problemi çözmek için üç sanal ekip arkadaşı ve bir fakülte danışmanı ile işbirliği yapmaları gerekmektedir.

PISA 2015 İPÇ ana uygulamasında da altı birim kullanılmıştır (117 madde). Bu birimler arasındaki korelasyon 0,76-0,81 arasında değişmektedir (He vd., 2017). Her bir birim, öğrencilerin üzerinde çalıştıkları birçok küçük senaryodan oluşmaktadır. Öğrenciler birimlerde, sınıf içi bir yarışmadalar gibi olabildiğince çok soruya cevap vermek için birlikte çalışmak zorunda kalmışlardır. Birimler öğrencilere bütünlük içerisinde sunulmuş ve her biri 30 dakika süren üç ayrı kümeye ayrılmıştır, yani iki birim bir küme oluşturmuştur. İPÇ değerlendirmesine katılan tüm öğrenciler, değerlendirmelerinde bir ya da iki İPÇ kümesini cevaplamışlardır.

İPÇ değerlendirmesinde hiçbir serbest cevaplı madde bulunmamaktadır. Tüm maddeler, öğrencilerin ekip üyelerine cevap vermeleri veya simgeleri uygun alana taşınmaları ya da görsel görüntü alanındaki bir seçeneği seçmeleri için çeşitli yollardan çoktan seçmeli bir seçim yapmalarını gerektirmiştir. Etkileşimli bir değerlendirme olduğu için, öğrencilerin bir sonraki maddeye geçmeden önce her bir maddeye cevap vermeleri ve maddeleri atlamamaları istenmiştir. Bu nedenle öğrencilerin sınav esnasında bir maddeyi cevaplamadan diğerine geçişleri engellenmiştir. Öğrencilerden sadece sınav kitapçığının sonuna ulaştıklarında cevaplarını gözden geçirmeleri istenmiştir. Böylece işbirliği becerisi, öğrencilerin bir veya daha fazla sanal katılımcıyla olan etkileşimlerindeki cevapları yoluyla değerlendirilmiştir.

PISA 2015 İPÇ yeterliği, işbirliği becerilerinin ve PISA 2012 bireysel problem çözme süreç becerilerin birleşmesiyle meydana gelen bir yapıdır. PISA 2015'deki tanımı, PISA 2012 problem çözme tanımını temel almıştır, ancak bireysel ve grup bilişinin kuramsal temellerini birleştirerek işbirliği alanındaki etkisini genişletmiştir. PISA 2012 problem çözme çerçevesinin dört süreci korunmuş ve bir İPÇ becerileri matrisi oluşturmak için

belirlenen üç temel yeterlik alanı eklenmiştir. Bu becerilerin her biri değerlendirme aracı tarafından ölçülebilen yeterlilik seviyeleri ile tanımlanır. Bu beceri matrisinin yapısı Tablo 4’te verilmiştir. PISA 2015 İPÇ uygulamasındaki maddelerin doğru cevaplanması için gereken nitelikler Tablo 4’e göre belirlenmiş ve raporlanmıştır. Bu araştırmada gerçek veri setleri için Q matrisi, Tablo 4 ve raporlar (OECD, 2017a; 2017b) doğrultusunda oluşturulmuştur.

Tablo 4

PISA 2015 İPÇ Beceri Matrisi (OECD, 2017a, s.12; 2017b, s.50)

		İşbirlikli Problem Çözme Yetkinlik Alanları		
		(1) Ortak anlayışın oluşturulması ve sürdürülmesi	(2) Problemi çözmek için uygun eylemlerin belirlenmesi	(3) Takım organizasyonu oluşturulması ve sürdürülmesi
Problem çözme süreçleri	(A) Keşfetme ve anlama	(A1) Takım üyelerinin bakış açılarını ve yeteneklerini keşfetme	(A2) Hedeflerle birlikte problemi çözmek için işbirlikli etkileşimin türünü keşfetme	(A3) Problemi çözmek için rolleri anlama
	(B) Temsil ve formüle etme	(B1) Ortak bir temsilin oluşturulması ve problemin müzakere edilmesi (ortak zemin)	(B2) Tamamlanacak görevleri belirleme ve tanımlama	(B3) Roller ve takım organizasyonunu tanımlama (iletişim protokolü/katılım kuralları)
	(C) Planlama ve yürütme	(C1) Takım üyeleriyle gerçekleştirecek eylemler hakkında iletişim kurma	(C2) Planların uygulanması	(C3) Belirlenen kuralları takip etme (örneğin takım üyelerini görevlerini yapmada harekete geçirme)
	(D) İzleme ve yansıtma	(D1) Ortak anlayışı izleme ve onarma	(D2) Eylem sonuçlarını izleme ve problem çözmede başarıyı değerlendirme	(D3) İzleme, geri bildirim sağlama ve takım organizasyonu ve rolleri uyarlama

Not: OECD (2013), “PISA 2015 Draft Collaborative Problem Solving Framework” ve OECD (2017), “PISA 2015 Results (Volume V): Collaborative Problem Solving”, kaynaklarından alınmıştır.

Tablo 4 ve raporlar (OECD, 2017a; 2017b) temel alınarak PISA 2015 İPÇ yapısını temsil eden üç farklı Q matrisi oluşturulabilmektedir. PISA 2015 İPÇ uygulamasında A3 hücrelerini ölçmeye yönelik herhangi bir madde olmadığı için 12 nitelikli olarak oluşturulması gereken Q matrisi 11 nitelikli olarak oluşturulmuştur. Araştırma koşullarını oluşturan ilk 17 madde için bu üç farklı Q matrisi Tablo 5’te verilmiştir. Tablo 5’teki üç nitelikli Q matrisi, Tablo 4’teki tüm 1’ler, 2’ler ve 3’ler birer nitelik olarak kabul edilip oluşturulmuştur. Yedi nitelikli Q matrisi Tablo 4’teki tüm 1’ler, 2’ler, 3’ler, A’lar, B’ler, C’ler ve D’ler birer nitelik olarak kabul edilip oluşturulmuştur. On bir nitelikli Q matrisi Tablo 4’teki tüm hücreleri, yani A1, A2, B1, B2, B3, C1, C2, C3, D1, D2 ve D3’ü birer nitelik olarak kabul edilip oluşturulmuştur. Bu doğrultuda Tablo 5’teki birinci satır, birinci

maddenin doğru cevaplanması için bir öğrencinin üç nitelikli Q matrisinde ikinci niteliğe, yedi nitelikli Q matrisinde B ve ikinci niteliğe, 11 nitelikli Q matrisinde ise B2 niteliğine sahip olması gerektiğini ifade etmektedir.

Tablo 5

Gerçek Veri için İlk 20 Maddeye Ait Q Matrisleri

3 nitelikli Q matrisi			7 nitelikli Q matrisi							11 nitelikli Q matrisi										
1	2	3	A	B	C	D	1	2	3	A1	A2	B1	B2	B3	C1	C2	C3	D1	D2	D3
0	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
0	1	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	1	0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1
0	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0
1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0
0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0
0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0
1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0

Bilişsel tanılama alanında çeşitli modellerin sınıflandırma oranları üzerinde yapılan araştırmalar, genellikle farklı koşullar için elde edilen simülasyon verileri üzerinde yürütülmüş ardından gerçek veri analizinden elde edilen bulgular ile karşılaştırılmıştır (bkz. de la Torre & Douglas, 2004, 2008; de la Torre vd., 2010). Bu araştırma kapsamında da gerçek veri ile simülasyon veri setlerine ait koşullar arasında bazı benzerlik ve farklılıklar vardır. Gerçek veri seti için ele alınan faktörler ve bu faktörlere ilişkin düzeyler Tablo 6’da verilmiştir. Bu araştırma kapsamında gerçek veri setlerinde $3 \times 3 \times 3 = 27$ karşılaştırma durumu oluşturulmuştur.

Tablo 6

Gerçek Veri Seti için Ele Alınan Faktörler ve Bu Faktörlere İlişkin Düzeyler

Faktörler	Karşılaştırma Koşulları		
Nitelik sayısı (Ni)	3	7	11
Örneklem büyüklüğü (N)	30	100	500
Modeller	POBT	YSA	DINO

Tablo 6’da yer alan faktörlerin açıklanmasının faydalı olacağı düşünüldüğünden bu faktörler aşağıda sırasıyla bahsedilmiştir.

Mevcut araştırma kapsamında nitelik sayısı faktörüne ait koşullar PISA 2015 İPÇ yapısına bağlı olarak belirlenmiştir. Alanyazın incelendiğinde eğitimde ölçme ve değerlendirme dersine ait yedi niteliğin 25 madde ile ölçüldüğü ölçme aracını geliştiren Başokçu’nun (2011), 471 öğrenciye ait veriyi DINA model ile analiz ettiği görülmektedir. Benzer şekilde Toker (2010), 4498 Türk öğrenciye ait TIMSS-2007 matematik verisini En Küçük Kareler Uzaklık Yöntemi (The Least Square Distance Method, LSDM) kullanarak analiz etmiştir. Toker (2010), araştırmasında TIMSS 2007 matematik uygulamasını 20 nitelik ve 49 maddeden oluşan Q matrisi ile analiz etmiştir. Buradan hareketle gerçek veri setlerinde yapılan bilişsel tanılama çalışmalarında nitelik sayısının yüksek olduğu söylenebilir. Nitekim Sessoms ve Henson (2018), gerçek veri setlerinde yapılan BTM çalışmalarının % 47’sinin sekiz veya daha fazla niteliği ölçtüğünü bulmuştur. Sessoms ve Henson’ın (2018) ele aldığı çalışmalar büyük ölçekli (ulusal ve uluslararası) değerlendirmelerken, sınıf düzeyindeki değerlendirmelerde nitelik sayısının artırılması öğrenci ve öğretmene daha detaylı dönüt verilmesi bakımından yararlı olacağı düşünülmektedir.

Mevcut araştırmanın odağında küçük örneklemelerde bilişsel tanılamamanın hangi yöntemle daha iyi yapılabileceğinin belirlenmesi bulunduğundan örneklem faktörüne ait koşullar küçük seçilmiştir. Araştırma kapsamında ele alınan örneklem büyüklüklerinin dağılımlarına ait betimsel bulgular EK-1’de verilmiştir.

3.4. Verilerin Analizi

Verilerin düzenlenmesinde Microsoft Excel 2019, faktöriyel ANOVA’nın yapılmasında IBM SPSS 26.0 (IBM Corp., 2019), doğrulayıcı faktör analizi için Mplus 6.12 (Muthén ve Muthén, 1998-2011) programlarından faydalanılmıştır. Verilerin üretilmesinde, gerçek veriye ait Q matrislerinin doğrulanmasında ve analizlerinde R-Studio programı kullanılmıştır. Gerçek veriye ait Q matrislerinin doğrulanmasında GDINA (Ma, de la Torre, Sorrel, & Jiang, 2020), DINO analizi için CDM (Robitzsch, Kiefer, George & Uenlue, 2020), POBT analizi için NPCD (Zheng, Chiu, & Douglas, 2019), YSA analizi için neuralnet (Fritsch, vd., 2019), kayıp veri oluşturma ve atama için missForest (Stekhoven, 2016) ve TestDataImputation (Dai, Wang & Svetina, 2019), son olarak

grafikler için ggplot2 (Wickham, vd., 2020) ve lattice (Sarkar, vd., 2020) paketlerinden yararlanılmıştır.

Birinci araştırma probleminin cevaplanması için simülasyon veri setlerinde POBT, YSA ve DINO model analizleri yapılarak çeşitli koşullar altında elde edilen sınıflandırma oranları önce yorumlanmış, ardından sınıflandırma oranlarının ortalamalarının koşullarda nasıl değiştiğinin belirlenmesi için 3x3x3x3 faktöriyel ANOVA yapılmıştır. Burada öncelikle YSA analizleri için ek açıklamalara ihtiyaç olduğu düşünülmektedir. Araştırma kapsamında YSA eğitimi en az 100 en çok 1000 olacak şekilde tekrarlanmıştır. Hata fonksiyonu olarak çapraz entropi (bkz. Günther & Fritsch, 2010) seçilmiş ve fonksiyon hem gizil katman/katmanlar hem de çıktı katmanları için lojistik fonksiyon (bkz. Günther & Fritsch, 2010) ile birlikte incelenmiştir. Ayrıca YSA eğitiminde geleneksel geri yayılım algoritması kullanılmadığı için ağırlık öğrenme oranı parametresi, AMORE paketi kullanılmadığı için de ağırlık momentum parametresi için özel bir değer kullanılmamıştır. YSA analizlerinde gizil katman ve nöron sayısının belirlenmesi ön denemeler ile olmaktadır. Intrator ve Intrator (2001), girdi ve çıktı katmanları arasındaki ilişkilerin çoğunu en fazla iki gizil katmanın açıklayabildiğini belirtmektedirler. Bu nedenle araştırma kapsamında gizil katman sayısı en fazla iki olacak şekilde sınırlandırılmıştır. Olası nöron sayıları, çıkış katmanı nöron sayısından (nitelik sayısı) başlayıp giriş katmanı nöron sayısının (madde sayısı) iki katı olacak şekilde değişimlenmiştir. Tek bir gizil katmanda çapraz entropi kayıp fonksiyonunu en aza indiren veya fonksiyonu kararlı hale getiren nöron sayısı belirlendikten sonra, iki katmanlı bir yapının kayıp fonksiyonunu daha aza indirip indirmeyeceğini belirlemek için iki gizil tabaka ile benzer bir araştırma gerçekleştirilmiştir. Bu durum, bazı koşulların tek katmanlı bazılarının ise iki katmanlı yapılara sahip olmasına neden olmuştur. Gizil katman ve nöron sayısını belirlemek için izlenen bu adımlar her bir koşul için ayrı ayrı gerçekleştirilmiştir.

Bilişsel tanılamada yapılan sınıflandırmaların güvenilirliği çeşitli indekslerle ifade edilmektedir. Örneğin Cui, Leighton, Gierl ve Hunka (2006), nitelikler arası ardışıklık ilişkisinin bulunduğu parametrik bir sınıflandırmada hiyerarşi tutarlılık indeksini (hierarchy consistency index) önermişlerdir. Araştırma kapsamında nitelikler arası ardışıklık söz konusu olmadığı için bu indeks kullanılmamıştır. Nitelikler arası ardışıklık ilişkisinin bulunmadığı durumlar için Cui, Gierl ve Chang (2012), sınıflandırma doğruluğu (Classification accuracy, Pa) ve sınıflandırma tutarlılığı (Classification consistency, Pc) indekslerini önermişlerdir. Bu indeksler parametrik modellerin analizi sonucu niteliklerin

sonsal olasılık (a posteriori) dağılımlarının kullanılmasıyla elde edilmektedirler. Cui vd. (2012) ve Wang, Song, Chen, Meng ve Ding (2015) bu indekslerin değişimleyerek bireysel nitelik düzeyinde güvenilirliği incelemek için kullanılmasını önermişlerdir. Daha sonra Johnson ve Sinharay (2018), Cui, vd. (2012) ve Wang, vd.'nin (2015) önerdikleri indeksleri doğru pozitif ve doğru negatif ile ağırlıklandırarak indekslerin son halini belirlemişlerdir. Son olarak, Johnson ve Sinharay (2020), niteliklerin sonsal olasılık dağılımlarının ortalamalarına dayalı üç farklı güvenilirlik ölçüsünün kullanılmasını önermişlerdir. POBT ve YSA, niteliklerin sonsal olasılık dağılımlarını hesaplayamadıkları için araştırma kapsamında Pa ve Pc indeksleri kullanılamamıştır.

Parametrik olmayan bilişsel tanılama çalışmaları incelendiğinde kümeleme analizinin kullanıldığı çalışmalarda sınıflandırma güvenilirliğini belirlemek için Düzeltilmiş Rand Endeksi'nin (the Adjusted Rand Index-ARI, Hubert & Arabie, 1985) ve ω 'nın (bkz. Chiu, 2008; Chiu & Köhn, 2016) kullanıldığı görülmektedir. Doğru sınıflandırma oranları yalnızca bireyler etiketli kümeler halinde gruplandırıldığında uygulanabilmektedir. Kümeleme analizinde bireyler etiketli sınıflara yerleştirilmekten öte, uzayda birbirlerine yakınlıkları ile sınıflandırıldıkları için kümeleme analizinin kullanıldığı çalışmalarda sınıflar arası uyum birliği ve sınıf içi homojenlik ölçüsü olan ARI ve ω 'nın (Chiu, Douglas, & Li, 2009) kullanılması yerinde bir seçimdir.

Parametrik olmayan bilişsel tanılama çalışmalarında sınıflandırma güvenilirliğini belirlemek için kullanılan diğer indeksler nitelik (NSO) ve örüntü (ÖSO) düzeyi sınıflandırma oranlarıdır (bkz. Wang vd., 2015). Nitelik ve örüntü düzeyi sınıflandırma oranları, elde edilen ve bilinen gerçek sınıflandırma arasındaki uyum oranını yansıtır (Akbay, 2016). Nitelik ve örüntü düzeyi sınıflandırma oranlarının isimlendirmesi çalışmalara göre değişmektedir. Örneğin, Chiu ve Douglas (2013) çalışmalarında NSO ve ÖSO'yu sırasıyla nitelik temelli uyum oranı (AAR; Attribute-wise agreement rate) ve örüntü temelli uyum oranı (PAR; pattern-wise agreement rate) olarak ifade etmiştir. Akbay (2016) ise NSO ve ÖSO'yu sırasıyla nitelik düzeyi doğru sınıflandırma oranı (attribute-level correct classification rate-ACR) ve vektör düzeyi doğru sınıflandırma oranı (vector-level correct classification rate-VCR) olarak ifade etmiştir. Mevcut araştırma kapsamında öğrenciler etiketli sınıflara yerleştirilmişlerdir. Bu nedenle POBT, YSA ve DINO modelin sınıflandırma performansları, simülasyon veri seti için NSO ve ÖSO ile incelenmiştir. NSO ve ÖSO'nun formülleri aşağıda verilmiştir.

$$NSO = \frac{1}{N \times K} \sum_{i=1}^N \sum_{k=1}^K I[\widehat{\alpha}_{ik} = \alpha_{ik}] \quad (15)$$

$$\ddot{OSO} = \frac{1}{N} \sum_{i=1}^N I[\widehat{\alpha}_i = \alpha_i] \quad (16)$$

Formül 15 ve Formül 16’da N , örneklem; K , nitelik sayısı; i , birey; α_{ik} , i . bireyin gerçekte sahip olduğu k . nitelik ve $\widehat{\alpha}_{ik}$, i . bireyin sahip olduğu tahmin edilen k . nitelik. Formülden anlaşılacağı üzere sınıflandırma oranları, doğru olarak tahmin edilen öğrencilerin sahip olduğu ve olmadığı niteliklerin sayılarının tüm karşılaştırmalara bölünmesine dayalıdır.

İncelenen alanyazında NSO ve ÖSO’ların yorumlanmasına yönelik herhangi bir ölçütle karşılaşılmamıştır. Araştırma kapsamında dikkate alınan NSO ve ÖSO, elde edilen ve bilinen gerçek sınıflandırma arasındaki uyum oranını yansıttıkları için bu indekslerin yorumlanmasında Johnson ve Sinharay’ın (2018) Goodman ve Kruskal güvenirlilik katsayısı yorumlama ölçütleri dikkate alınmıştır. Johnson ve Sinharay (2018), BTM alanında Goodman ve Kruskal’ın (1954) Λ uyum oranı güvenirlilik katsayısının yorumlanmasında şu aralıkların dikkate alınmasını önermişlerdir: güvenilir değil <0,25; zayıf güvenirlilik (0,25 – 0,50); kabul edilebilir güvenirlilik (0,51 – 0,65); iyi güvenirlilik (0,66 – 0,80); çok iyi güvenirlilik (0,81 – 0,90); mükemmel güvenirlilik > 0,90.

POBT, YSA ve DINO modeline ait NSO ve ÖSO’ların yorumlanmasından sonra faktöriyel ANOVA’ya geçilmiştir. Birden fazla bağımsız değişkene ait ortalamaları karşılaştırmak ve bu bağımsız değişkenlerin etkileşimlerini incelemek için faktöriyel ANOVA kullanılmaktadır. Faktöriyel ANOVA’nın bağımsız ölçümler faktöriyel desen (Independent Measures Factorial Design), tekrarlı ölçümler deseni (Related Measures Design) ve faktöriyel karma ölçümler deseni (Factorial Mixed Measures Designs) olmak üzere üç türü bulunmaktadır (Rutherford, 2011). Araştırma kapsamında POBT’nin, YSA’nın ve DINO modelinin ortalama NSO ve ÖSO’ları faktörler ve bu faktörlerin koşulları üzerinde tek tek karşılaştırıldığı için bağımsız ölçümler faktöriyel desen seçilmiştir. Araştırmada, ortalamaların faktörler arasında karşılaştırılarak ana etkilerin (main effects) ve etkileşim etkilerinin (interaction effects) istatistiksel olarak anlamlılığı incelendiği 3x3x3x3 faktöriyel ANOVA analizi yapılmıştır.

ANOVA, bağımlı değişkenin sürekli bir veri yapısına sahip olması, en az iki kategoriye sahip en az iki bağımsız değişkenin olması, gözlemlerin bağımsızlığı, grup içi dağılımların normal olması ve varyansların homojenliği gibi bazı varsayımlara sahiptir. Araştırma kapsamında faktöriyel ANOVA’ya her bir replikasyondan elde edilen veriler dâhil edilmiştir. İncelemeler sonucunda, verilerin ilk üç varsayımı sağlarken normallik ve

varyansların homojenliği varsayımlarının ihlal edildiği görülmüştür. Bu varsayımın sağlanmamasında büyük örneklem etkisinin (Kirk, 2014) olduğu düşünülmektedir. Buna ek olarak, ANOVA özellikle örneklem büyüklüklerinin eşit olduğu durumlarda örneklem büyüklüğü 12'den büyük olduğunda normallik varsayımının ihlalleri açısından (Clinch ve Keselman, 1982; Tan, 1982), örneklem büyüklüğü beşten büyük olduğunda (Clinch ve Keselman, 1982; Tomarken ve Serlin, 1986) ise orta dereceli varyansların homojenliği varsayımı ihlalleri açısından güçlüdür (Rutherford, 2011). Mevcut araştırmada her bir koşul için replikasyon sayısı eşit ve 100 olduğu için faktöriyel ANOVA'ya devam edilmiştir. Birinci tip hata oranı üzerinde son iki varsayımın ihlalinin kaynaklanabilecek olası etkilerden kaçınmak için anlamlılık düzeyi $p < 0,01$ olarak belirlenmiştir. İstatistiksel olarak anlamlı farkların bulunduğu durumlarda bu farkın hangi model (POBT, YSA veya DINO model) lehine olduğunun belirlenmesinde varyansların eşit olmadığı durumlar için Tamhane's T2, Dunnett's T3, Games–Howell ve Dunnett's C. gibi birçok post-hoc çoklu aralık testi geliştirilmiştir (Sparks, 1963). Araştırma kapsamında bu testler arasından Tamhane's T2 ve Games–Howell testleri dikkate alınmıştır. Faktöriyel ANOVA sonucu hesaplanan etki büyüklükleri Cohen'e (1988) göre küçük etki (0,20); orta etki (0,50); geniş etki büyüklüğü (0,80) olarak yorumlanmıştır.

İkinci araştırma probleminin cevaplanması için POBT, YSA ve DINO modeline ait sınıflandırma oranlarının ortalamaları faktörler ve bu faktörlere ait koşullarda karşılaştırılmıştır. Bu karşılaştırmalar için sekiz tane 3x3 faktöriyel ANOVA yapılmıştır. Etkileşim etkisi istatistiksel olarak anlamlı bulunan faktörlerin her bir koşulu için basit etkilerinin (simple effect) istatistiksel olarak anlamlılığı incelenmiştir. Faktöriyel ANOVA analizleri için kurulan eşitliklerden bir tanesi örnek olarak Formül 17'de verilmiştir. Formül 17'de, nitelik sayısı faktöründe POBT, YSA ve DINO modelinin NSO'ları arasında istatistiksel olarak anlamlı bir farkın olup olmadığı incelenmektedir. Basit etkilerin analizi için kullanılan sentaksların bir örneği EK-12'de verilmiştir.

$$NSO_{ijk} = \mu + Nitelik_j + Model_k + (NitelikModel)_{jk} + \varepsilon_{i(jk)} \quad (i = 1, \dots, n; j = 1, \dots, p; k = 1, \dots, q) \quad (17)$$

NSO_{ijk} : k . modelin j . nitelik koşulunda gözlenen NSO değeridir.

μ : NSO'lara ait genel ortalamadır.

$Nitelik_j$: nitelik etkisi, $\mu_{.j} - \mu$ 'ye eşittir. Başka bir ifadeyle nitelik etkisi, j . nitelik koşulundaki NSO'ların ortalaması ile genel ortalama arasındaki farktır.

Model_k: model etkisi, $\mu_{.k} - \mu$ 'ye eşittir. Başka bir ifadeyle model etkisi, k . modele ait NSO'ların ortalaması ile genel ortalama arasındaki farktır.

(NitelikModel)_{jk}: etkileşim etkisi, $\mu_{jk} - \mu_j - \mu_{.k} + \mu$ 'ye eşittir. Başka bir ifadeyle model ve nitelik etkileşimlerine ait ortalama j . nitelik koşulu ve k . modele ait ortalmaların çıkartılıp genel ortalamanın eklenmesiyle elde edilir.

$\varepsilon_{i(jk)}$: rastgele hata etkisidir.

Üçüncü araştırma probleminin cevaplanması için gerçek veri setlerinden (PSA 2015 İPÇ) elde edilen POBT, YSA ve DINO modeline ait nitelik (NSOB) ve örüntü düzeyi sınıflandırma oranlarının benzerliği (ÖSOB) yorumlanmıştır. Gerçek veri setlerinin analizine başlamadan önce maddeler arası çoklu bağlantı sorunu olup olmadığına maddeler arası korelasyon, tolerans (tolerance) ve VIF (variance inflation factor) değerleri incelenerek karar verilmiştir. Yapılan analizler sonucunda maddeler arasında çoklu bağlantı olmadığı belirlenmiştir. Maddeler arası çoklu bağlantı durumu incelendikten sonra PISA 2015 İPÇ uygulamasına ait 18170 kişi üzerinde ve çalışma koşullarını oluşturan 500, 100 ve 30 örneklemde üç ve onbir nitelik durumları faktör sayısı olarak ele alınıp doğrulayıcı faktör analizi (DFA) yapılmıştır. 18170 kişilik, 500 kişilik ve 100 kişilik örneklemde üç ve onbir faktörlü olarak yapılan DFA sonucu χ^2 değerlerinin anlamlı ($p < 0,05$), RMSEA değerlerinin 0,030 ile 0,059 arasında, CFI değerlerinin 0,958 ile 0,911 arasında, TLI değerlerinin ise 0,956 ile 0,908 arasında olduğu görülmektedir. Otuz örneklemde üç ve onbir faktörlü olarak yapılan DFA sonucunda χ^2 değerlerinin anlamlı ($p < 0,05$), RMSEA değerlerinin sırasıyla 0,118 ve 0,117; CFI değerlerinin sırasıyla 0,853 ve 0,860; TLI değerlerinin ise sırasıyla 0,849 ve 0,853 olduğu görülmektedir. Bu değerlerin örneklem büyüklüğü nedeniyle düşük çıktığı düşünülmektedir. Araştırma odağında küçük örneklem koşulu bulunduğu için analize devam edilmiştir. Faktör analizinden sonra, PISA 2015 teknik raporlarında maddelerin ölçtükleri niteliklerin yanlış ifade edilmesi ihtimaline yönelik raporlar doğrultusunda elde edilen Q matrisleri üzerinde PVAF yöntemine (de la Torre & Chiu, 2015) dayalı Q matris geçerlemesi (doğrulaması) yapılmıştır. Üçüncü araştırma probleminin cevaplanmasında PISA 2015 teknik raporlarında ifade edilen Q matrisleri ile birlikte Q matris doğrulaması sonucu elde edilen Q matrislerinden faydanılarak bulgular raporlanmıştır.

Gerçek veri setlerinde öğrencilerin gerçek nitelik sınıfları bilinmediği için Formül 15 ve Formül 16 gerçek veri analizine uygun olarak değiştirilmiştir. Bunun için α_{ik} ,

karşılaştırılan modellerden birinin kestirdiği α değeriyle $\widehat{\alpha}_{lk}$ ise diğer modelin kestirdiği α değeri ile değiştirilmiştir.



BÖLÜM IV

BULGULAR VE YORUM

Bu bölümde bulgular, araştırma problemleri doğrultusunda üç başlık altında verilmiştir.

4.1. Simülasyon Verisinde POBT'nin, YSA'nın ve DINO Modelinin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması

POBT'ye, YSA'ya ve DINO modeline ait NSO ve ÖSO'ların grafiksel olarak özeti EK-4'te verilmiştir. Bu araştırma probleminin cevaplanması için POBT'ye, YSA'ya ve DINO modeline ait NSO ve ÖSO'lar araştırma koşullarında ayrı ayrı incelenmiştir. POBT'ye, YSA'ya ve DINO modeline ait NSO ve ÖSO'lar öncelikle her bir model için sırasıyla Tablo 7, Tablo 10 ve Tablo 13'de özetlenip yorumlanmıştır. Bu tablolarda verilen NSO ve ÖSO'ların, her bir koşul için üretilen 100 yinelemenin (replikasyonun) analizi sonucunda elde edilen 100'er NSO ve ÖSO'ların ortalaması olduğunun belirtilmesinde yarar vardır. POBT, YSA ve DINO modeline ait NSO ve ÖSO'lar genel olarak yorumlandıktan sonra koşullara ait ana ve etkileşim etkilerinin belirlenmesi için 3x3x3x3 faktöriyel ANOVA yapılmıştır.

4.1.1. POBT'nin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması

POBT analizi sonucu çeşitli koşullarda elde edilen sınıflandırma oranlarına ait detaylı bilgiler EK-5'te, AIC ve BIC değerleri ise EK-8'de raporlanmıştır. Değişen koşullarda POBT'ye ait NSO ve ÖSO'lar Tablo 7'de özetlenmiştir.

Tablo 7 incelendiğinde POBT'ye ait NSO'ların 0,78 ile 0,99 arasında, ÖSO'ların ise 0,23 ile 0,98 arasında değiştiği görülmektedir. POBT'ye ait NSO'ların iyi ile mükemmel güvenilirlik ölçütleri arasında, ÖSO'ların ise zayıf ile mükemmel güvenilirlik ölçütleri arasında oldukları söylenebilir. POBT'ye ait NSO'ların tüm koşullarda 0,78 üzerindiyken ÖSO'ların güvenilirlik derecesi kabul edilebilir düzeyin üstünde (0,65 üzeri) olma durumlarının koşullara göre değiştiği görülmektedir.

Tablo 7

Değişen koşullarda POBT'ye ait NSO ve ÖSO'lar

Ni	N	M	NSO			ÖSO		
			Kayıp veri oranı			Kayıp veri oranı		
			%0	%5	%10	%0	%5	%10
3	30	15	0,930	0,903	0,883	0,810	0,751	0,707
		30	0,956	0,949	0,941	0,879	0,862	0,840
		45	0,978	0,979	0,979	0,938	0,939	0,943
	100	15	0,877	0,875	0,902	0,708	0,699	0,752
		30	0,957	0,949	0,954	0,883	0,861	0,873
		45	0,971	0,983	0,968	0,919	0,950	0,910
	500	15	0,907	0,869	0,924	0,761	0,692	0,801
		30	0,945	0,954	0,944	0,854	0,875	0,849
		45	0,992	0,982	0,970	0,977	0,947	0,915
5	30	15	0,842	0,796	0,847	0,471	0,380	0,477
		30	0,883	0,885	0,899	0,573	0,572	0,608
		45	0,946	0,948	0,944	0,775	0,783	0,765
	100	15	0,881	0,862	0,853	0,554	0,526	0,510
		30	0,922	0,910	0,901	0,686	0,665	0,619
		45	0,977	0,952	0,943	0,897	0,794	0,767
	500	15	0,878	0,849	0,876	0,560	0,478	0,552
		30	0,909	0,917	0,898	0,643	0,684	0,621
		45	0,961	0,940	0,936	0,828	0,757	0,742
7	30	15	0,816	0,818	0,808	0,285	0,286	0,288
		30	0,844	0,878	0,870	0,357	0,446	0,413
		45	0,923	0,925	0,920	0,606	0,610	0,581
	100	15	0,823	0,838	0,793	0,306	0,358	0,260
		30	0,907	0,874	0,835	0,538	0,415	0,355
		45	0,957	0,936	0,929	0,758	0,662	0,631
	500	15	0,805	0,777	0,797	0,271	0,234	0,257
		30	0,894	0,858	0,832	0,495	0,384	0,336
		45	0,930	0,916	0,930	0,630	0,565	0,630

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısını temsil etmektedir.

POBT'nin güvenilirlik derecesi kabul edilebilir düzeyin üstünde ÖSO'lara sahip olduğu koşulların,

- Nitelik sayısı üç iken tüm örneklem büyüklüğü, madde sayısı ve kayıp veri oranı koşulları,
- Nitelik sayısı beş iken tüm örneklem büyüklükleri ve kayıp veri oranlarında madde sayısı 45 olduğu,
- Nitelik sayısı beş ve madde sayısı 30 iken örneklem büyüklüğü 100 ise kayıp veri oranı sıfır veya 0,05 olduğu, örneklem büyüklüğü 500 ise kayıp veri oranı 0,05 olduğu,
- Nitelik sayısı yedi iken örneklem büyüklüğü 100 ve madde sayısı 45 ise kayıp veri oranı sıfır veya 0,05 olduğu görülmektedir.

4.1.1.1. Değişen koşullar altında POBT'ye ait NSO'lar için faktöriyel ANOVA bulguları

POBT'ye ait NSO'ların araştırma kapsamında ele alınan faktörlerde ve bu faktörlerin koşullarında nasıl değiştiğinin belirlenmesi için 3x3x3x3 faktöriyel ANOVA yapılmıştır. Bulgular, ana ve etkileşim etkileri tablosu ve etkileşim grafiğiyle sunulmuştur. Ana ve etkileşim etkilerine ait Tablo 8 incelendiğinde madde sayısı ($F_{(2,8019)} = 20641,23$; $p < 0,01$), nitelik sayısı ($F_{(2,8019)} = 11289,33$; $p < 0,01$), örneklem büyüklüğü ($F_{(2,8019)} = 166,63$; $p < 0,01$) ve kayıp veri oranı ($F_{(2,8019)} = 376,85$; $p < 0,01$) faktörlerinin ana etkilerinin istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 8 incelendiğinde nitelik sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 252,51$; $p < 0,01$), nitelik ile madde sayısı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 340,97$; $p < 0,01$), nitelik sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 69,72$; $p < 0,01$), madde sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 7,04$; $p < 0,01$), örneklem büyüklüğü ile kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 93,15$; $p < 0,01$) ve madde sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 88,75$; $p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 8

Değişen koşullar altında POBT'ye ait NSO'lar için faktöriyel ANOVA sonuçları

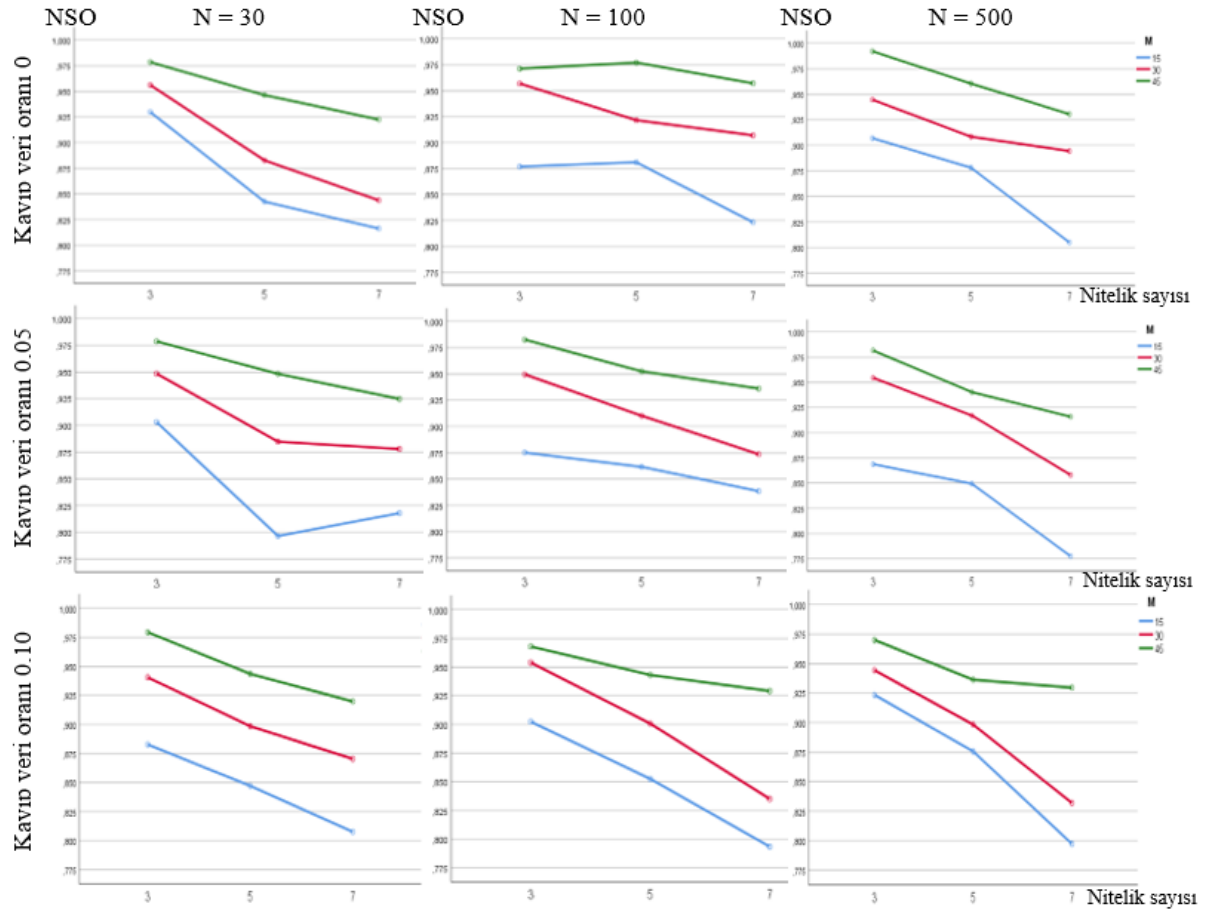
Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Ni	7,323	2	3,661	11289,330	0,00	0,738
N	0,108	2	0,054	166,633	0,00	0,040
M	13,389	2	6,694	20641,229	0,00	0,837
KayıpVeri	0,244	2	0,122	376,851	0,00	0,086
Ni * N	0,328	4	0,082	252,505	0,00	0,112
Ni * M	0,442	4	0,111	340,973	0,00	0,145
Ni * KayıpVeri	0,090	4	0,023	69,719	0,00	0,034
M * N	0,009	4	0,002	7,036	0,00	0,003
N * KayıpVeri	0,121	4	0,030	93,150	0,00	0,044
M * KayıpVeri	0,115	4	0,029	88,750	0,00	0,042
Ni * N * M	0,220	8	0,028	84,941	0,00	0,078
Ni * N * KayıpVeri	0,305	8	0,038	117,632	0,00	0,105
Ni * M * KayıpVeri	0,152	8	0,019	58,666	0,00	0,055
M * N * KayıpVeri	0,229	8	0,029	88,422	0,00	0,081
Ni * N * M * KayıpVeri	0,245	16	0,015	47,119	0,00	0,086
Hata	2,601	8019	0,0003			

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısı, KayıpVeri: Kayıp veri oranını temsil etmektedir.

Tablo 8'den faktörler arasındaki üçlü etkileşim etkileri incelendiğinde nitelik ve madde sayısı ile örneklem büyüklüğü faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 84,94$; $p < 0,01$), nitelik sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 117,63$; $p < 0,01$), nitelik ve madde sayısı ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 58,67$; $p < 0,01$) ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 88,42$; $p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Son olarak, Tablo 8 incelendiğinde dörtlü etkileşim etkilerinin, yani nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(2,8019)} = 47,12$; $p < 0,01$) ve bu etkileşimin küçük etkiye

sahip olduğu görülmektedir ($\eta_p^2 = 0,09$). Nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı koşullarının etkileşimlerine ait grafik Şekil 2’de verilmiştir.



Şekil 2. Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında POBT’ye ait ortalama NSO’lar

Şekil 2 incelendiğinde tüm kayıp veri oranı, örneklem büyüklüğü ve nitelik sayısı koşullarında POBT’nin en yüksek ortalama NSO’lara madde sayısı 45 iken, en düşük ortalama NSO’lara ise madde sayısı 15 iken sahip olduğu görülmektedir. Buradan hareketle madde sayısı arttıkça POBT’ye ait ortalama NSO’ların arttığı söylenebilir. Şekil 2’den nitelik sayısı arttıkça madde sayısı koşullarındaki ortalama NSO’ların genel olarak azaldığı görülmektedir. Buna bağlı olarak nitelik sayısı arttıkça POBT’ye ait ortalama NSO’ların azaldığı söylenebilir. Şekil 2’den kayıp veri oranındaki değişimin POBT’ye ait ortalama NSO’lara etkisi incelendiğinde, kayıp veri oranı arttıkça ortalama NSO’ların genel olarak azaldığı söylenebilir. Son olarak, örneklem büyüklüğünün artmasıyla ortalama NSO’ların azalmaları, artmaları veya sabit kalmaları hakkında herhangi bir sistematikliğe rastlanmamıştır.

4.1.1.2. Değişen koşullar altında POBT'ye ait ÖSO'lar için faktöriyel ANOVA bulguları

POBT'ye ait ÖSO'ların araştırma kapsamında ele alınan faktörlerde ve bu faktörlerin koşullarında nasıl değiştiğinin belirlenmesi için 3x3x3x3 faktöriyel ANOVA yapılmıştır. Bulgular ana ve etkileşim etkileri tablosu ve etkileşim grafiğiyle sunulmuştur. Ana ve etkileşim etkilerine ait Tablo 9 incelendiğinde nitelik ($F_{(2,8019)} = 36211,11; p<0,01$) ve madde sayısı ($F_{(2,8019)} = 17013,28; p<0,01$), örneklem büyüklüğü ($F_{(2,8019)} = 309,29; p<0,01$) ile kayıp veri oranı ($F_{(2,8019)} = 317,23; p<0,01$) faktörlerinin ana etkilerinin istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 9 incelendiğinde nitelik sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 189,66; p<0,01$), nitelik ile madde sayısı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 539,53; p<0,01$), nitelik sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 35,38; p<0,01$), madde sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 14,62; p<0,01$), örneklem büyüklüğü ile kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 75,02; p<0,01$) ve madde sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 56,92; p<0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 9'dan faktörler arasındaki üçlü etkileşim etkileri incelendiğinde nitelik ve madde sayısı ile örneklem büyüklüğü faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 44,10; p<0,01$), nitelik sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 72,04; p<0,01$), nitelik ve madde sayısı ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 48,34; p<0,01$) ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 57,82; p<0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 9

Değişen koşullar altında POBT'ye ait ÖSO'lar için faktöriyel ANOVA sonuçları

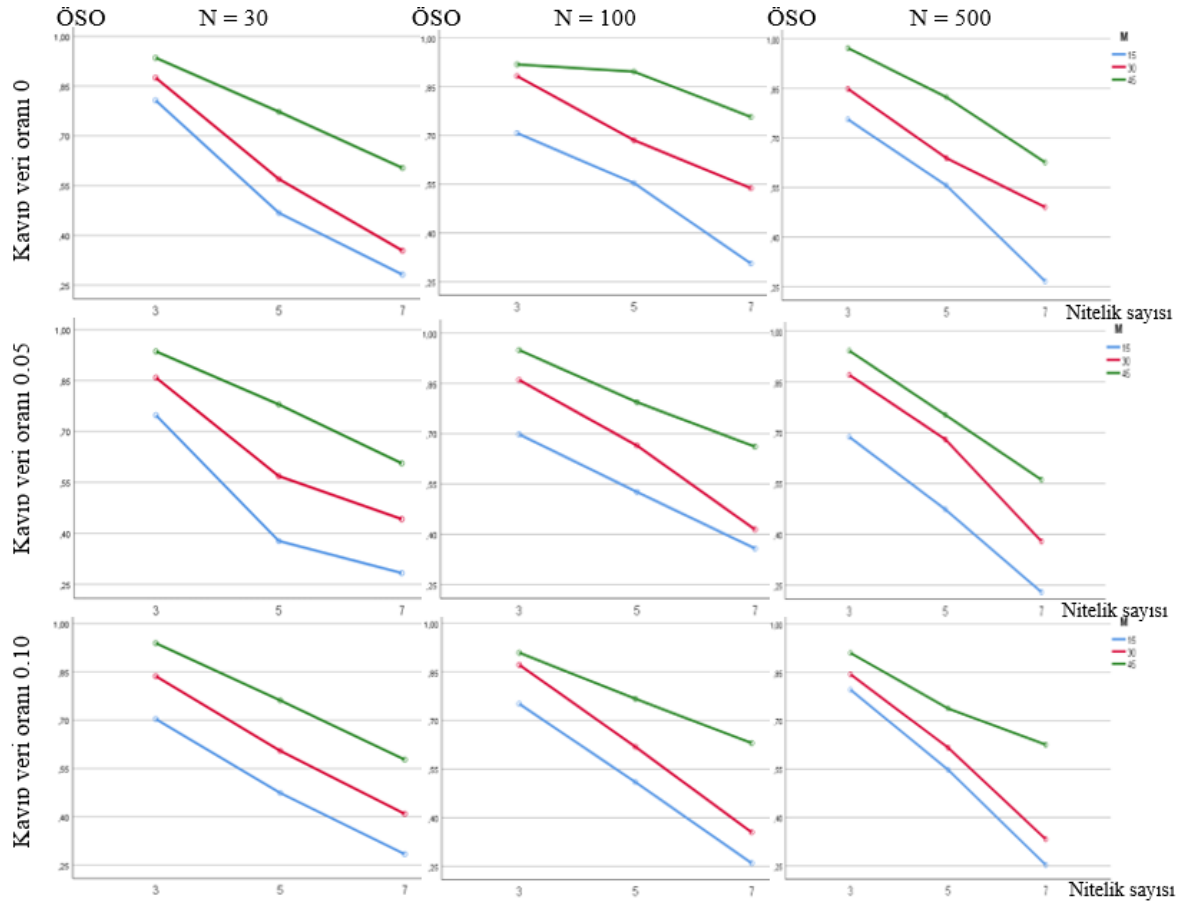
Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Ni	221,478	2	110,739	36211,108	0,000	0,900
N	1,892	2	0,946	309,289	0,000	0,072
M	104,058	2	52,029	17013,280	0,000	0,809
KayıpVeri	1,940	2	0,970	317,229	0,000	0,073
Ni * N	2,320	4	0,580	189,657	0,000	0,086
Ni * M	6,600	4	1,650	539,525	0,000	0,212
Ni * KayıpVeri	0,433	4	0,108	35,380	0,000	0,017
M * N	0,179	4	0,045	14,616	0,000	0,007
N * KayıpVeri	0,918	4	0,229	75,018	0,000	0,036
M * KayıpVeri	0,696	4	0,174	56,924	0,000	0,028
Ni * N * M	1,079	8	0,135	44,098	0,000	0,042
Ni * N * KayıpVeri	1,762	8	0,220	72,037	0,000	0,067
Ni * M * KayıpVeri	1,183	8	0,148	48,340	0,000	0,046
M * N * KayıpVeri	1,415	8	0,177	57,823	0,000	0,055
Ni * N * M * KayıpVeri	1,577	16	0,099	32,239	0,000	0,060
Hata	24,523	8019	0,003			

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısı, KayıpVeri: Kayıp veri oranını temsil etmektedir.

Son olarak, Tablo 9 incelendiğinde dörtlü etkileşim etkilerinin, yani nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(2,8019)} = 32,24$; $p < 0,01$) ve bu etkileşimin küçük etkiye sahip olduğu görülmektedir ($\eta_p^2 = 0,06$). Nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı koşullarının etkileşimlerine ait grafik Şekil 3'te verilmiştir.

Şekil 3 incelendiğinde tüm kayıp veri oranı, örneklem büyüklüğü ve nitelik sayısı koşullarında POBT'nin en yüksek ortalama ÖSO'lara madde sayısı 45 iken, en düşük ortalama ÖSO'lara ise madde sayısı 15 iken sahip olduğu görülmektedir. Buradan hareketle madde sayısı arttıkça POBT'ye ait ortalama ÖSO'ların arttığı söylenebilir. Şekil 3'ten nitelik sayısı arttıkça madde sayısı koşullarındaki ÖSO'ların azaldığı görülmektedir. Buna bağlı olarak nitelik sayısı arttıkça POBT'ye ait ortalama ÖSO'ların azaldığı

söylenbilir. Şekil 3'ten kayıp veri oranındaki değişimin POBT'ye ait ortalama ÖSO'lara etkisi incelendiğinde, kayıp veri oranı arttıkça ortalama ÖSO'ların genel olarak azaldığı söylenbilir. Son olarak, örneklem büyüklüğünün artmasıyla ortalama ÖSO'ların azalmaları, artmaları veya sabit kalmaları hakkında herhangi bir sistematikliğe rastlanmamıştır.



Şekil 3. Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında POBT'ye ait ortalama ÖSO'lar

4.1.2. YSA'nın NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması

YSA analizi sonucu çeşitli koşullarda elde edilen sınıflandırma oranlarına ilişkin detaylı bilgiler EK-6'da verilmiştir. Değişen koşullarda YSA'ya ait NSO ve ÖSO'lar Tablo 10'da özetlenmiştir. Tablo 10 incelendiğinde YSA'ya ait NSO'ların 0,43 ile 0,80 arasında, ÖSO'ların ise 0,01 ile 0,21 arasında değiştiği görülmektedir.

Tablo 10

Değişen koşullarda YSA'ya ait NSO ve ÖSO'lar

Ni	N	M	NSO			ÖSO		
			Kayıp veri oranı			Kayıp veri oranı		
			%0	%5	%10	%0	%5	%10
3	30	15	0,622	0,580	0,577	0,214	0,178	0,167
		30	0,429	0,513	0,496	0,074	0,093	0,111
		45	0,480	0,476	0,467	0,107	0,111	0,092
	100	15	0,508	0,579	0,576	0,115	0,168	0,160
		30	0,560	0,447	0,429	0,161	0,065	0,073
		45	0,484	0,478	0,475	0,108	0,104	0,099
	500	15	0,575	0,512	0,574	0,163	0,114	0,162
		30	0,430	0,530	0,432	0,076	0,115	0,073
		45	0,472	0,482	0,479	0,100	0,099	0,101
5	30	15	0,654	0,684	0,681	0,127	0,165	0,148
		30	0,536	0,549	0,591	0,040	0,044	0,072
		45	0,523	0,484	0,482	0,041	0,024	0,022
	100	15	0,668	0,644	0,654	0,136	0,119	0,127
		30	0,540	0,540	0,524	0,044	0,041	0,033
		45	0,488	0,487	0,485	0,025	0,025	0,028
	500	15	0,680	0,699	0,688	0,152	0,169	0,156
		30	0,535	0,524	0,595	0,041	0,036	0,069
		45	0,522	0,514	0,511	0,041	0,040	0,037
7	30	15	0,788	0,761	0,757	0,193	0,160	0,152
		30	0,632	0,654	0,629	0,046	0,048	0,039
		45	0,617	0,565	0,644	0,036	0,012	0,052
	100	15	0,795	0,782	0,757	0,214	0,188	0,156
		30	0,662	0,668	0,661	0,058	0,060	0,057
		45	0,57	0,679	0,562	0,02	0,077	0,015
	500	15	0,793	0,786	0,760	0,203	0,195	0,158
		30	0,670	0,628	0,636	0,061	0,038	0,041
		45	0,643	0,566	0,633	0,051	0,017	0,044

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısını temsil etmektedir.

YSA'ya ait NSO'ların zayıf ile iyi güvenilirlik ölçütleri arasında olduğu, ÖSO'ların ise güvenilirlik derecesi kabul edilebilir düzeyin üstünde olmadıkları söylenebilir. Ek olarak, Tablo 10'de YSA'ya ait ÖSO'ların tüm koşullarda güvenilirlik derecesi kabul edilebilir

düzeyin altındayken NSO'ların güvenilirlik derecesi kabul edilebilir düzeyin üstünde (0,65 üzeri) olma durumlarının koşullara göre değiştiği görülmektedir. YSA'nın güvenilirlik derecesi kabul edilebilir düzeyin üstünde NSO'lara sahip olduğu koşulların,

- Nitelik sayısı beş ve madde sayısı 15 iken örneklem büyüklüğü 30 ise kayıp veri oranı 0,05 veya 0,10; örneklem büyüklüğü 100 ise kayıp veri oranı sıfır, örneklem büyüklüğü 500 ise tüm kayıp veri oranı koşulları,
- Nitelik sayısı yedi iken 30 örneklem büyüklüğünde madde sayısı 15 ise tüm kayıp veri oranı koşulları,
- Nitelik sayısı yedi iken 100 örneklemde madde sayısı 15 ve 30 ise tüm kayıp veri oranları, madde sayısı 45 ise kayıp veri oranı 0,05 olduğu,
- Nitelik sayısı yedi iken 500 örneklemde madde sayısı 15 ise tüm kayıp veri oranları, madde sayısı 30 ise kayıp veri oranının sıfır olduğu görülmektedir.

4.1.2.1. Değişen koşullar altında YSA'ya ait NSO'lar için faktöriyel ANOVA bulguları

YSA'ya ait NSO'ların araştırma kapsamında ele alınan faktörlerde ve bu faktörlerin koşullarında nasıl değiştiğinin belirlenmesi için 3x3x3x3 faktöriyel ANOVA yapılmıştır. Bulgular ana ve etkileşim etkileri tablosu ve etkileşim grafiğiyle sunulmuştur. Tablo 11 incelendiğinde nitelik ($F_{(2,8019)} = 11253,59$; $p < 0,01$) ve madde sayısı ($F_{(2,8019)} = 8671,70$; $p < 0,01$), örneklem büyüklüğü ($F_{(2,8019)} = 20,93$; $p < 0,01$) ile kayıp veri oranı ($F_{(2,8019)} = 7,77$; $p < 0,01$) ana etkilerinin istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 11 incelendiğinde nitelik sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 61,55$; $p < 0,01$), nitelik ile madde sayısı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 271,42$; $p < 0,01$), nitelik sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 22,76$; $p < 0,01$), madde sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 20,21$; $p < 0,01$), örneklem büyüklüğü ile kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 33,41$; $p < 0,01$) ve madde sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 7,84$; $p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 11'den faktörler arasındaki üçlü etkileşim etkileri incelendiğinde nitelik ve madde sayısı ile örneklem büyüklüğü faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 16,38$; $p < 0,01$); nitelik sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin

($F_{(2,8019)} = 40,43$; $p < 0,01$); nitelik ve madde sayısı ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 43,63$; $p < 0,01$); madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 76,07$; $p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

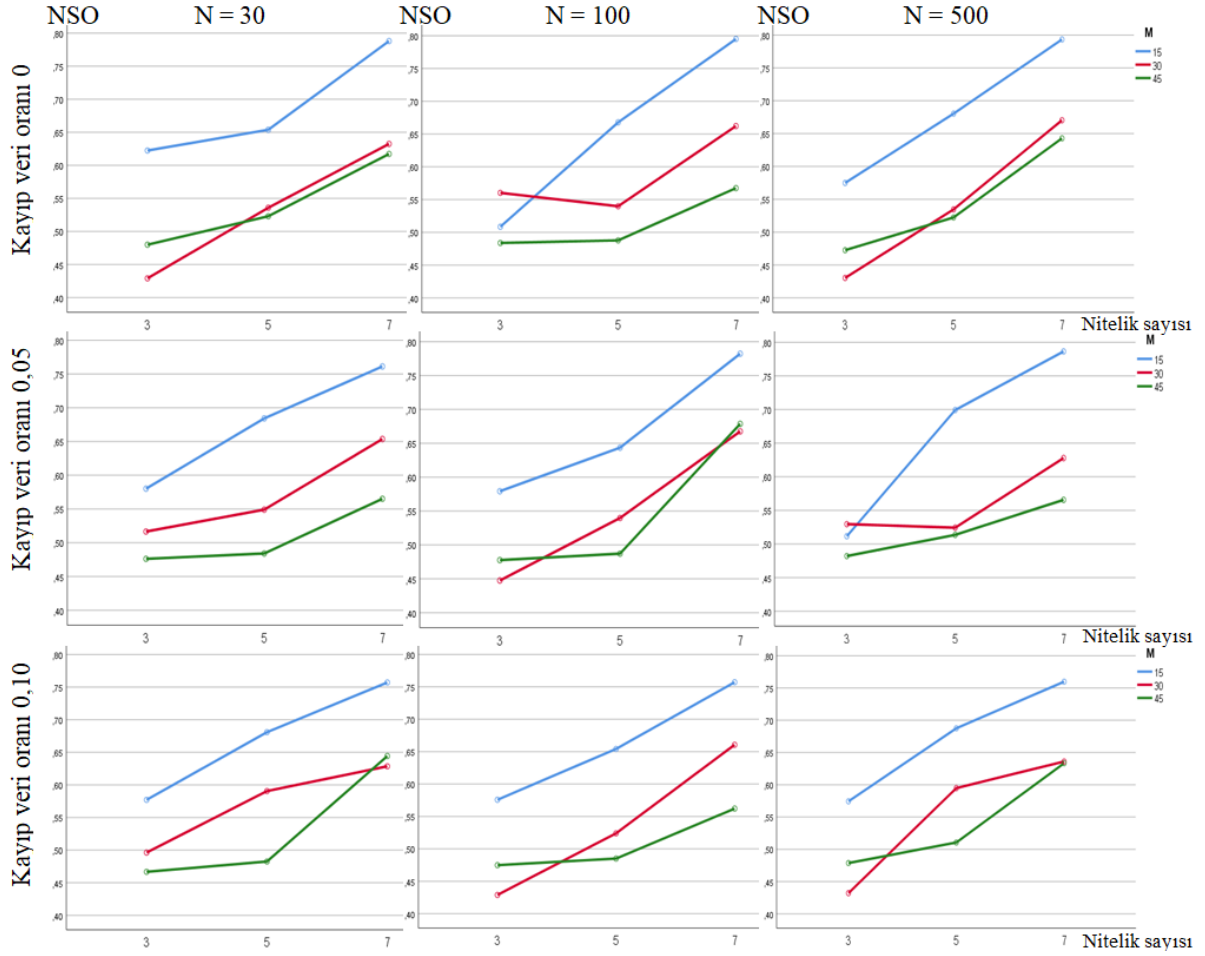
Tablo 11

Değişen koşullar altında YSA'ya ait NSO'lar için faktöriyel ANOVA sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Ni	40,316	2	20,158	11253,592	0,00	0,737
N	0,075	2	0,037	20,932	0,00	0,005
M	31,066	2	15,533	8671,702	0,00	0,684
KayıpVeri	0,028	2	0,014	7,767	0,00	0,002
Ni * N	0,441	4	0,110	61,550	0,00	0,030
Ni * M	1,945	4	0,486	271,422	0,00	0,119
Ni * KayıpVeri	0,163	4	0,041	22,759	0,00	0,011
M * N	0,145	4	0,036	20,209	0,00	0,010
N * KayıpVeri	0,239	4	0,060	33,409	0,00	0,016
M * KayıpVeri	0,056	4	0,014	7,838	0,00	0,004
Ni * N * M	0,235	8	0,029	16,378	0,00	0,016
Ni * N * KayıpVeri	0,579	8	0,072	40,430	0,00	0,039
Ni * M * KayıpVeri	0,625	8	0,078	43,631	0,00	0,042
M * N * KayıpVeri	1,090	8	0,136	76,074	0,00	0,071
Ni * N * M * KayıpVeri	2,590	16	0,162	90,385	0,00	0,153
Hata	14,364	8019	0,002			

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısı, KayıpVeri: Kayıp veri oranını temsil etmektedir.

Son olarak, Tablo 11 incelendiğinde dörtlü etkileşim etkilerinin, yani nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(2,8019)} = 90,39$; $p < 0,01$) ve bu etkileşimin küçük etkiye sahip olduğu görülmektedir ($\eta_p^2 = 0,15$). Nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı koşullarının etkileşimlerine ait grafik Şekil 4'te verilmiştir.



Şekil 4. Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında YSA'ya ait ortalama NSO'lar

Şekil 4 incelendiğinde kayıp veri oranı, örneklem büyüklüğü ve nitelik sayısı koşullarında genel olarak YSA'nın en yüksek ortalama NSO'lara madde sayısı 15 iken, en düşük ortalama NSO'lara ise madde sayısı 45 iken sahip olduğu görülmektedir. Buradan hareketle madde sayısı arttıkça YSA'ya ait ortalama NSO'ların azaldığı söylenebilir. Şekil 4'ten nitelik sayısı arttıkça madde sayısı koşullarındaki ortalama NSO'ların genel olarak arttığı görülmektedir. Buna bağlı olarak nitelik sayısı arttıkça YSA'ya ait ortalama NSO'ların arttığı söylenebilir. Şekil 4'ten kayıp veri oranındaki değişimin YSA'ya ait ortalama NSO'lara etkisi incelendiğinde, kayıp veri oranı arttıkça ortalama NSO'ların az bir miktar azaldığı söylenebilir. Son olarak, örneklem büyüklüğünün artmasıyla ortalama NSO'ların azalmaları, artmaları veya sabit kalmaları hakkında herhangi bir sistematikliğe rastlanmamıştır.

4.1.2.2. Değişen koşullar altında YSA'ya ait ÖSO'lar için faktöriyel ANOVA bulguları

YSA'ya ait ÖSO'ların araştırma kapsamında ele alınan faktörlerde ve bu faktörlerin koşullarında nasıl değiştiğinin belirlenmesi için 3x3x3x3 faktöriyel ANOVA yapılmıştır. Bulgular ana ve etkileşim etkilerinin tablosu ve etkileşim grafiğiyle sunulmuştur. Tablo 12 incelendiğinde madde ($F_{(2,8019)} = 4246,67; p<0,01$) ve nitelik sayısı (sırasıyla $F_{(2,8019)} = 655,92; p<0,01$), örneklem büyüklüğü ($F_{(2,8019)} = 5,88; p<0,01$) ve kayıp veri oranı ($F_{(2,8019)} = 16,92; p<0,01$) ana etkilerinin istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 12 incelendiğinde nitelik sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 38,59; p<0,01$); nitelik ile madde sayısı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 215,79; p<0,01$); nitelik sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 14,39; p<0,01$); madde sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 9,55; p<0,01$); örneklem büyüklüğü ile kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 6,15; p<0,01$); madde sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 4,68; p<0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 12'den faktörler arasındaki üçlü etkileşim etkileri incelendiğinde nitelik ve madde sayısı ile örneklem büyüklüğü faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 11,19; p<0,01$); nitelik sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 9,65; p<0,01$); nitelik ve madde sayısı ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 16,72; p<0,01$) ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 21,11; p<0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Son olarak, Tablo 12 incelendiğinde dörtlü etkileşim etkilerinin, yani nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(2,8019)} = 33,22; p<0,01$) ve bu etkileşimin küçük etkiye sahip olduğu görülmektedir ($\eta_p^2 = 0,06$). Nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı koşullarının etkileşimlerine ait grafik Şekil 5'te verilmiştir.

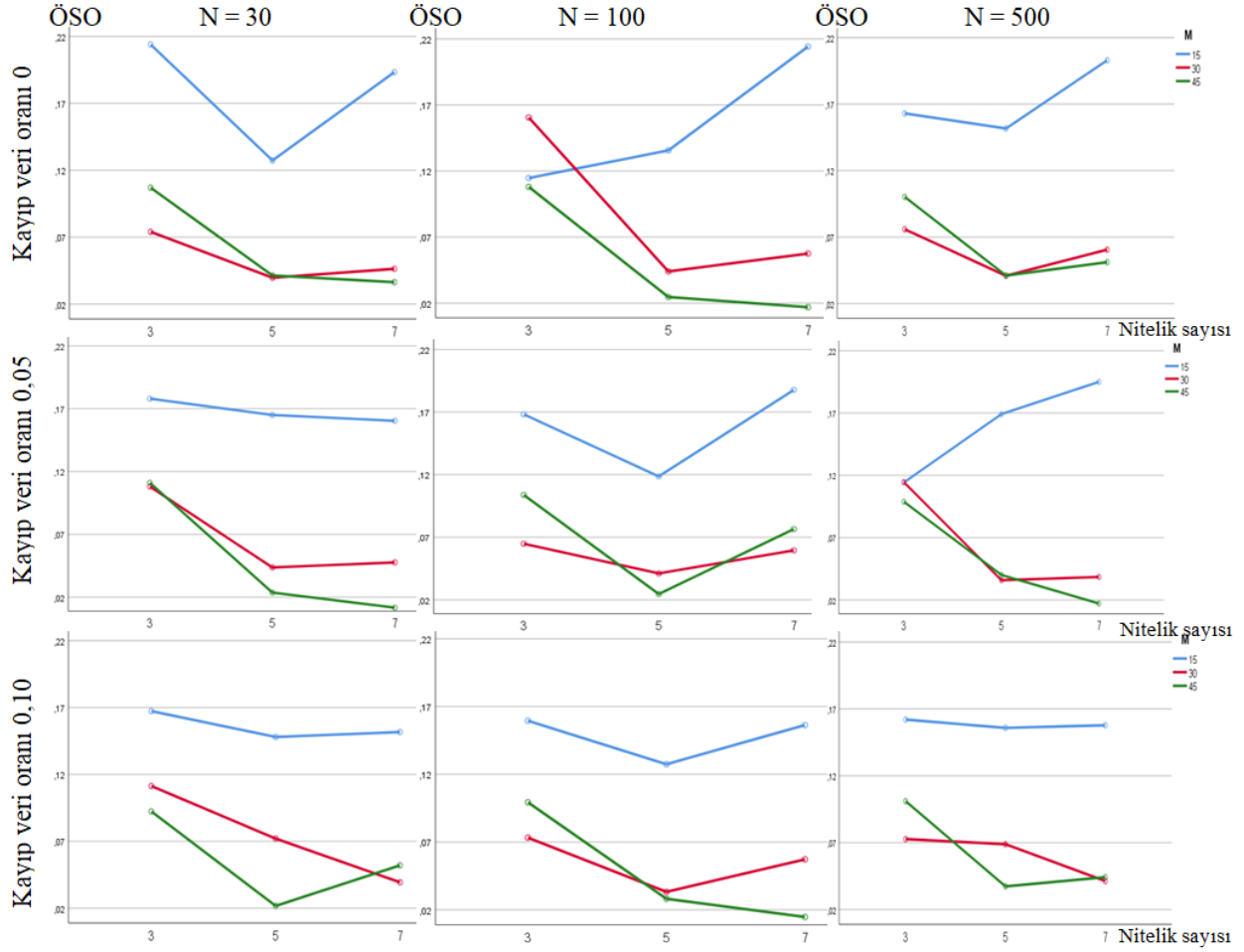
Tablo 12

Değişen koşullar altında YSA'ya ait ÖSO'lar için faktöriyel ANOVA sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Ni	2,865	2	1,433	655,920	0,000	0,141
N	0,026	2	0,013	5,882	0,003	0,001
M	18,549	2	9,275	4246,666	0,000	0,514
KayıpVeri	0,074	2	0,037	16,920	0,000	0,004
Ni * N	0,337	4	0,084	38,589	0,000	0,019
Ni * M	1,885	4	0,471	215,786	0,000	0,097
Ni * KayıpVeri	0,126	4	0,031	14,392	0,000	0,007
M * N	0,083	4	0,021	9,551	0,000	0,005
N * KayıpVeri	0,054	4	0,013	6,153	0,000	0,003
M * KayıpVeri	0,041	4	0,010	4,682	0,001	0,002
Ni * N * M	0,195	8	0,024	11,185	0,000	0,011
Ni * N * KayıpVeri	0,169	8	0,021	9,653	0,000	0,010
Ni * M * KayıpVeri	0,292	8	0,037	16,724	0,000	0,016
M * N * KayıpVeri	0,369	8	0,046	21,114	0,000	0,021
Ni * N * M * KayıpVeri	1,161	16	0,073	33,220	0,000	0,062
Hata	17,513	8019	0,002			

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısı, KayıpVeri: Kayıp veri oranını temsil etmektedir.

Şekil 5 incelendiğinde YSA'nın en yüksek ortalama ÖSO'lara genel olarak madde sayısı 15 iken sahip olduğu görülmektedir. Şekil 5'ten nitelik sayısı arttıkça madde sayısı koşullarındaki ÖSO'ların önce azaldığı sonra arttığı görülmektedir. Yani nitelik sayısı üçten beşe arttırıldığında ortalama ÖSO'lar azalırken, nitelik sayısı beşten yediye arttırıldığında ortalama ÖSO'lar genel olarak artmaktadır. Buna bağlı olarak nitelik sayısı arttıkça YSA'ya ait ortalama ÖSO'ların önce azaldığı sonra arttığı söylenebilir. Şekil 5'ten kayıp veri oranındaki değişimin YSA'ya ait ortalama ÖSO'lara etkisi incelendiğinde, kayıp veri oranı arttıkça ortalama ÖSO'ların az bir miktar azaldığı söylenebilir. Son olarak, örneklem büyüklüğünün artmasıyla ortalama ÖSO'ların azalmaları, artmaları veya sabit kalmaları hakkında herhangi bir sistematikliğe rastlanmamıştır.



Şekil 5. Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında YSA'ya ait ortalama ÖSO'lar

4.1.3. DINO Modelinin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması

DINO modelinin analizi sonucu çeşitli koşullarda elde edilen sınıflandırma oranlarına ait detaylı bilgiler EK-7'de, ortalama RMSEA, sapma (deviance), loglikelihood, AIC ve BIC değerleri EK-8'de raporlanmıştır. Değişen koşullarda DINO modeline ait NSO ve ÖSO'lar Tablo 13'te özetlenmiştir.

Tablo 13 incelendiğinde DINO modeline ait NSO'ların 0,84 ile 0,99 arasında, ÖSO'ların ise 0,35 ile 0,97 arasında değiştiği görülmektedir. DINO modeline ait NSO'ların iyi ile mükemmel güvenilirlik ölçütleri arasında, ÖSO'ların ise zayıf ile mükemmel güvenilirlik ölçütleri arasında oldukları söylenebilir. DINO modeline ait NSO'ların tüm koşullarda 0,83

üzerindeyken ÖSO'ların güvenilirlik derecesi kabul edilebilir düzeyin üstünde (0,65 üzeri) olma durumlarının koşullara göre değiştiği görülmektedir.

Tablo 13

Değişen koşullarda DINO modeline ait NSO ve ÖSO'lar

Ni	N	M	NSO			ÖSO		
			Kayıp veri oranı			Kayıp veri oranı		
			%0	%5	%10	%0	%5	%10
3	30	15	0,945	0,933	0,928	0,850	0,819	0,806
		30	0,978	0,979	0,967	0,934	0,938	0,900
		45	0,989	0,984	0,969	0,966	0,951	0,911
	100	15	0,941	0,933	0,926	0,840	0,818	0,799
		30	0,981	0,983	0,972	0,948	0,953	0,920
		45	0,989	0,987	0,978	0,966	0,962	0,937
	500	15	0,947	0,940	0,929	0,852	0,833	0,805
		30	0,974	0,974	0,968	0,925	0,927	0,910
		45	0,988	0,985	0,982	0,963	0,956	0,948
5	30	15	0,895	0,885	0,880	0,588	0,568	0,539
		30	0,936	0,928	0,926	0,728	0,704	0,703
		45	0,957	0,943	0,961	0,803	0,746	0,829
	100	15	0,886	0,877	0,869	0,562	0,527	0,521
		30	0,943	0,939	0,928	0,764	0,745	0,728
		45	0,967	0,954	0,947	0,848	0,801	0,777
	500	15	0,901	0,891	0,882	0,607	0,586	0,559
		30	0,944	0,942	0,931	0,765	0,756	0,718
		45	0,965	0,961	0,955	0,847	0,828	0,809
7	30	15	0,854	0,854	0,843	0,375	0,367	0,347
		30	0,909	0,904	0,892	0,557	0,525	0,484
		45	0,938	0,934	0,929	0,665	0,648	0,620
	100	15	0,868	0,863	0,979	0,393	0,388	0,356
		30	0,917	0,909	0,900	0,575	0,547	0,517
		45	0,940	0,933	0,922	0,674	0,647	0,602
	500	15	0,874	0,865	0,857	0,415	0,395	0,380
		30	0,923	0,916	0,907	0,603	0,579	0,541
		45	0,949	0,944	0,937	0,712	0,688	0,658

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısını temsil etmektedir.

DINO modelinin güvenilirlik derecesi kabul edilebilir düzeyin üstünde ÖSO'lara sahip olduğu koşulların,

- Nitelik sayısı üç iken tüm örneklem büyüklüğü, madde sayısı ve kayıp veri oranı koşulları,
- Nitelik sayısı beş iken tüm örneklem büyüklükleri ve kayıp veri oranlarında madde sayısı 30 veya 45 olduğu,
- Nitelik sayısı yedi ve madde sayısı 45 iken örneklem büyüklüğü 30 ve 100 ise kayıp veri oranının sıfır olduğu, örneklem büyüklüğü 500 ise tüm kayıp veri oranı koşulları olduğu görülmektedir.

4.1.2.1. Değişen koşullar altında DINO modeline ait NSO'lar için faktöriyel ANOVA bulguları

DINO modeline ait NSO'ların araştırma kapsamında ele alınan faktörlerde ve bu faktörlerin koşullarında nasıl değiştiğinin belirlenmesi için 3x3x3x3 faktöriyel ANOVA yapılmıştır. Bulgular, ana ve etkileşim etkilerinin tablosu ve etkileşim grafiğiyle sunulmuştur. Ana ve etkileşim etkilerine ait Tablo 14 incelendiğinde nitelik ($F_{(2,8019)} = 13621,53$; $p < 0,01$) ve madde sayısı ($F_{(2,8019)} = 15049,99$; $p < 0,01$), örneklem büyüklüğü ($F_{(2,8019)} = 175,51$; $p < 0,01$) ile kayıp veri oranı ($F_{(2,8019)} = 600,27$; $p < 0,01$) ana etkilerinin istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 14 incelendiğinde nitelik sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 43,00$; $p < 0,01$); nitelik ile madde sayısı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 279,51$; $p < 0,01$); nitelik sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 10,75$; $p < 0,01$); madde sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 16,61$; $p < 0,01$); örneklem büyüklüğü ile kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 6,07$; $p < 0,01$); madde sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 12,86$; $p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 14'ten faktörler arasındaki üçlü etkileşim etkileri incelendiğinde nitelik ve madde sayısı ile örneklem büyüklüğü faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 19,71$; $p < 0,01$); nitelik sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 10,88$; $p < 0,01$); nitelik ve madde sayısı ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 6,78$; $p < 0,01$) ve madde sayısı, örneklem büyüklüğü ile kayıp veri

oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 2,95$; $p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 14

Değişen koşullar altında DINO modeline ait NSO'lar için faktöriyel ANOVA sonuçları

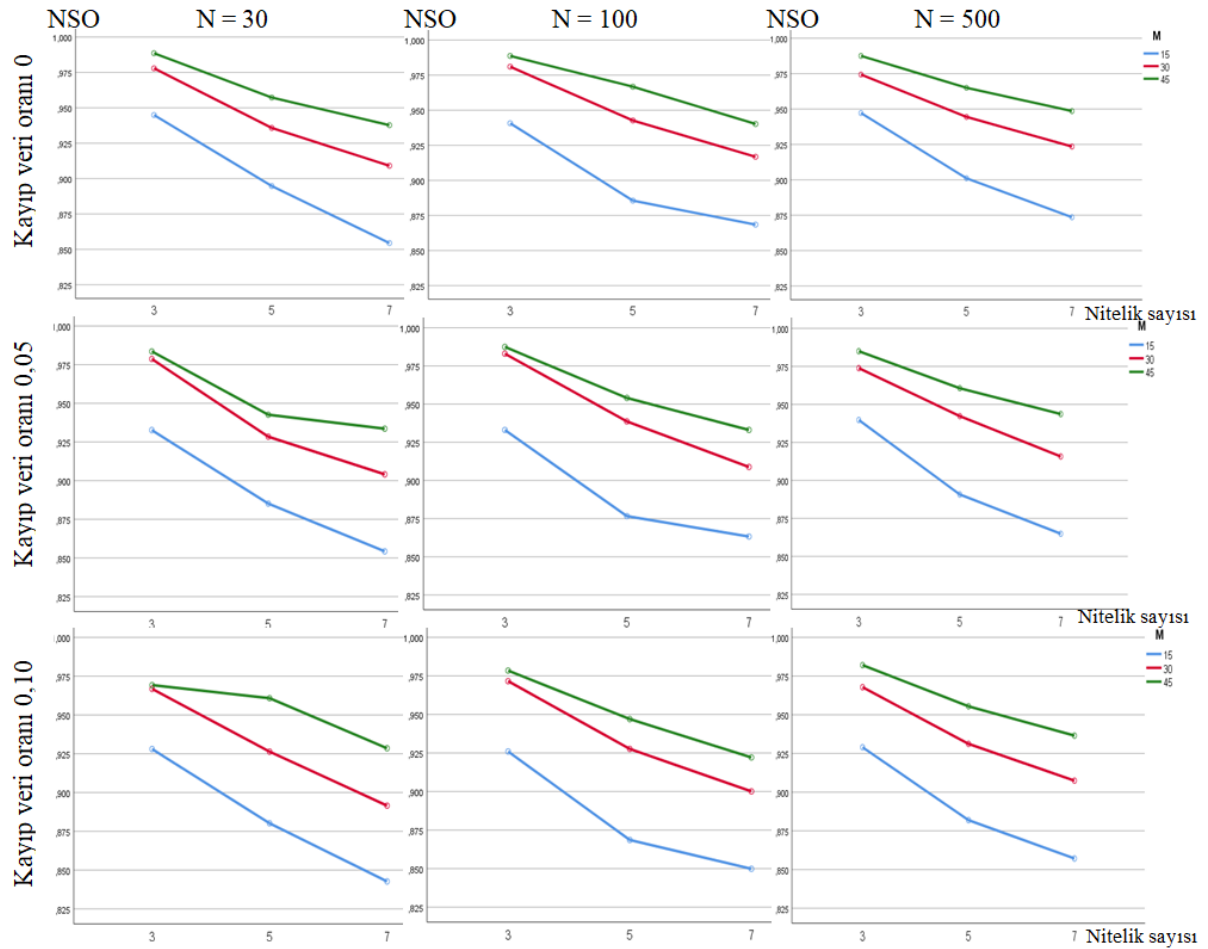
Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Ni	5,556	2	2,778	13621,534	0,000	0,773
N	0,072	2	0,036	175,506	0,000	0,042
M	6,139	2	3,069	15049,987	0,000	0,790
KayıpVeri	0,245	2	0,122	600,270	0,000	0,130
Ni * N	0,035	4	0,009	42,998	0,000	0,021
Ni * M	0,228	4	0,057	279,510	0,000	0,122
Ni * KayıpVeri	0,009	4	0,002	10,748	0,000	0,005
M * N	0,014	4	0,003	16,610	0,000	0,008
N * KayıpVeri	0,005	4	0,001	6,067	0,000	0,003
M * KayıpVeri	0,010	4	0,003	12,861	0,000	0,006
Ni * N * M	0,032	8	0,004	19,710	0,000	0,019
Ni * N * KayıpVeri	0,018	8	0,002	10,879	0,000	0,011
Ni * M * KayıpVeri	0,011	8	0,001	6,784	0,000	0,007
M * N * KayıpVeri	0,005	8	0,001	2,952	0,003	0,003
Ni * N * M * KayıpVeri	0,014	16	0,001	4,194	0,000	0,008
Hata	1,635	8019	0,0002			

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısı, KayıpVeri: Kayıp veri oranını temsil etmektedir.

Son olarak, Tablo 14 incelendiğinde dörtlü etkileşim etkilerinin, yani nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(2,8019)} = 4,19$; $p < 0,01$) ve bu etkileşimin küçük etkiye sahip olduğu görülmektedir ($\eta_p^2 = 0,008$). Nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı koşullarının etkileşimlerine ait grafik Şekil 6'da verilmiştir.

Şekil 6 incelendiğinde tüm kayıp veri oranı, örneklem büyüklüğü ve nitelik sayısı koşullarında DINO modeline ait en yüksek ortalama NSO'lara madde sayısı 45 iken, en düşük ortalama NSO'lara ise madde sayısı 15 iken sahip olduğu görülmektedir. Buradan hareketle madde sayısı arttıkça DINO modeline ait ortalama NSO'ların arttığı söylenebilir. Şekil 6'dan nitelik sayısı arttıkça madde sayısı koşullarındaki ortalama NSO'ların azaldığı

görülmektedir. Buna bağlı olarak nitelik sayısı arttıkça DINO modeline ait ortalama NSO'ların azaldığı söylenebilir. Şekil 6'dan kayıp veri oranındaki değişimin DINO modeline ait ortalama NSO'lara etkisi incelendiğinde, kayıp veri oranı arttıkça ortalama NSO'ların genel olarak azaldığı söylenebilir. Son olarak, Şekil 6'da örneklem büyüklüğünün artmasıyla ortalama NSO'ların genel olarak arttığı görülmektedir.



Şekil 6. Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında DINO modeline ait ortalama NSO'lar

4.1.3.2. Değişen koşullar altında DINO modeline ait ÖSO'lar için faktöriyel ANOVA bulguları

DINO modeline ait ÖSO'ların araştırma kapsamında ele alınan faktörlerde ve bu faktörlerin koşullarında nasıl değiştiğinin belirlenmesi için 3x3x3 faktöriyel ANOVA yapılmıştır. Bulgular, ana ve etkileşim etkilerinin tablosu ve etkileşim grafiğiyle sunulmuştur. Ana ve etkileşim etkilerine ait Tablo 15 incelendiğinde nitelik ($F_{(2,8019)} = 37138,04$; $p < 0,01$) ve madde sayısı ($F_{(2,8019)} = 13116,71$; $p < 0,01$), örneklem büyüklüğü

($F_{(2,8019)} = 180,69$; $p < 0,01$) ve kayıp veri oranı ($F_{(2,8019)} = 443,56$; $p < 0,01$) ana etkilerinin istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 15

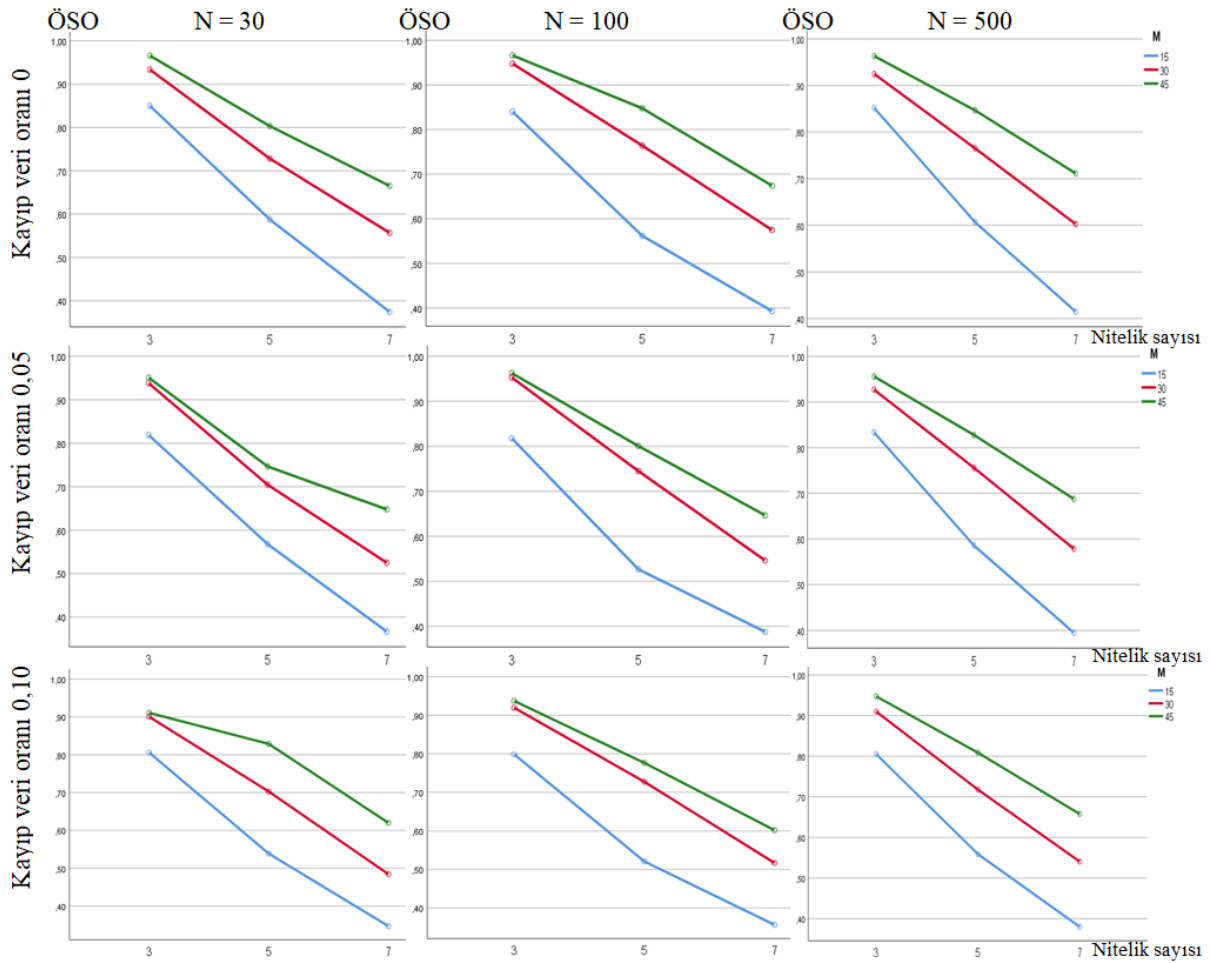
Değişen koşullar altında DINO modeline ait ÖSO'lar için faktöriyel ANOVA sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Ni	188,471	2	94,235	37138,044	0,000	0,903
N	0,917	2	0,458	180,685	0,000	0,043
M	66,566	2	33,283	13116,710	0,000	0,766
KayıpVeri	2,251	2	1,126	443,558	0,000	0,100
Ni * N	0,395	4	0,099	38,914	0,000	0,019
Ni * M	5,886	4	1,471	579,881	0,000	0,224
Ni * KayıpVeri	0,179	4	0,045	17,654	0,000	0,009
M * N	0,244	4	0,061	24,050	0,000	0,012
N * KayıpVeri	0,044	4	0,011	4,309	0,002	0,002
M * KayıpVeri	0,043	4	0,011	4,189	0,002	0,002
Ni * N * M	0,271	8	0,034	13,341	0,000	0,013
Ni * N * KayıpVeri	0,182	8	0,023	8,976	0,000	0,009
Ni * M * KayıpVeri	0,216	8	0,027	10,629	0,000	0,010
M * N * KayıpVeri	0,092	8	0,012	4,547	0,000	0,005
Ni * N * M * KayıpVeri	0,263	16	0,016	6,473	0,000	0,013
Hata	20,348	8019	0,003			

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğü, M: Madde sayısı, KayıpVeri: Kayıp veri oranını temsil etmektedir.

Tablo 15 incelendiğinde nitelik sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 38,91$; $p < 0,01$); nitelik ile madde sayısı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 579,88$; $p < 0,01$); nitelik sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 17,65$; $p < 0,01$); madde sayısı ile örneklem büyüklüğü faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 24,05$; $p < 0,01$); örneklem büyüklüğü ile kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 4,31$; $p < 0,01$); madde sayısı ve kayıp veri oranı faktörlerinin arasındaki etkileşimin ($F_{(4,8019)} = 4,19$; $p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 15'ten faktörler arasındaki üçlü etkileşim etkileri incelendiğinde nitelik ve madde sayısı ile örneklem büyüklüğü faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 13,34; p < 0,01$); nitelik sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 8,98; p < 0,01$); nitelik ve madde sayısı ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 10,63; p < 0,01$) ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin ($F_{(2,8019)} = 4,55; p < 0,01$) istatistiksel olarak anlamlı olduğu görülmektedir.



Şekil 7. Kayıp veri oranı, nitelik ve madde sayısı ile örneklem büyüklüğü koşullarında DINO modeline ait ortalama ÖSO'lar

Son olarak, Tablo 15 incelendiğinde dördümlü etkileşim etkilerinin, yani nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı faktörleri arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(2,8019)} = 6,47; p < 0,01$) ve bu etkileşimin küçük etkiye sahip olduğu görülmektedir ($\eta_p^2 = 0,01$). Nitelik ve madde sayısı, örneklem büyüklüğü ile kayıp veri oranı koşullarının etkileşimlerine ait grafik Şekil 7'de verilmiştir.

Şekil 7 incelendiğinde tüm kayıp veri oranı, örneklem büyüklüğü ve nitelik sayısı koşullarında DINO modeline ait en yüksek ortalama ÖSO'lara madde sayısı 45 iken, en düşük ortalama ÖSO'lara ise madde sayısı 15 iken sahip olduğu görülmektedir. Buradan hareketle madde sayısı arttıkça DINO modeline ait ortalama ÖSO'ların arttığı söylenebilir. Şekil 7'den nitelik sayısı arttıkça madde sayısı koşullarındaki ortalama ÖSO'ların azaldığı görülmektedir. Buna bağlı olarak nitelik sayısı arttıkça DINO modeline ait ortalama ÖSO'ların azaldığı söylenebilir. Şekil 7'den kayıp veri oranındaki değişimin DINO modeline ait ortalama ÖSO'lara etkisi incelendiğinde, kayıp veri oranı arttıkça ortalama ÖSO'ların genel olarak azaldığı söylenebilir. Son olarak, Şekil 7'de örneklem büyüklüğünün artmasıyla ortalama ÖSO'ların genel olarak arttığı görülmektedir.

4.2. Simülasyon Verisinde Nitelik Sayısı (3, 5 ve 7), Örneklem Büyüklüğü (30, 100 ve 500), Madde Sayısı (15, 30 ve 45) ve Kayıp Veri Oranı (0; 0,05 ve 0,10) Koşullarında POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması

Bu araştırma probleminin cevaplanmasında her bir alt problem için 3x3 faktöriyel ANOVA yapılmıştır. Bulgular ilgili başlıklar altında sunulmuştur.

4.2.1. Nitelik Sayılarına (3, 5 ve 7) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması

NSO'ların POBT, YSA ve DINO modeli ile nitelik sayısı faktörü ve bu faktörün koşulları (3, 5 ve 7) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 16'da, basit etkileri Tablo 17'de sunulmuştur. Tablo 16 incelendiğinde nitelik sayısı faktörünün ve modellerin ana etkilerinin istatistiksel olarak anlamlı farklılık gösterdiği görülmektedir (sırasıyla $F_{(2,24291)} = 153,13$; $F_{(2,24291)} = 89683,31$; $p < 0,01$). Etki büyüklükleri incelendiğinde nitelik sayısının küçük etkiye ($\eta^2 = 0,01$), modellerin ise geniş etkiye ($\eta^2 = 0,88$) sahip oldukları görülmektedir.

Tablo 16

Nitelik Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Nitelik sayısı	1,018	2	0,509	153,127	0,000	0,012
Model	595,986	2	297,993	89683,310	0,000	0,881
Nitelik sayısı * Model	52,177	4	13,044	3925,781	0,000	0,393
Hata	80,712	24291	0,003			

Tablo 16 incelendiğinde nitelik sayısı ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,8019)} = 3925,78$; $p < 0,01$) ve bu etkileşimin orta düzey etkiye ($\eta_p^2 = 0,39$) sahip olduğu görülmektedir. Nitelik sayısı ile modeller arasındaki etkileşimin anlamlı olması nedeniyle nitelik sayısı faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir.

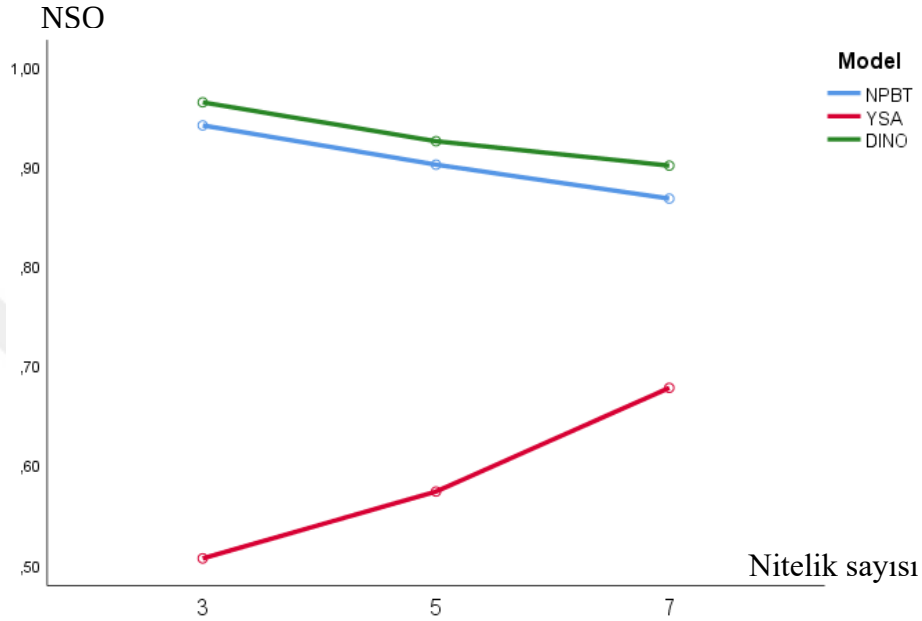
Tablo 17

Nitelik Sayısı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin NSO'larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	F	p	Anlamlı Fark
3 Nitelik						DINO-POBT
Modeller	360,254	2	180,127	54210,649	0,000	DINO-YSA
Hata	80,712	24291	0,003			POBT-YSA
5 Nitelik						DINO-POBT
Modeller	209,406	2	104,703	31511,120	0,000	DINO-YSA
Hata	80,712	24291	0,003			POBT-YSA
7 Nitelik						DINO-POBT
Modeller	78,503	2	39,252	11813,103	0,000	DINO-YSA
Hata	80,712	24291	0,003			POBT-YSA

Basit etkiler analizine ait Tablo 17 incelendiğinde nitelik sayısı üç, beş ve yedi olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 54210,65$; $F_{(2,8019)} = 31511,12$; $F_{(2,8019)} = 11813,10$; $p < 0,01$). Ek olarak, Tablo 17'de POBT, YSA ve DINO modeline ait NSO'lar arasındaki en büyük farkın nitelik sayısı üç iken, en az farkın ise nitelik sayısı yedi iken olduğu görülmektedir.

Nitelik sayısı koşulları ile modellerin etkileşimlerine ait grafik Şekil 8’de verilmiştir. Şekil 8 incelendiğinde nitelik sayısı arttıkça POBT’ye ve DINO modeline ait NSO’ların azaldığı YSA’ya ait NSO’ların ise arttığı görülmektedir. Tüm nitelik sayısı koşullarında en yüksek ortalama NSO’ya DINO modelinin, en düşük ortalama NSO’ya ise YSA’nın sahip olduğu görülmektedir.



Şekil 8. Nitelik sayısının 3, 5 ve 7 olduğu koşullarda POBT’nin, YSA’nın ve DINO modelinin NSO’ları

ÖSO’ların POBT, YSA ve DINO modeli ile nitelik sayısı faktörü ve bu faktörün koşulları (3, 5 ve 7) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 18’de, basit etkileri ise Tablo 17’de sunulmuştur. Tablo 17 incelendiğinde ÖSO’ların nitelik sayısına ve modele göre istatistiksel olarak anlamlı farklılık gösterdiği görülmektedir (sırasıyla $F_{(2,24291)} = 12421,70$; $F_{(2,24291)} = 77570,81$; $p < 0,01$). Etki büyüklükleri incelendiğinde nitelik sayısının orta düzey etkiye ($\eta^2 = 0,51$), modellerin ise geniş etkiye ($\eta^2 = 0,87$) sahip olduğu görülmektedir.

Tablo 18 incelendiğinde nitelik sayısı ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,8019)} = 2453,27$; $p < 0,01$) ve bu etkileşimin küçük etkiye ($\eta_p^2 = 0,29$) sahip olduğu görülmektedir. Nitelik sayısı ile modeller arasındaki etkileşimin anlamlı olması nedeniyle nitelik sayısı faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir.

Tablo 18

Nitelik Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Nitelik sayısı	295,924	2	147,962	12421,698	0,000	0,506
Model	1847,983	2	923,991	77570,806	0,000	0,865
Nitelik sayısı * Model	116,889	4	29,222	2453,265	0,000	0,288
Hata	289,344	24291	0,012			

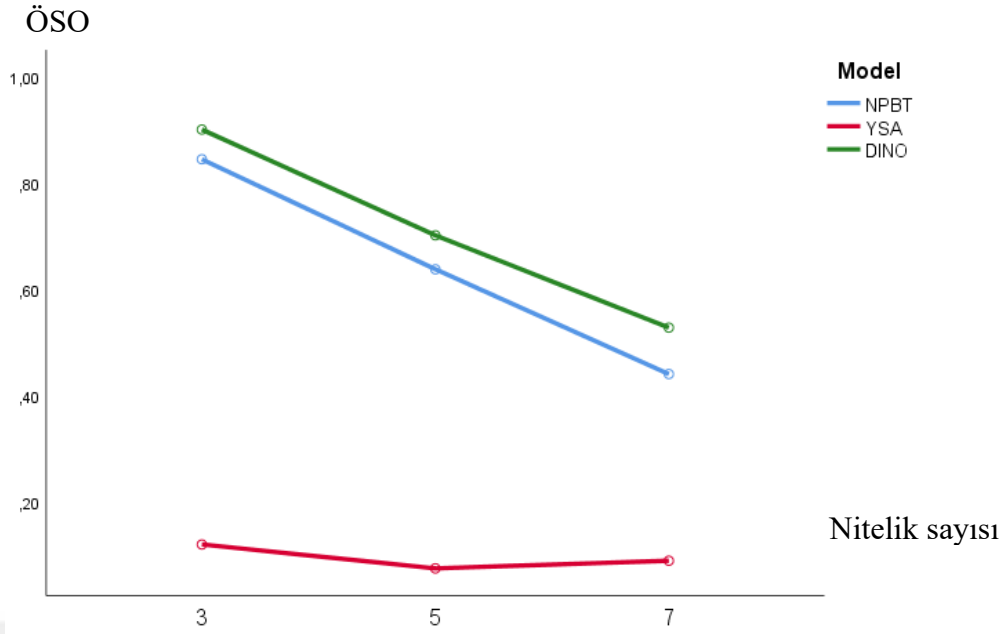
Basit etkiler analizine ait bulgulara Tablo 19'da yer verilmiştir. Basit etkiler analizine ait Tablo 19 incelendiğinde nitelik sayısı üç, beş ve yedi olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 43151,60$; $F_{(2,8019)} = 27059,06$; $F_{(2,8019)} = 12266,67$; $p < 0,01$). Ek olarak, Tablo 19'da POBT, YSA ve DINO modeline ait ÖSO'lar arasındaki en büyük farkın nitelik sayısı üç iken, en az farkın ise nitelik sayısı yedi iken olduğu görülmektedir.

Tablo 19

Nitelik Sayısı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	F	p	Anlamlı Fark
3 Nitelik						DINO-POBT
Modeller	1028,008	2	514,004	43151,603	0,000	DINO-YSA
Hata	289,344	24291	0,012			POBT-YSA
5 Nitelik						DINO-POBT
Modeller	644,633	2	322,316	27059,064	0,000	DINO-YSA
Hata	289,344	24291	0,012			POBT-YSA
7 Nitelik						DINO-POBT
Modeller	292,231	2	146,116	12266,669	0,000	DINO-YSA
Hata	289,344	24291	0,012			POBT-YSA

Nitelik sayısı koşulları ile modellerin etkileşimlerine ait Şekil 9 incelendiğinde nitelik sayısı arttıkça POBT, YSA ve DINO modelin ÖSO'larının azaldığı görülmektedir. Tüm nitelik sayısı koşullarında en yüksek ortalama NSO'ya DINO modelinin, en düşük ortalama NSO'ya ise YSA'nın sahip olduğu görülmektedir.



Şekil 9. Nitelik sayısının 3, 5 ve 7 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları

4.2.2. Örneklem Büyüklüklerine (30, 100 ve 500) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması

NSO'ların POBT, YSA ve DINO model ile örneklem büyüklüğü faktörü ve bu faktörün koşulları (30, 100 ve 500) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 20'de, basit etkileri Tablo 21'de sunulmuştur. Tablo 20 incelendiğinde NSO'ların örneklem büyüklüğüne göre istatistiksel olarak anlamlı farklılık göstermezken ($F_{(2,24291)} = 4,66$; $p = 0,01$) modele göre istatistiksel olarak anlamlı farklılık ($F_{(2,24291)} = 54159,50$; $p < 0,01$) gösterdiği görülmektedir. Ek olarak, modellerin geniş etkiye ($\eta^2 = 0,82$) sahip olduğu görülmektedir.

Tablo 20

Örneklem Büyüklüğü Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Örneklem büyüklüğü	0,051	2	0,026	4,655	0,010	0,000
Model	595,986	2	297,993	54159,499	0,000	0,817
Örneklem büyüklüğü * Model	0,203	4	0,051	9,244	0,000	0,002
Hata	133,652	24291	0,006			

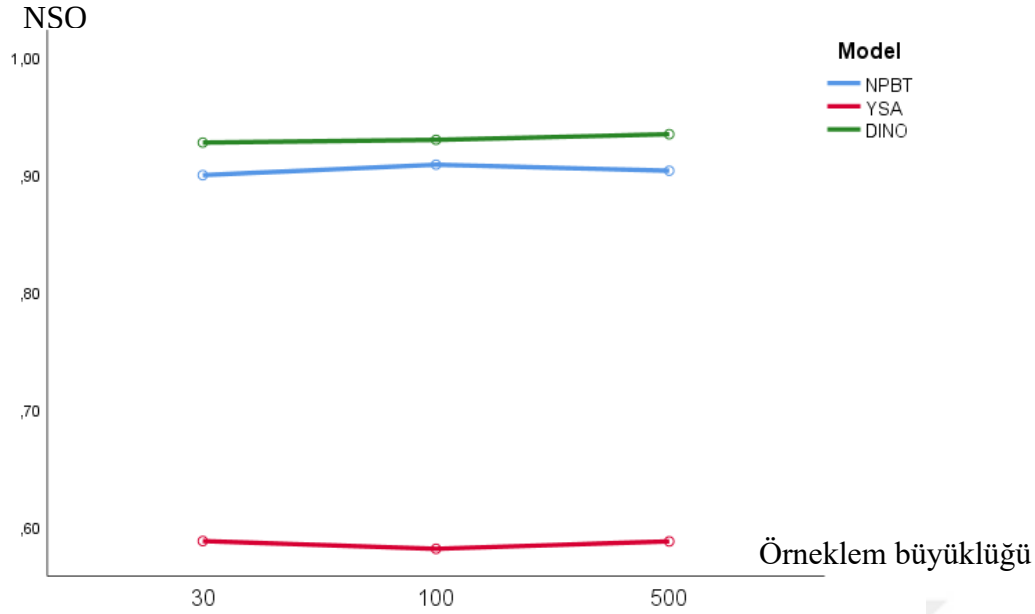
Tablo 20 incelendiğinde örneklem büyüklüğü ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,24291)} = 9,24$; $p < 0,01$) ve bu etkileşimin küçük etkiye ($\eta_p^2 = 0,002$) sahip olduğu görülmektedir. Örneklem büyüklüğü ile modeller arasındaki etkileşimin anlamlı olması nedeniyle örneklem büyüklüğü faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir.

Tablo 21

Örneklem Büyüklüğü Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin NSO'larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	<i>F</i>	<i>p</i>	Anlamlı Fark
30 Örneklem						DINO-POBT
Modeller	191,686	2	95,843	17419,223	0,000	DINO-YSA
Hata	133,652	24291	0,006			POBT-YSA
100 Örneklem						DINO-POBT
Modeller	205,804	2	102,902	18702,193	0,000	DINO-YSA
Hata	133,652	24291	0,006			POBT-YSA
500 Örneklem						DINO-POBT
Modeller	198,700	2	99,350	18056,570	0,000	DINO-YSA
Hata	133,652	24291	0,006			POBT-YSA

Basit etkiler analizine ait Tablo 21 incelendiğinde örneklem büyüklüğü 30, 100 ve 500 olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 17419,22$; $F_{(2,8019)} = 18702,19$; $F_{(2,8019)} = 18056,57$; $p < 0,01$). Ek olarak, Tablo 22'de POBT, YSA ve DINO modeline ait NSO'lar arasındaki en büyük farkın örneklem büyüklüğü 100 iken, en az farkın ise örneklem büyüklüğü 30 iken olduğu görülmektedir. Örneklem büyüklüğü koşulları ile modellerin etkileşimlerine ait grafik Şekil 10'da verilmiştir. Şekil 10 incelendiğinde örneklem büyüklüğü arttıkça POBT'ye, YSA'ya ve DINO modeline ait NSO'ların değişmediği söylenebilir. Tüm örneklem büyüklüğü koşullarında en yüksek ortalama NSO'ya DINO modelinin, en düşük ortalama NSO'ya ise YSA'nın sahip olduğu görülmektedir.



Şekil 10. Örneklem büyüklüğünün 30, 100 ve 500 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin NSO'ları

ÖSO'ların POBT, YSA ve DINO modeli ile örneklem büyüklüğü faktörü ve bu faktörün koşulları (30, 100 ve 500) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 22'de, basit etkileri ise Tablo 23'te sunulmuştur. Tablo 22 incelendiğinde ÖSO'ların örneklem büyüklüğüne ve modele göre istatistiksel olarak anlamlı farklılık gösterdiği görülmektedir (sırasıyla $F_{(2,24291)} = 16,85$; $F_{(2,24291)} = 32094,84$; $p < 0,01$). Etki büyüklükleri incelendiğinde örneklem büyüklüğünün küçük etkiye ($\eta^2 = 0,001$), modellerin orta düzey etkiye ($\eta^2 = 0,73$) sahip oldukları görülmektedir.

Tablo 22

Örneklem Büyüklüğü Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Örneklem büyüklüğü	0,970	2	0,485	16,849	0,000	0,001
Model	1847,983	2	923,991	32094,844	0,000	0,725
Örneklem büyüklüğü * Model	1,864	4	0,466	16,188	0,000	0,003
Hata	699,323	24291	0,029			

Tablo 22 incelendiğinde örneklem büyüklüğü ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,24291)} = 16,19$; $p < 0,01$) ve bu etkileşimin küçük etkiye ($\eta_p^2 = 0,003$) sahip olduğu görülmektedir. Örneklem büyüklüğü ile modeller

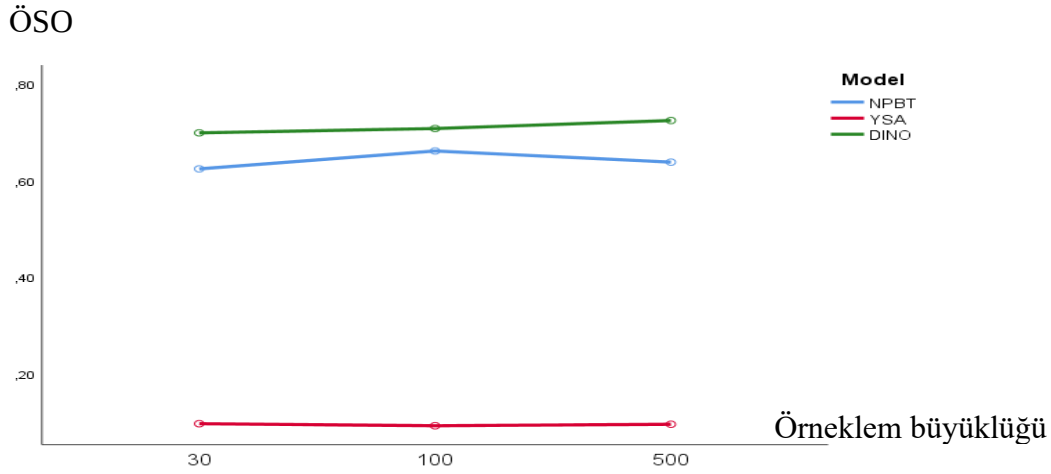
arasındaki etkileşimin anlamlı olması nedeniyle örneklem büyüklüğü faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir.

Tablo 23

Örneklem Büyüklüğü Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	F	p	Anlamlı Fark
30 Örneklem						DINO-POBT
Modeller	583,817	2	291,909	10139,445	0,000	DINO-YSA
Hata	699,323	24291	0,029			POBT-YSA
100 Örneklem						DINO-POBT
Modeller	636,068	2	318,034	11046,914	0,000	DINO-YSA
Hata	699,323	24291	0,029			POBT-YSA
500 Örneklem						DINO-POBT
Modeller	629,962	2	314,981	10940,861	0,000	DINO-YSA
Hata	699,323	24291	0,029			POBT-YSA

Basit etkiler analizine ait Tablo 23 incelendiğinde örneklem büyüklüğü 30, 100 ve 500 olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 10139,45$; $F_{(2,8019)} = 11046,91$; $F_{(2,8019)} = 10940,86$; $p < 0,01$). Ek olarak, Tablo 23'te POBT, YSA ve DINO modeline ait ÖSO'lar arasındaki en büyük farkın örneklem büyüklüğü 100 iken, en az farkın ise örneklem büyüklüğü 30 iken olduğu görülmektedir.



Şekil 11. Örneklem büyüklüğünün 30, 100 ve 500 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları

Örneklem büyüklüğü koşulları ile modellerin etkileşimlerine ait grafik Şekil 11'de verilmiştir. Şekil 11 incelendiğinde örneklem büyüklüğü 30 örneklemden 100 örnekleme arttırıldığında POBT'nin ve DINO modelin ÖSO'ları artarken YSA'nın ÖSO'larının azaldığı görülmektedir. Örneklem büyüklüğü 500 olduğunda ise POBT'nin ÖSO'larının azaldığı, YSA'nın ÖSO'larının değişmediği, DINO modelin ÖSO'larının ise arttığı görülmektedir. Ek olarak, tüm örneklem büyüklüğü koşullarında en yüksek ortalama ÖSO'ya DINO modelinin, en düşük ortalama ÖSO'ya ise YSA'nın sahip olduğu görülmektedir.

4.2.3. Madde Sayılarına (15, 30 ve 45) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması

NSO'ların POBT, YSA ve DINO modeli ile madde sayısı faktörü ve bu faktörün koşulları (15, 30 ve 45) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 24'te, basit etkileri Tablo 25'te sunulmuştur. Tablo 24 incelendiğinde NSO'larının madde sayısına ve modele göre istatistiksel olarak anlamlı farklılık gösterdiği görülmektedir (sırasıyla $F_{(2,24291)} = 89,08$; $F_{(2,24291)} = 86883,58$; $p < 0,01$). Etki büyüklükleri incelendiğinde madde sayısının küçük etkiye ($\eta^2 = 0,007$), modellerin ise geniş etkiye ($\eta^2 = 0,88$) sahip olduğu görülmektedir.

Tablo 24

Madde Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Madde sayısı	0,611	2	0,306	89,075	0,000	0,007
Model	595,986	2	297,993	86883,579	0,000	0,877
Madde sayısı * Model	49,983	4	12,496	3643,283	0,000	0,375
Hata	83,313	24291	0,003			

Tablo 24 incelendiğinde madde sayısı ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,8019)} = 3643,28$; $p < 0,01$) ve bu etkileşimin orta düzey etkiye ($\eta_p^2 = 0,38$) sahip olduğu görülmektedir. Madde sayısı ile modeller arasındaki etkileşimin anlamlı olması nedeniyle madde sayısı faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir. Basit etkiler analizine ait Tablo 25 incelendiğinde madde sayısı 15, 30 ve 45 olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 10952,18$; $F_{(2,8019)} = 35315,23$; $F_{(2,8019)} = 47902,74$; $p < 0,01$). Ek olarak, Tablo 25'te POBT, YSA ve DINO modeline ait NSO'lar arasındaki en büyük farkın madde sayısı 45 iken, en az farkın ise madde sayısı 15 iken olduğu görülmektedir.

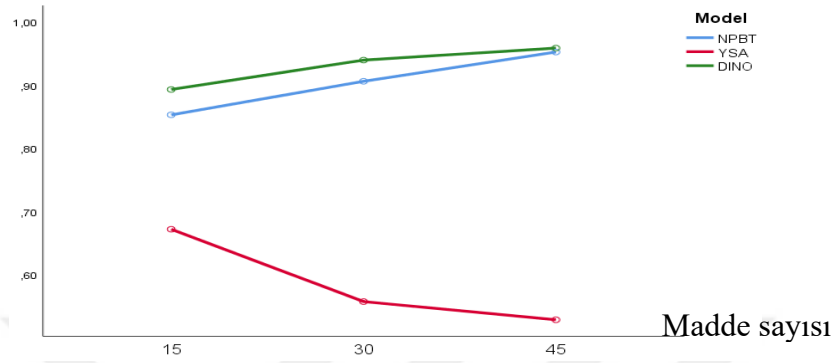
Tablo 25

Madde Sayısı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin NSO'larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	F	p	Anlamlı Fark
15 Madde						DINO-POBT
Modeller	75,128	2	37,564	10952,184	0,000	DINO-YSA
Hata	83,313	24291	0,003			POBT-YSA
30 Madde						DINO-POBT
Modeller	242,248	2	121,124	35315,225	0,000	DINO-YSA
Hata	83,313	24291	0,003			POBT-YSA
45 Madde						DINO-POBT
Modeller	328,593	2	164,297	47902,736	0,000	DINO-YSA
Hata	83,313	24291	0,003			POBT-YSA

Madde sayısı koşulları ile modellerin etkileşimlerine ait grafik Şekil 12'de verilmiştir. Şekil 12 incelendiğinde madde sayısı arttıkça POBT ve DINO modeline ait NSO'ların

artmakta, YSA'ya ait NSO'ların ise azalmakta olduğu görülmektedir. Madde 15 ve 30 iken en yüksek ortalama NSO'ya DINO modeli sahipken, en düşük ortalama NSO'ya YSA sahiptir. Madde sayısı 45 olduğunda ise en yüksek ortalama NSO'ya DINO modeli ve POBT sahipken, en düşük ortalama NSO'ya YSA sahiptir.



Şekil 12. Madde sayısının 15, 30 ve 45 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin NSO'ları

ÖSO'ların POBT, YSA ve DINO model ile madde sayısı faktörü ve bu faktörün koşulları (15, 30 ve 45) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 26'da, basit etkileri ise Tablo 27'de sunulmuştur. Tablo 26 incelendiğinde ÖSO'ların madde sayısına ve modele göre istatistiksel olarak anlamlı farklılık gösterdiği görülmektedir (sırasıyla $F_{(2,24291)} = 1621,85$; $F_{(2,24291)} = 43753,11$; $p < 0,01$). Etki büyüklükleri incelendiğinde madde sayısının küçük etkiye ($\eta^2 = 0,12$), modellerin ise orta düzey etkiye ($\eta^2 = 0,78$) sahip olduğu görülmektedir.

Tablo 26

Madde Sayısı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Madde sayısı	68,501	2	34,251	1621,845	0,000	0,118
Model	1847,983	2	923,991	43753,111	0,000	0,783
Madde sayısı * Model	120,672	4	30,168	1428,522	0,000	0,190
Hata	512,985	24291	0,021			

Tablo 26 incelendiğinde madde sayısı ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,8019)} = 1428,52$; $p < 0,01$) ve bu etkileşimin küçük etkiye ($\eta_p^2 = 0,19$) sahip olduğu görülmektedir. Madde sayısı ile modeller arasındaki etkileşimin anlamlı

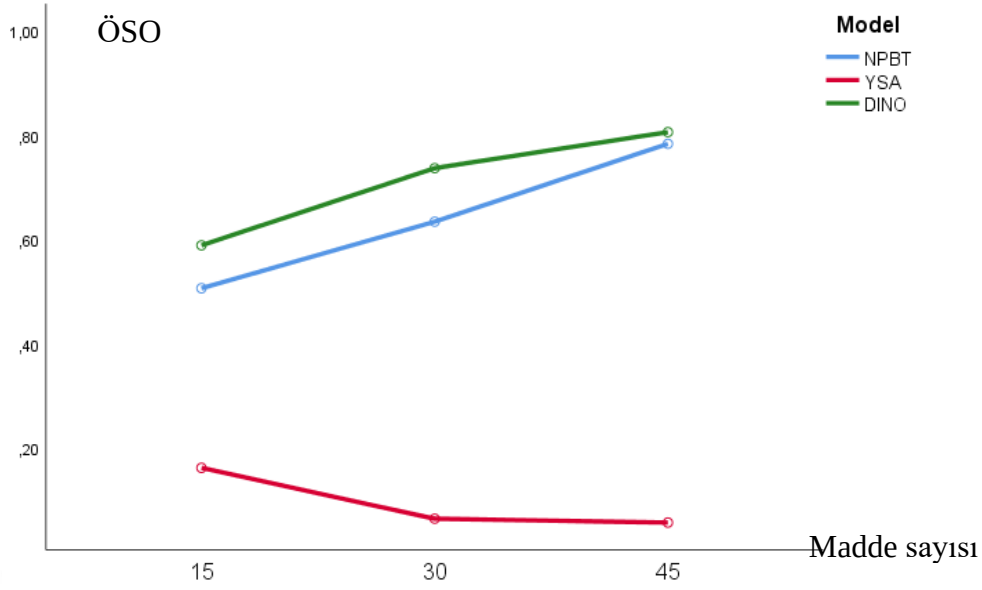
olması nedeniyle madde sayısı faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir. Basit etkiler analizine ait Tablo 27 incelendiğinde madde sayısı 15, 30 ve 45 olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 6564,78$; $F_{(2,8019)} = 16802,86$; $F_{(2,8019)} = 23242,52$; $p < 0,01$). Ek olarak, Tablo 27’de POBT, YSA ve DINO modeline ait ÖSO’lar arasındaki en büyük farkın madde sayısı 45 iken, en az farkın ise madde sayısı 15 iken olduğu görülmektedir.

Tablo 27

Madde Sayısı Faktörünün Her Bir Koşulunda POBT’nin, YSA’nın ve DINO Modelinin ÖSO’larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	F	p	Anlamlı Fark
15 Madde						DINO-POBT
Modeller	277,274	2	138,637	6564,776	0,000	DINO-YSA
Hata	512,985	24291	0,021			POBT-YSA
30 Madde						DINO-POBT
Modeller	709,696	2	354,848	16802,857	0,000	DINO-YSA
Hata	512,985	24291	0,021			POBT-YSA
45 Madde						DINO-POBT
Modeller	981,685	2	490,843	23242,523	0,000	DINO-YSA
Hata	512,985	24291	0,021			POBT-YSA

Madde sayısı koşulları ile modellerin etkileşimlerine ait grafik Şekil 13’te verilmiştir. Şekil 13 incelendiğinde madde sayısı arttıkça POBT’ye ve DINO modeline ait ÖSO’larının artmakta, YSA’ya ait ÖSO’ların ise azalmakta olduğu görülmektedir. Tüm nitelik sayısı koşullarında en yüksek ortalama ÖSO’ya DINO modelinin, en düşük ortalama ÖSO’ya ise YSA’nın sahip olduğu görülmektedir.



Şekil 13. Madde sayısının 15, 30 ve 45 madde olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları

4.2.4. Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Karşılaştırılması

NSO'ların POBT, YSA ve DINO modeli ile kayıp veri oranı faktörü ve bu faktörün koşulları (3, 5 ve 7) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 28'de, basit etkileri Tablo 29'da sunulmuştur. Tablo 28 incelendiğinde kayıp veri oranı faktörünün ve modellerin ana etkilerinin istatistiksel olarak anlamlı farklılık gösterdiği görülmektedir (sırasıyla $F_{(2,24291)} = 38,61$; $F_{(2,24291)} = 54266,06$; $p < 0,01$). Etki büyüklükleri incelendiğinde kayıp veri oranı faktörünün küçük etkiye ($\eta^2 = 0,003$), modellerin ise geniş etkiye ($\eta^2 = 0,82$) sahip oldukları görülmektedir.

Tablo 28

Kayıp Veri Oranı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin NSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Kayıp veri oranı	0,424	2	0,212	38,614	0,000	0,003
Model	595,986	2	297,993	54266,062	0,000	0,817
Kayıp veri oranı * Model	0,093	4	0,023	4,235	0,002	0,001
Hata	133,390	24291	0,005			

Tablo 28 incelendiğinde kayıp veri oranı ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,8019)} = 4,24$; $p < 0,01$) ve bu etkileşimin küçük etkiye ($\eta_p^2 = 0,001$) sahip olduğu görülmektedir. Kayıp veri oranı ile modeller arasındaki etkileşimin anlamlı olması nedeniyle kayıp veri oranı faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir.

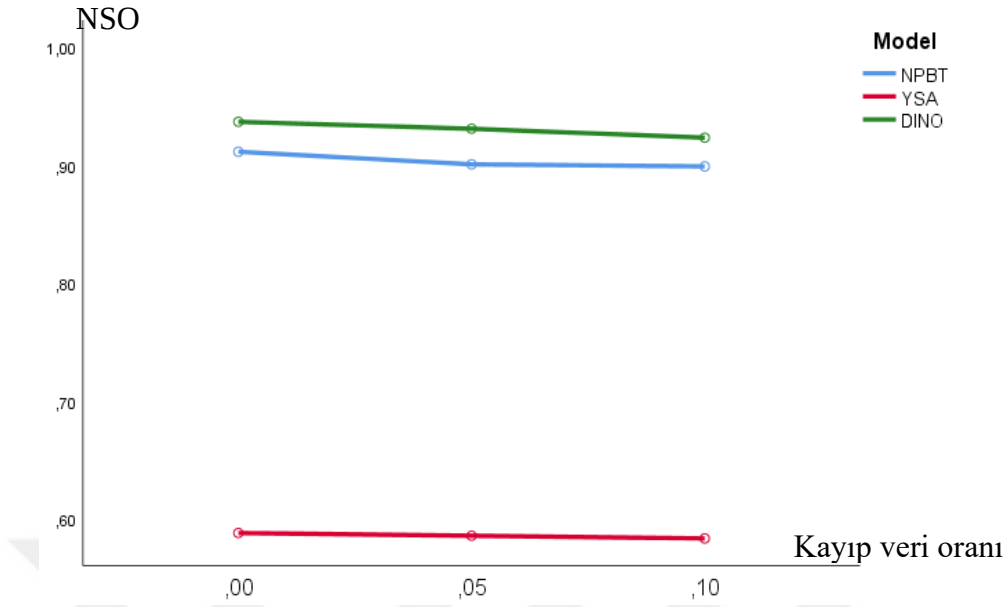
Basit etkiler analizine ait Tablo 29 incelendiğinde kayıp veri oranı 0; 0,05 ve 0,10 olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 18604,24$; $F_{(2,8019)} = 17985,43$; $F_{(2,8019)} = 17684,87$; $p < 0,01$). Ek olarak, Tablo 29’da POBT, YSA ve DINO modeline ait NSO’lar arasındaki en büyük farkın kayıp veri oranı sıfır iken, en az farkın ise kayıp veri oranının 0,10 iken olduğu görülmektedir.

Tablo 29

Kayıp Veri Oranı Faktörünün Her Bir Koşulunda POBT’nin, YSA’nın ve DINO Modelinin NSO’larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	F	p	Anlamlı Fark
Kayıp veri oranı 0						DINO-POBT
Modeller	204,324	2	102,162	18604,235	0,000	DINO-YSA
Hata	133,390	24291	0,005			POBT-YSA
Kayıp veri oranı 0,05						DINO-POBT
Modeller	197,528	2	98,764	17985,427	0,000	DINO-YSA
Hata	133,390	24291	0,005			POBT-YSA
Kayıp veri oranı 0,10						DINO-POBT
Modeller	194,227	2	97,114	17684,871	0,000	DINO-YSA
Hata	133,390	24291	0,005			POBT-YSA

Kayıp veri oranı koşulları ile modellerin etkileşimlerine ait grafik Şekil 14’te verilmiştir. Şekil 14 incelendiğinde tüm kayıp veri oranı koşullarında DINO modelinin POBT ve YSA’ya göre en yüksek NSO’lara sahipken, YSA’nın en düşük NSO’lara sahip olduğu görülmektedir.



Şekil 14. Kayıp verinin 0; 0,05 ve 0,10 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin NSO'ları

ÖSO'ların POBT, YSA ve DINO model ile kayıp veri oranı faktörü ve bu faktörün koşulları (0; 0,05 ve 0,10) arasında karşılaştırma sonuçlarına ait ana ve etkileşim etkileri Tablo 30'da, basit etkileri ise Tablo 31'de sunulmuştur. Tablo 30 incelendiğinde ÖSO'ların kayıp veri oranına ve modele göre istatistiksel olarak anlamlı farklılık gösterdiği görülmektedir (sırasıyla $F_{(2,24291)} = 56,22$; $F_{(2,24291)} = 32160,65$; $p < 0,01$). Etki büyüklükleri incelendiğinde kayıp veri oranının küçük etkiye ($\eta^2 = 0,005$), modellerin ise orta düzey etkiye ($\eta^2 = 0,73$) sahip oldukları görülmektedir.

Tablo 30

Kayıp Veri Oranı Faktöründe POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına ait 3x3 Faktöriyel ANOVA Sonuçları

Kaynak	Kareler Toplamı	sd	Ortalama Kare	F	p	η_p^2
Kayıp veri oranı	3,230	2	1,615	56,219	0,000	0,005
Model	1847,983	2	923,991	32160,645	0,000	0,726
Kayıp veri oranı * Model	1,035	4	0,259	9,004	0,000	0,001
Hata	697,893	24291	0,029			

Tablo 30 incelendiğinde kayıp veri oranı ile modeller arasındaki etkileşimin istatistiksel olarak anlamlı olduğu ($F_{(4,8019)} = 9,00$; $p < 0,01$) ve bu etkileşimin küçük etkiye ($\eta_p^2 = 0,001$)

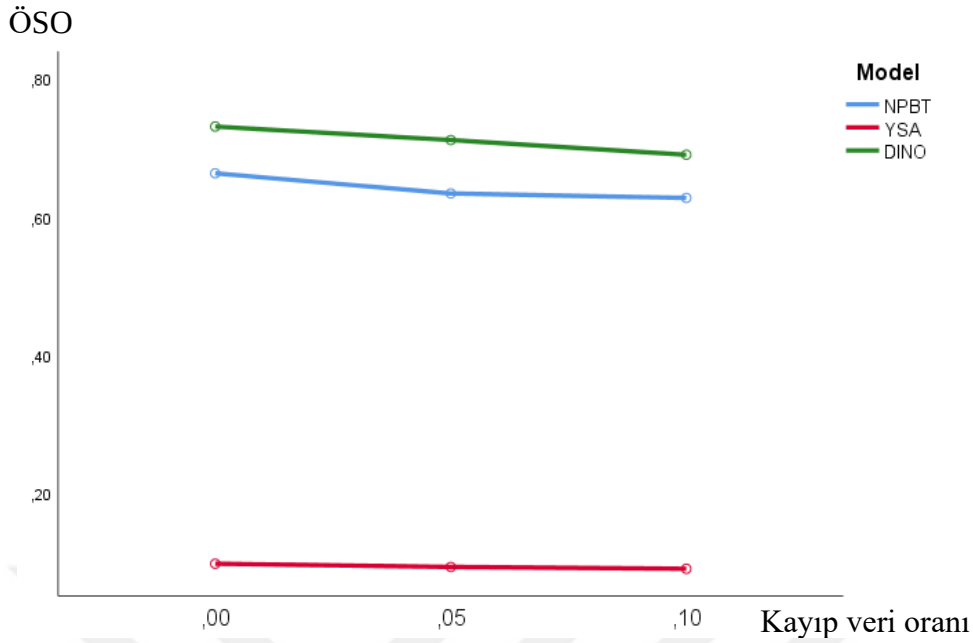
sahip olduğu görülmektedir. Kayıp veri oranı ile modeller arasındaki etkileşimin anlamlı olması nedeniyle kayıp veri oranı faktörünün her bir koşulunda modeller arasında istatistiksel olarak anlamlı fark olup olmadığı basit etkiler analiziyle incelenmiştir. Basit etkiler analizine ait Tablo 31 incelendiğinde kayıp veri oranı 0; 0,05 ve 0,10 olduğunda modeller arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir (sırasıyla $F_{(2,8019)} = 11338,64$; $F_{(2,8019)} = 10646,10$; $F_{(2,8019)} = 10193,91$; $p < 0,01$).

Tablo 31

Kayıp Veri Oranı Faktörünün Her Bir Koşulunda POBT'nin, YSA'nın ve DINO Modelinin ÖSO'larına Ait Basit Etki ANOVA Sonuçları

Koşullar	Kareler Toplamı	Serbestlik Derecesi	Ortalama Kare	F	p	Anlamlı Fark
Kayıp Veri Oranı 0						DINO-POBT
Modeller	651,530	2	325,765	11338,644	0,000	DINO-YSA
Hata	697,893	24291	0,029			POBT-YSA
Kayıp Veri Oranı 0,05						DINO-POBT
Modeller	611,735	2	305,868	10646,096	0,000	DINO-YSA
Hata	697,893	24291	0,029			POBT-YSA
Kayıp Veri Oranı 0,10						DINO-POBT
Modeller	585,752	2	292,876	10193,914	0,000	DINO-YSA
Hata	697,893	24291	0,029			POBT-YSA

Ek olarak, Tablo 31'de POBT, YSA ve DINO modeline ait ÖSO'lar arasındaki en büyük farkın kayıp veri oranı sıfır iken, en az farkın ise kayıp veri oranının 0,10 iken olduğu görülmektedir. Kayıp veri oranı koşulları ile modellerin etkileşimlerine ait grafik Şekil 15'te verilmiştir. Şekil 15 incelendiğinde tüm kayıp veri oranı koşullarında DINO modelinin POBT ve YSA'ya göre en yüksek ÖSO'lara sahipken, YSA'nın en düşük ÖSO'lara sahip olduğu görülmektedir



Şekil 15. Kayıp veri oranının 0; 0,05 ve 0,10 olduğu koşullarda POBT'nin, YSA'nın ve DINO modelinin ÖSO'ları

Sonuç olarak, tüm koşullarda DINO modeline ait NSO'ların POBT ve YSA'ya ait NSO'lardan, POBT'ye ait NSO'ların ise YSA'ya ait NSO'lardan daha yüksek olduğu görülmektedir. Benzer şekilde DINO modeline ait ÖSO'lar POBT ve YSA'ya ait ÖSO'lardan, POBT'ye ait ÖSO'lar ise YSA'ya ait ÖSO'lardan daha yüksektir.

4.3. Gerçek Veri Setinde (PISA 2015 İPÇ) POBT'nin, YSA'nın ve DINO Modeli Sınıflandırma Oranlarının Benzerliğinin Nitelik Sayılarına (3, 7 ve 11) ve Örneklem Büyüklüklerine (30, 100 ve 500) Göre Durumları

Bu araştırma problemi nitelik ve örneklem koşullarına göre alt problemlere ayrılarak cevaplanmıştır. Her bir alt problemde teknik raporlar doğrultusunda oluşturulan Q matrisi ile geçerlemesi yapılmış Q matrislerinden elde edilen bulgular ve NSOB ve ÖSOB'lar aynı tabloda verilmiştir. Alt problemlerin cevaplanmasında öncelikle teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen POBT, YSA ve DINO modeline ait NSOB ve ÖSOB'lar yorumlanmış ardından POBT, YSA ve DINO modeline ait NSOB ve ÖSOB'lar birbirleriyle karşılaştırılmıştır. Daha sonra benzer incelemeler geçerlemesi yapılmış Q matrislerinden elde edilen bulgular üzerinde yapılmış ve bu bulgular yorumlanmıştır. Son olarak, teknik raporlar doğrultusunda oluşturulan Q matrisi ile geçerlemesi yapılmış Q matrislerinden elde edilen bulgular karşılaştırılmıştır.

4.3.1. Nitelik Sayılarına (3, 7 ve 11) Göre POBT, YSA ve DINO Modeline ait Sınıflandırma Oranlarının Benzerliği

Farklı nitelik sayılarına (3, 7 ve 11) göre POBT, YSA ve DINO modeline ait NSOB ve ÖSOB'lar Tablo 32'de verilmiştir. Tablo 32 incelendiğinde her bir örneklem büyüklüğü için nitelik sayısı arttıkça genel olarak teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile YSA'ya, POBT ile DINO modeline ve YSA ile DINO modeline ait NSOB'ların, önce azaldığı sonra arttığı söylenebilir. NSOB'ların nitelik sayısının yedi olduğunda azalmasının sebebinin karmaşık Q matris yapısı olduğu düşünülmektedir. Mevcut araştırma kapsamında oluşturulan Q matrislerinden üç ve 11 nitelikli Q matrisleri basit yapıdadır, yani her bir maddenin doğru cevaplanması için tek bir niteliğe sahip olmak gerekmektedir. Yedi nitelikli Q matrisi ise karmaşık yapıdadır ve her bir madde birden fazla nitelik (iki) ile ölçülmüştür. Böylece yedi nitelikli Q matrisinde nitelikler, diğer Q matris yapısındaki niteliklere göre daha fazla ölçülmüş, dolayısıyla daha kompleks bir Q matris yapısı oluşmuştur.

Tablo 32 incelendiğinde teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile DINO modeli NSOB'larının güvenilirlik derecesi kabul edilebilir düzeyin üstünde (0,65 üzeri) olduğu, POBT ile YSA'ya ve YSA ile DINO modeline ait NSOB'ların 0,65 üzeri olmalarının nitelik sayısına göre değiştiği görülmektedir. POBT ile YSA nitelik düzeyi sınıflandırma oranlarının benzerliği örneklem büyüklüğü 30 ve 500 iken nitelik sayısı 11 olduğunda güvenilirlik derecesi kabul edilebilir düzeyde (0,50 – 0,65), diğer tüm durumlarda ise güvenilirlik derecesi kabul edilebilir düzeyin üstünde olmaktadır. YSA ile DINO modelinin nitelik düzeyi sınıflandırma oranlarının benzerliği örneklem büyüklüğü 30 ve 500 iken nitelik sayısı 11 olduğunda güvenilirlik derecesi kabul edilebilir düzeyde, diğer tüm koşullarda ise kabul edilebilir düzeyin üstündedir.

Tablo 32

Gerçek Veride POBT, YSA ve DINO Modeline ait NSOB ve ÖSOB'ların Nitelik Sayısına Göre Değişimi

N	Ni	İlk Q matrisi						Geçerlemesi yapılmış Q matrisi					
		POBT-YSA		POBT-DINO		YSA-DINO		POBT-YSA		POBT-DINO		YSA-DINO	
		NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB
30	3	0,730	0,471	0,941	0,875	0,705	0,472	0,944	0,867	0,822	0,533	0,789	0,467
	7	0,714	0,201	0,808	0,302	0,721	0,207	0,786	0,233	0,771	0,400	0,805	0,267
	11	0,570	0,002	0,827	0,070	0,573	0,003	0,800	0,033	0,854	0,267	0,624	0,00
100	3	0,773	0,532	0,843	0,684	0,722	0,481	0,783	0,470	0,867	0,710	0,717	0,380
	7	0,691	0,127	0,776	0,316	0,658	0,062	0,741	0,330	0,771	0,430	0,819	0,410
	11	0,706	0,038	0,805	0,087	0,685	0,014	0,765	0,090	0,805	0,080	0,722	0,080
500	3	0,781	0,515	0,879	0,664	0,696	0,426	0,717	0,436	0,861	0,668	0,761	0,496
	7	0,682	0,141	0,777	0,326	0,679	0,157	0,790	0,362	0,860	0,594	0,824	0,418
	11	0,654	0,026	0,833	0,123	0,624	0,018	0,637	0,030	0,852	0,334	0,612	0,018

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğünü temsil etmektedir.

Teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen NSOB'lar incelendiğinde POBT ile DINO model arasındaki NSOB'ların POBT ile YSA ve YSA ile DINO arasındaki NSOB'lardan daha yüksek olduğu görülmektedir. Başka bir ifadeyle POBT ile DINO modeli arasındaki nitelik düzeyi sınıflandırma oranının benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten fazladır. Ayrıca Tablo 32'de POBT ile YSA'ya ve YSA ile DINO modeline ait NSOB'ların da birbirlerine benzediği görülmektedir.

Tablo 32 incelendiğinde her bir örneklem büyüklüğü için nitelik sayısı arttıkça teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile YSA'ya, POBT ile DINO modeline ve YSA ile DINO modeline ait ÖSOB'ların genel olarak azaldığı söylenebilir. POBT ile DINO model ÖSOB'larının örneklem büyüklüğü 30, 100 ve 500 ve nitelik sayısı sadece üç iken güvenilirlik derecesi kabul edilebilir düzeyin üstünde olduğu söylenebilir. POBT ile YSA'ya ve YSA ile DINO modeline ait ÖSOB'lar ise hiçbir koşulda güvenilirlik derecesi kabul edilebilir düzeyin üstünde değildir. ÖSOB'lar incelendiğinde POBT ile DINO model arasındaki ÖSOB'ların POBT ile YSA ve YSA ile DINO model arasındaki ÖSOB'lardan daha yüksek olduğu görülmektedir. Başka bir ifadeyle POBT ile DINO modeli arasındaki örüntü düzeyi sınıflandırma oranının benzerliği POBT ile YSA ve YSA ile DINO model arasındaki

benzerlikten fazladır. Ayrıca POBT ile YSA'ya ve YSA ile DINO modeline ait ÖSOB'lar da birbirlerine benzemektedir.

Tablo 32 incelendiğinde her bir örneklem büyüklüğü için nitelik sayısı arttıkça genel olarak geçerlemesi yapılmış Q matrisine dayalı analizler sonucu elde edilen POBT ile YSA'ya ve POBT ile DINO modeline ait NSOB'ların, önce azaldığı sonra arttığı söylenebilir. YSA ile DINO modeline ait NSOB'ların ise nitelik sayısı arttıkça önce arttığı sonra azaldığı görülmektedir. PVAF yöntemiyle yapılan geçerleme sürecinden sonra tüm Q matrisleri karmaşık yapıdadır. Bu nedenle geçerlemesi yapılmış Q matrisine dayalı analizler sonucu elde edilen bulgular, teknik raporlar doğrultusunda oluşturulan Q matrislerin analizlerinden elde edilen bulgularla benzerlik göstermenin yanı sıra bazı farklılıklar içermektedir.

Tablo 32 incelendiğinde geçerlemesi yapılmış Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile DINO modeli NSOB'larının tümünün güvenilirlik derecesi kabul edilebilir düzeyin üstünde (0,65 üzeri) olduğu, POBT ile YSA'ya ve YSA ile DINO modeline ait NSOB'ların ise 0,65 üzeri olmalarının nitelik sayısına göre değiştiği görülmektedir. POBT ile YSA nitelik düzeyi sınıflandırma oranlarının benzerliği örneklem büyüklüğü 500 iken nitelik sayısı 11 olduğunda güvenilirlik derecesi kabul edilebilir düzeyde (0,50 – 0,65), diğer tüm durumlarda ise güvenilirlik derecesi kabul edilebilir düzeyin üstünde olmaktadır. YSA ile DINO modelinin nitelik düzeyi sınıflandırma oranlarının benzerliği örneklem büyüklüğü 30 ve 500 iken nitelik sayısı 11 olduğunda güvenilirlik derecesi kabul edilebilir düzeyde, diğer tüm koşullarda ise kabul edilebilir düzeyin üstündedir.

Geçerlemesi yapılmış Q matrisine dayalı yapılan analizler sonucu elde edilen NSOB'lar incelendiğinde POBT ile DINO model arasındaki NSOB'ların POBT ile YSA ve YSA ile DINO arasındaki NSOB'lardan genel olarak daha yüksek olduğu görülmektedir. Başka bir ifadeyle POBT ile DINO modeli arasındaki nitelik düzeyi sınıflandırma oranının benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten fazladır.

Tablo 32 incelendiğinde her bir örneklem büyüklüğü için nitelik sayısı arttıkça geçerlemesi yapılmış Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile YSA'ya, POBT ile DINO modeline ve YSA ile DINO modeline ait ÖSOB'ların genel olarak azaldığı söylenebilir. POBT ile DINO model ÖSOB'larının örneklem büyüklüğü 100 ve 500 ve nitelik sayısı sadece üç iken; POBT ile YSA'ya ait ÖSOB'ların ise örneklem büyüklüğü 30 ve nitelik sayısı sadece üç iken güvenilirlik derecesi kabul edilebilir düzeyin üstünde olduğu

söylenir. YSA ile DINO modeline ait ÖSOB'lar ise hiçbir koşulda güvenilirlik derecesi kabul edilebilir düzeyin üstünde değildir. ÖSOB'lar incelendiğinde POBT ile DINO model arasındaki ÖSOB'ların POBT ile YSA ve YSA ile DINO model arasındaki ÖSOB'lardan genel olarak daha yüksek olduğu görülmektedir. Başka bir ifadeyle POBT ile DINO modeli arasındaki örüntü düzeyi sınıflandırma oranının benzerliği POBT ile YSA ve YSA ile DINO model arasındaki benzerlikten fazladır.

Sonuç olarak, geçerlemesi yapılan Q matris sonucu elde edilen NSOB ve ÖSOB'lar genel olarak daha yüksektir. Teknik raporlardan elde edilen Q matrisleri ve geçerlemesi yapılmış Q matrislerinin analizleri sonucu bazı benzer bulgular elde edilmiştir. Bu bulgular:

- POBT ile DINO modeline ait NSOB ve ÖSOB'lar arasındaki benzerlik her iki Q matris bulgusunda yüksektir.
- Nitelik sayısı arttıkça genel olarak POBT ile YSA'ya ve POBT ile DINO modeline ait NSOB'ların, önce azaldığı sonra arttığı söylenebilir.
- Nitelik sayısı arttıkça genel olarak POBT ile YSA'ya, POBT ile DINO modeline ve YSA ile DINO modeline ait ÖSOB'ların genel olarak azaldığı söylenebilir.

4.3.2. Örneklem Büyüklüklerine (30, 100 ve 500) Göre POBT, YSA ve DINO Modeline ait Sınıflandırma Oranlarının Benzerliği

Farklı örneklem büyüklüklerine (30, 100 ve 500) göre POBT, YSA ve DINO modeline ait NSOB ve ÖSOB'lar Tablo 33'te verilmiştir. Tablo 33'te görüldüğü üzere örneklem büyüklüğü arttıkça teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile YSA'ya, POBT ile DINO modeline ve YSA ile DINO modeline ait NSOB ve ÖSOB değerlerinin, artış veya azalışlarına ait herhangi bir sistematikliğe rastlanmamıştır. POBT ve YSA'ya ait NSO ve ÖSO'ların teorikte ve uygulamada örneklem büyüklüğünden etkilenmedikleri göz önüne alındığında böyle bir bulgu beklenmektedir.

Tablo 33 incelendiğinde teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile YSA'ya ve YSA ile DINO modeline ait NSOB'ların güvenilirlik derecesi kabul edilebilir düzeyin üstünde (0,65 üzeri) olması örneklem büyüklüğüne bağlı iken ÖSOB'ların hiçbir koşulda güvenilirlik derecesi kabul edilebilir düzeyin üstünde olmadığı görülmektedir. POBT ile YSA'ya ait NSOB'lar nitelik sayısı 11 iken örneklem büyüklüğü 30 olduğunda güvenilirlik derecesi kabul edilebilir düzeyde (0,50 – 0,65) olmakta, diğer durumlarda ise güvenilirlik derecesi kabul edilebilir

düzeyin üstünde olmaktadır. YSA ile DINO modeline ait NSOB'lar ise nitelik sayısı 11 ve örneklem büyüklüğü 30 ve 500 iken güvenilirlik derecesi kabul edilebilir düzeyde, diğer tüm koşullarda güvenilirlik derecesi kabul edilebilir düzeyin üstündedir. POBT ile DINO modeline ait NSOB'ların ise tüm nitelik ve örneklem koşullarında güvenilirlik derecesi kabul edilebilir düzeyin üstündeyken ÖSOB'ların nitelik sayısı sadece üç iken tüm örneklem büyüklüklerinde güvenilirlik derecesi kabul edilebilir düzeyin üstünde olduğu görülmektedir.

Tablo 33

Gerçek Veride POBT, YSA ve DINO modeli NSOB ve ÖSOB'ların Örneklem Büyüklüğüne Göre Değişimi

N	Ni	İlk Q matrisi						Geçerlemesi yapılmış Q matrisi					
		POBT-YSA		POBT-DINO		YSA-DINO		POBT-YSA		POBT-DINO		YSA-DINO	
		NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB	NSOB	ÖSOB
3	30	0,730	0,471	0,941	0,875	0,705	0,472	0,944	0,867	0,822	0,533	0,789	0,467
	100	0,773	0,532	0,843	0,684	0,722	0,481	0,783	0,470	0,867	0,710	0,717	0,380
	500	0,781	0,515	0,879	0,664	0,696	0,426	0,717	0,436	0,861	0,668	0,761	0,496
7	30	0,714	0,201	0,808	0,302	0,721	0,207	0,786	0,233	0,771	0,400	0,805	0,267
	100	0,691	0,127	0,776	0,316	0,658	0,062	0,741	0,330	0,771	0,430	0,819	0,410
	500	0,682	0,141	0,777	0,326	0,679	0,157	0,790	0,362	0,860	0,594	0,824	0,418
11	30	0,570	0,002	0,827	0,070	0,573	0,003	0,800	0,033	0,854	0,267	0,624	0,00
	100	0,706	0,038	0,805	0,087	0,685	0,014	0,765	0,090	0,805	0,080	0,722	0,080
	500	0,654	0,026	0,833	0,123	0,624	0,018	0,637	0,030	0,852	0,334	0,612	0,018

Not: Ni: Nitelik sayısı, N: Örneklem büyüklüğünü temsil etmektedir.

Teknik raporlar doğrultusunda oluşturulan Q matrisine dayalı yapılan analizler sonucu elde edilen POBT ile DINO model arasındaki NSOB ve ÖSOB'ların benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten daha fazla olduğu görülmektedir. Başka bir ifadeyle POBT ile DINO modeli arasındaki nitelik ve örüntü düzeyi sınıflandırma oranlarının benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten fazladır. POBT ile YSA'ya ve YSA ile DINO modeline ait NSOB ve ÖSOB'lar da birbirlerine benzemektedir.

Tablo 33'te görüldüğü üzere örneklem büyüklüğü arttıkça geçerlemesi yapılmış Q matrisine dayalı analizler sonucu elde edilen POBT ile YSA'ya, POBT ile DINO modeline ve YSA ile DINO modeline ait NSOB ve ÖSOB değerlerinin, artış veya azalışlarına ait herhangi bir sistematikliğe rastlanmamıştır. POBT ve YSA'ya ait NSO ve ÖSO'ların teorikte ve uygulamada örneklem büyüklüğünden etkilenmedikleri göz önüne alındığında böyle bir bulgu beklenmektedir.

Tablo 33 incelendiğinde geçerlemesi yapılmış Q matrisine dayalı analizler sonucu elde edilen POBT ile YSA'ya ve YSA ile DINO modeline ait NSOB'ların güvenilirlik derecesi kabul edilebilir düzeyin üstünde (0,65 üzeri) olması örneklem büyüklüğüne bağlı iken ÖSOB'ların (üç nitelik ve 30 örneklem koşulu dışında) hiçbir koşulda güvenilirlik derecesi kabul edilebilir düzeyin üstünde olmadığı görülmektedir. POBT ile YSA nitelik düzeyi sınıflandırma oranlarının benzerliği örneklem büyüklüğü 500 iken nitelik sayısı 11 olduğunda güvenilirlik derecesi kabul edilebilir düzeyde (0,50 – 0,65), diğer tüm durumlarda ise güvenilirlik derecesi kabul edilebilir düzeyin üstündedir. YSA ile DINO modeline ait NSOB'lar ise nitelik sayısı 11 ve örneklem büyüklüğü 30 ve 500 iken güvenilirlik derecesi kabul edilebilir düzeyde, diğer tüm koşullarda güvenilirlik derecesi kabul edilebilir düzeyin üstündedir. POBT ile DINO modeline ait NSOB'ların ise tüm nitelik ve örneklem koşullarında güvenilirlik derecesi kabul edilebilir düzeyin üstündeyken ÖSOB'ların nitelik sayısı sadece üç iken 100 ve 500 örneklem büyüklüklerinde güvenilirlik derecesi kabul edilebilir düzeyin üstünde olduğu görülmektedir.

Geçerlemesi yapılmış Q matrisine dayalı analizler sonucu elde edilen POBT ile DINO model arasındaki NSOB ve ÖSOB'ların benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten daha fazla olduğu görülmektedir. Başka bir ifadeyle POBT ile DINO modeli arasındaki nitelik ve örüntü düzeyi sınıflandırma oranlarının benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten fazladır. POBT ile YSA'ya ve YSA ile DINO modeline ait NSOB ve ÖSOB'lar da birbirlerine benzemektedir.

Sonuç olarak, geçerlemesi yapılan Q matris sonucu elde edilen NSOB ve ÖSOB'lar genel olarak daha yüksektir. Teknik rapor doğrultusunda oluşturulan Q matris ve geçerlemesi yapılmış Q matrisinin analizi sonucu aşağıdaki bulguların ortak olduğu belirlenmiştir. Bu bulgular:

- POBT ile YSA'ya, POBT ile DINO modeline ve YSA ile DINO modeline ait NSOB ve ÖSOB değerlerinin, artış veya azalışlarına ait herhangi bir sistematikliğe rastlanmamıştır.
- POBT ile DINO model arasındaki NSOB ve ÖSOB'ların benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten daha fazla olduğu görülmektedir.

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

Bu bölümde araştırma sonucunda elde edilen bulgulara ilişkin sonuç, tartışma ve önerilere yer verilmiştir.

5.1. Sonuç ve Tartışma

Tanılayıcı ve biçimlendirici değerlendirme ve 21. yüzyıl becerilerinin ölçülmesi ve değerlendirilmesi son yıllarda önem kazanmıştır. Buradan hareketle sınıf veya kurs düzeyi öğretme ve öğrenme süreçlerinin ve karmaşık yapıya sahip becerilerin tanılayıcı ve biçimlendirici değerlendirmeleri gündeme gelmektedir. Bu bağlamda BTM'ler yaygın bir şekilde kullanılmaktadır. Fakat BTM'lerin küçük örneklemin, az madde sayısının veya karmaşık yapıların (fazla nitelik sayısı) bulunduğu bazı durumlarda öğrenciler hakkında yanlış sonuçlar verebildikleri görülmektedir (Shu vd., 2013). Bu nedenle parametrik olmayan model ve tekniklerin, 21. yüzyıl becerilerinin veya sınıf düzeyindeki tanılayıcı ve biçimlendirici değerlendirmelerde güvenilir ve kullanışlı birer yöntem olabileceği düşünülmektedir. Buradan hareketle mevcut araştırmada, POBT ve YSA'ya ait nitelik ve örüntü düzeyi sınıflandırma oranlarının çeşitli koşullarda DINO temelli simülasyon veri setlerinde öncelikle birbirleriyle, sonra DINO model ile karşılaştırılmaları, ardından POBT, YSA ve DINO modelin sınıflandırma oranlarının benzerliklerinin PISA 2015 İPÇ veri seti üzerinde incelenmesi amaçlanmıştır.

Mevcut araştırmada simülasyon ve gerçek veri setlerinden elde edilen sonuçların karşılaştırılabilir olması için simülasyon veri setleri gerçek veri setlerinin yapısına benzer şekilde üretilmiştir. Bu düşünceden hareketle öncelikle PISA 2015 İPÇ yeterliğinin yapısı

incelenmiştir. PISA 2015 İPÇ yeterliğini oluşturan problem çözme süreçleri ve iş birliği becerilerini temsil eden nitelikler arasında herhangi bir ardışıklık veya önkoşul ilişkisi bulunmadığı görülmüştür. Yani, problem çözme süreç becerilerinden bir veya birkaçına sahip bir öğrenci diğer becerilere sahip olmasa da bir problemi çözebilmektedir (Yavuz, 2014) ve bu PISA 2015 İPÇ yapısının telafisel olduğunu göstermektedir. PISA 2015 İPÇ yeterliğiyle ilgili teknik raporlarda tüm maddeler paylaşılmadığından doğru cevaplanması için birden fazla niteliğe sahip olunması gereken maddelerde, hangi niteliğin maddenin doğru cevaplanmasına daha çok katkı sağladığı belirlenmemektedir. Bu nedenle İPÇ yeterliğinin analizinde ve simülasyon veri setlerinin üretilmesinde temel olarak telafisel modellerden maddenin doğru cevaplanmasında her bir niteliğin eşit katkı yaptığını varsayan DINO modeli seçilmiştir. Mevcut araştırmada simülasyon verisinin DINO model temel alınarak üretilmesi ve veri setlerinin analizi için POBT, YSA ve DINO modelin bu veri yapısına göre şekillenmesi bu çalışmayı YSA ile yapılan çalışmalardan (bkz. Cui vd., 2016; Guo vd., 2017; McCoy & Willse, 2014; Paulsen, 2019; Shu vd., 2013) ayırmaktadır. Mevcut araştırmadan elde edilen bulgular telafisel olmayan modellerle yapılan çalışmalarla da karşılaştırılarak tartışılmıştır.

Mevcut araştırmada gerçek veri setlerinde madde parametreleri (s ve g) 0,33 ile 0,35 arasında değiştiği için simülasyon veri setlerinin maddelerine ait s ve g parametreleri 0 ile 0,4 aralığında belirlenmiştir. S ve g parametreleri 0,5'i geçmediği sürece ideal cevap örüntüleri en olası yeterlik sınıfının belirlenmesinde yardımcı olmaktadır (Chiu & Douglas, 2013). Mevcut araştırmada s ve g parametre değerlerinin 0,5'i geçmese de bu değere yakın olmalarının NSO ve ÖSO'ların düşük hesaplanmalarına neden olduğu düşünülmektedir. Çalışmalarında s ve g parametrelerinin model veya tekniklerin sınıflandırma performanslarına etkilerini araştıran Akbay (2016) ve Chiu ve Douglas'ın (2013) elde ettiği sonuçlar da bu düşüncüyü destekler niteliktedir. DINO model temel alınarak üretilen simülasyon veri setlerinde Akbay (2016) iki farklı ayırt edicilik düzeyinin (s ve g parametrelerinin her ikisi için 0 ile 0,15 ve 0 ile 0,25 arası), Chiu ve Douglas (2013) ise üç farklı ayırt edicilik düzeyinin (s ve g parametrelerinin her ikisi için 0 ile 0,1; 0 ile 0,3 ve 0 ile 0,5 arası) sınıflandırmaya olan etkilerini incelemişlerdir. Araştırmacılar s ve g parametre değerleri arttıkça parametrik ve parametrik olmayan model ve tekniklerin sınıflandırma oranlarının azaldığını belirlemişlerdir. DINA model temelli simülasyon veri setleri üzerinde çalışan Chiu vd. (2017), Cui vd. (2016), Paulsen (2019), Wang (2015) ve Wang ve Douglas (2015) da benzer sonuçlara ulaşmışlardır. Paulsen (2019) çalışmasında

ayırt ediciliği yüksek olan maddelerin bulunduğu koşullarda POBT'nin, YSA'nın ve DINA modelinin tüm koşullarda 0,80 üzeri NSO'lara, düşük ayırt edici maddelerin bulunduğu koşullarda ise POBT ve DINO modelin iki koşul dışında YSA'nın ise tüm koşullarda güvenilirlik derecesi kabul edilebilir düzeyin üstünde olmayan NSO'lara sahip olduklarını tespit etmiştir. Ayrıca Paulsen (2019) yüksek ayırt edici maddelerin bulunduğu durumlarda POBT'nin ve DINA modelin tüm koşullarda, YSA'nın ise iki koşulda 0,80 üzeri ÖSO'lara sahip olduğunu, düşük ayırt edici maddelerin bulunduğu tüm koşullarda ise POBT, YSA ve DINA modelin güvenilirlik derecesi kabul edilebilir düzeyin üstünde olmayan ÖSO'lara sahip olduklarını belirlemiştir. Orta düzeyde ayırt ediciliğe sahip maddelerin kullanıldığı mevcut araştırmada POBT ve DINO modelin 0,77 üzeri NSO'lara sahip olmaları, YSA'nın ise 0,43 ile 0,80 arasında NSO'lara sahip olması Paulsen'in (2019) bulgularını destekler niteliktedir. Ayrıca Paulsen'in (2019) çalışmasıyla benzer şekilde mevcut araştırmada bazı koşullarda POBT ve DINO modelinin güvenilirlik derecesi kabul edilebilir düzeyin üstünde ÖSO'lara sahipken YSA'nın hiçbir koşulda güvenilirlik derecesi kabul edilebilir düzeyin üstünde ÖSO'lara sahip olmadığı görülmektedir. Ayrıca mevcut araştırmada elde edilen NSO ve ÖSO'ların Chiu vd. (2017) çalışmalarında elde edilen NSO ve ÖSO'lar ile benzerlik gösterdiği de görülmektedir.

Mevcut araştırmada gerçek veri setlerinin analizlerinde ilk önce PISA 2015 İPÇ yapısının ifade edilişine bağlı olarak basit ve karmaşık Q matrisleri kullanılmıştır. Ardından Q matris geçerlemesi yapılarak analizler tekrar edilmiştir. Geçerleme işlemi sonucunda karmaşık yapıda Q matrisleri elde edilmiştir. Bu nedenle simülasyon verisinin üretilmesi ve analizinde karmaşık Q matrisi kullanılmasına karar verilmiştir. Bir Q matrisi, bir maddenin doğru cevaplanması için en az iki niteliğe sahip olunmasını gerektiriyorsa karmaşık yapıdadır. Mevcut araştırma kapsamında karmaşık Q matrislerinin kullanılmasının NSO ve ÖSO'ların düşük hesaplanmalarının başka bir nedeni olduğu düşünülmektedir. Nitekim çalışmalarında basit ve karmaşık Q matrislerinin model ve yöntemlerin sınıflandırma performanslarına etkilerini inceleyen Akbay (2016), Paulsen (2019) ve Shu vd. (2013) karmaşık yapıda Q matrislerinin kullanıldığı koşullarda yöntemlerin basit yapıda Q matrislerinin kullanıldığı koşullara göre daha düşük NSO ve ÖSO'lara sahip olduklarını belirlemişlerdir.

Mevcut araştırmada kayıp veri türünün tamamıyla rastgele olarak belirlenmesinin kayıp verinin bulunduğu veri setlerinde NSO ve ÖSO'ların düşük hesaplanmalarının başka bir nedeni olduğu düşünülmektedir. Bilişsel tanılamada kayıp veri türü (bkz. Dai, 2017;

Sünbül, 2018) ve kayıp veri ile baş etme yöntemleri (bkz. Ayers vd., 2009; Dai, 2017; Sünbül, 2018) sınıflandırma doğruluğunu etkilemektedir. Örneğin rastgele ve tamamıyla rastgele kayıp veri türünün sınıflandırma doğruluğuna etkisini inceleyen Sünbül (2018) tamamıyla rastgele kayıp veri türüne sahip veri setlerinde DINA modelin sınıflandırma oranlarının daha düşük olduğunu tespit etmiştir. Kayıp veri türüne ek olarak mevcut araştırmada kullanılan veri atama yönteminin de NSO ve ÖSO'ların düşük hesaplanmalarının bir nedeni olduğu düşünülmektedir. Dai (2017), DINA modelinin iki yönlü atamanın yapıldığı veri setlerindeki sınıflandırma oranlarının EM atamanın yapıldığı veri setlerindeki sınıflandırma oranlarından daha düşük olduğunu tespit etmiştir. Dai'nin çalışmasına benzer şekilde mevcut araştırmada iki yönlü veri atama yönteminin kullanılmasının NSO ve ÖSO'ların düşük hesaplanmalarına neden olduğu düşünülmektedir.

Mevcut araştırmada POBT'ye ait NSO ve ÖSO'ların düşük hesaplanmalarının nedenleri arasında madde ayırt ediciliği, Q matris yapısı, kayıp veri türü ve kayıp veri ile baş etme yolu haricinde seçilen uzaklık ölçütleri ve nitelikler arası ilişkinin büyüklüğü de sayılabilir. Chiu ve Douglas (2013) çalışmalarında ağırlıklandırılmış ve cezalandırılmış Hamming uzaklıklarını, McCoy ve Willse (2014) ise Hamming ve ağırlıklandırılmış Hamming uzaklıklarını karşılaştırmışlardır. Her iki çalışmada da ağırlıklandırılmış Hamming uzaklığına ait sınıflandırma oranları diğerlerinden daha düşüktür. Mevcut araştırmada ağırlıklandırılmış Hamming uzaklığı tercih edildiği için POBT'ye ait NSO ve ÖSO'ların olduklarından düşük hesaplandıkları düşünülmektedir. Mevcut araştırmada nitelikler arası ilişki orta düzey ($p = 0,5$) olarak belirlenmiştir. Lim ve Drasgow (2017) ve Paulsen (2019) çalışmalarında nitelikler arası ilişkinin daha az olduğu durumlarda POBT'ye ait sınıflandırma oranlarının daha yüksek olduğunu belirlemişlerdir. Buradan hareketle mevcut araştırmada hesaplanan NSO ve ÖSO'ların nitelikler arası ilişkinin daha düşük olduğunda hesaplanan NSO ve ÖSO'lardan daha düşük olduğu düşünülmektedir.

Mevcut araştırmada YSA'ya ait NSO ve ÖSO'ların düşük hesaplanmalarının nedeni olarak madde ayırt ediciliği, Q matris yapısı, kayıp veri türü ve kayıp veri ile baş etme yoluna ek olarak ağ yapısı (katman ve katman/katmanlardaki nöron sayısı, giriş katmanı nöron sayılarının başlangıç değerleri) ve ağın eğitim süreci (ağın öğrenme hızı, momentumu, iterasyon sayısı, aktivasyon fonksiyonu vb.) de sayılabilir (Arı & Berberler, 2017; Shu vd., 2013). Alanyazında YSA'ların kaç katmanlı olacağı ve katmanın/katmanların kaç nörondan oluşacağı deneme yöntemiyle belirlenmektedir.

Ayrıca başlangıç nöron sayılarının belirlenmesi için çeşitli formüller öne sürülmektedir (bkz. Kavzoglu & Mather, 2003). Alanyazında başlangıç nöron sayısının belirlenmesinde net bir yöntemin kabul görmemiş olması mevcut araştırma kapsamında elde edilen YSA'ya ait NSO ve ÖSO'ların düşük olmalarına sebep olabileceği düşünülmektedir.

Mevcut araştırmada NSO ve ÖSO'ların düşük hesaplanmalarına ek olarak NSO'ların ÖSO'lardan yüksek oldukları da dikkati çekmektedir. Simülasyon veri setlerinde GNPC, NPC, G-DINA ve DINA modelinin sınıflandırma oranlarını ve gerçek veri setlerinde bu modellerin sınıflandırma oranlarına ait benzerliği inceleyen Chiu vd. (2017) çalışmalarında benzer şekilde NSO'ların ÖSO'lardan yüksek olduğunu belirlemişlerdir. Sadece simülasyon veri setlerinde çalışan Akbay (2016), Cui vd. (2016), Dai (2017), Lim ve Drasgow (2017) ve Paulsen (2019) de benzer sonuç üzerine vurgu yapmışlardır.

Mevcut araştırmada simülasyon veri setlerine benzer şekilde gerçek veri setlerinde de POBT, YSA ve DINO modeline ait nitelik ve örüntü düzeyi sınıflandırma oranlarının benzerlikleri (sırasıyla NSOB ve ÖSOB) düşük hesaplanmıştır. POBT, YSA ve DINO modeline ait NSO ve ÖSO'larına benzer şekilde PISA 2015 İPÇ madde ayırt ediciliğinin orta düzeyde ve Q matrislerinin karmaşık yapıda olması NSOB ve ÖSOB'ların düşük hesaplanmalarının bir nedeni olduğu düşünülmektedir. Ayrıca PISA 2015 İPÇ yeterliğini oluşturan nitelik sayılarının fazla olması da NSOB ve ÖSOB'ların düşük hesaplanmasının bir nedeni olabileceği düşünülmektedir. Olası nitelik kombinasyonların üstel bir şekilde artması nedeniyle DiBello, Roussos ve Stout (2007) nitelik sayısının üst sınırı için on niteliğin iyi bir pratik kural olabileceğini öne sürmüşlerdir. Mevcut araştırmada 11 niteliğin bulunduğu koşullarda ÖSOB'ların sıfıra yaklaşması bu durumu destekler niteliktedir. Madde ayırt ediciliği, Q matris yapısı ve nitelik sayısına ek olarak NSOB ve ÖSOB'ların düşük hesaplanmalarının bir nedeni POBT'nin, YSA'nın ve DINO modelin öğrencileri sınıflandırmak için kullandıkları yöntemlerin farklı olması da olabilir. Nitekim POBT ile DINO modeline ait NSOB ve ÖSOB'ların POBT ile YSA ve YSA ile DINO modeline ait NSOB ve ÖSOB'lardan yüksek olduğu görülmektedir.

Mevcut araştırmada simülasyon veri setlerine benzer şekilde gerçek veri setlerinde de POBT, YSA ve DINO modeline ait NSOB'ların ÖSOB'lardan daha yüksek olduğu belirlenmiştir. Chiu ve Douglas (2013) ve Chiu ve diğerlerinin (2017) yaptıkları çalışmalarda da POBT ile DINA modeline ait NSOB'ların ÖSOB'lardan yüksek olduğu belirlenmiştir.

Bu bölümde simülasyon ve gerçek veri setlerinde NSO, ÖSO, NSOB ve ÖSOB'ların düşük kestirilmelerinin, NSO'ların ÖSO'lardan ve NSOB'ların ÖSOB'lardan yüksek hesaplanmalarının olası nedenleri tartışıldıktan sonra araştırma problemlerine ilişkin sonuç ve tartışmalara yer verilmiştir. Bu kapsamda her bir araştırma problemine ilişkin sonuç ve tartışma ilgili başlık altında verilmiştir.

5.1.1. Simülasyon Verisinde POBT'nin, YSA'nın ve DINO Modelinin NSO ve ÖSO'larının Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) göre Durumları

Birinci araştırma probleminde POBT'nin, YSA'nın ve DINO modelin nitelik (NSO) ve örüntü (ÖSO) düzeyi sınıflandırma oranları nitelik sayısı, örneklem büyüklüğü, madde sayısı ve kayıp veri oranı koşullarında ayrı ayrı incelenmiştir. Yapılan faktöriyel ANOVA'lar sonucunda tüm koşullar için etkileşim etkisi istatistiksel olarak anlamlı bulunmuştur. Bu bölümde ise genel bulgular tartışılmıştır. Ek olarak, mevcut araştırma kapsamında ele alınan bazı koşullardaki gözlem sayısı, parametre sayısını tahmin etmek için yeterli olmadığından model kurulamamaktadır. Bu nedenle araştırma kapsamında model temelli kayıp veri atama yöntemleri kullanılamamış, onların yerine iki yönlü veri atama yöntemi kullanılmıştır. Bulgular bu sınırlılık çerçevesinde tartışılmıştır.

5.1.1.1. POBT'ye ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) göre Durumları

Bu araştırmada nitelik sayısı (3, 5 ve 7) arttıkça POBT'ye ait NSO ve ÖSO'ların genel olarak azaldıkları belirlenmiştir. Bu sonuç alanyazındaki çalışmalarla da benzerlik göstermektedir. Örneğin DINO model temelli simülasyon veri setlerinde yaptıkları çalışmalarda Chiu ve Douglas (2013) nitelik sayısını üç ve dört olarak, Lim ve Drasgow (2017) üç, beş ve yedi olarak ele almışlardır. DINA model temelli simülasyon veri setlerinde yaptıkları çalışmalarda ise Chiu vd. (2017) nitelik sayısını üç ve beş olarak, Paulsen (2019) ise iki, dört ve altı olarak ele almıştır. Simülasyon verilerinin üretilmesinde temel alınan modellerin ve araştırma kapsamında incelenen nitelik sayılarının değişmesine

rağmen bu çalışmalarda da nitelik sayısı arttıkça POBT'ye ait NSO ve ÖSO'ların azaldıkları belirlenmiştir.

Mevcut araştırmada örneklem büyüklüğü (30, 100 ve 500) arttıkça POBT'ye ait NSO ve ÖSO'ların değişimleri hakkında herhangi bir sistematikliğe rastlanmamıştır. Örneklem büyüklüğü arttıkça NSO ve ÖSO'ların bazı koşullarda arttığı, bazı koşullarda sabit kaldığı, bazı koşullarda ise önce artıp sonra azaldığı veya önce azalıp sonra arttığı tespit edilmiştir. Bu sonuca paralel olarak Lim ve Drasgow (2017) DINO model temelli 250, 1000 ve 5000 kişilik, Chiu vd. (2017) 30, 100 ve 500 kişilik ve Paulsen (2019) 25, 50, 150 ve 1000 kişilik DINA temelli simülasyon veri setlerinde benzer bulgulara ulaşmışlardır. Teorik olarak POBT, gözlemlenen ve ideal cevap örüntüleri arasındaki uzaklığa göre sınıflandırma yaptığından örneklem büyüklüğüyle ilişkisiz olması beklenir. Bu durum, mevcut araştırmadaki ve diğer çalışmalardan elde edilen sonuçları desteklemektedir.

Mevcut araştırmada madde sayısı (15, 30 ve 45) arttıkça POBT'ye ait NSO ve ÖSO'ların arttıkları belirlenmiştir. Bu sonuç, alanyazın tarafından da desteklenmektedir. Örneğin, Chiu ve Douglas'ın (2013) DINO model temelli 20 ve 40 maddelik Lim ve Drasgow (2017) ise 20, 40 ve 60 maddelik simülasyon veri setlerinde çalışarak mevcut araştırma ile benzer bulguları elde ettikleri görülmektedir. Ayrıca farklı madde sayılarının sınıflandırma oranları üzerindeki etkilerini DINA temelli simülasyon veri setlerinde inceleyen Chiu vd. (2017), Wang (2015), Wang ve Douglas (2015) ve Paulsen (2019) de benzer bulguları elde etmişlerdir. Buradan simülasyon verilerinin üretilmesinde temel alınan modellerin ve araştırma kapsamında incelenen madde sayılarının değişmesine rağmen madde sayısı arttıkça POBT'ye ait NSO ve ÖSO'ların arttığı söylenebilmektedir.

Mevcut araştırmada kayıp veri oranı (0; 0,05 ve 0,10) arttıkça NSO ve ÖSO'ların genel olarak azaldıkları belirlenmiştir. Teorik olarak POBT, gözlemlenen ve ideal cevap örüntüleri arasındaki uzaklığa göre sınıflandırma yaptığından örneklem büyüklüğüyle ilişkisiz olması beklenir. Buna paralel olarak POBT'nin kayıp veri oranından etkilenmemesi veya az miktarda etkilenmesi de beklenmektedir. Mevcut araştırma bulgusu bu durumu destekler niteliktedir. Alanyazında kayıp veri oranının POBT'nin NSO ve ÖSO'larına etkisini inceleyen herhangi bir çalışmaya ulaşılamadığından bu bulgunun farklı veri setlerinde yapılacak çalışmalarla incelenmesi gerektiği düşünülmektedir.

5.1.1.2. YSA'ya ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Durumları

Bu araştırmada nitelik sayısı (3, 5 ve 7) arttıkça YSA'ya ait NSO'ların genel olarak arttığı görülmektedir. Bunun nedeni, az nitelik sayısının bulunduğu koşullarda çok az sayıda ideal cevap örüntüsü elde edilebilmesi olabilir. Çok az sayıda ideal cevap örüntüsü, YSA'daki bağlantı ağırlıklarının tahminlerinde kararsızlığa yol açabildiğinden (Cui vd., 2016) nitelik sayısı arttıkça YSA'ya ait NSO'ların arttığı düşünülmektedir. Ayrıca mevcut araştırmada nitelik sayısı üçten beşe arttırıldığında ÖSO'ların genel olarak azaldığı, nitelik sayısı beşten yediye arttırıldığında ise ÖSO'ların arttığı belirlenmiştir. Bu sonuç, ilgili alanyazındaki bazı araştırmalarla benzerlikler gösterirken bazılarından farklıdır. Örneğin Shu vd. (2013), DINA model temelli simülasyon veri setlerinde nitelik sayısı üçten altıya arttırıldığında YSA'ya ait ÖSO'ların, Cui vd. (2016) ise NSO ve ÖSO'ların azaldıklarını tespit etmişlerdir. Ayrıca Paulsen (2019) çalışmasında nitelik sayısı ikiden dörde arttırıldığında NSO ve ÖSO'ların azaldığını, nitelik sayısının dörtten altıya arttırıldığında ise NSO ve ÖSO'ların arttığını belirlemiştir. Alanyazında YSA'lara ait sınıflandırma oranlarının nitelik sayısına göre nasıl değiştiğinin belirlenmesinde yeterli sayıda çalışmanın bulunmaması ve var olan çalışmaların birbirlerinden farklı sonuçlar içermesi nedeniyle bu konuda daha fazla çalışmaya ihtiyaç olduğu düşünülmektedir.

Mevcut araştırmada örneklem büyüklüğünün (30, 100 ve 500) artırılmasıyla YSA'ya ait NSO ve ÖSO'ların artması, azalması veya sabit kalmasıyla ilgili herhangi bir sistematikliğe rastlanmamıştır. Benzer sonuçla Shu vd. (2013) ve Paulsen'in (2019) çalışmalarında da karşılaşılmaktadır. Shu vd. (2013) 20, 50, 150, 200, 500, 1000 kişilik, Paulsen (2019) ise 25, 50, 150, 1000 kişilik örneklemelerden oluşan veri setlerini inceleyerek küçük ve büyük örneklemde YSA'nın sınıflandırma oranlarının örneklem büyüklüğüyle ilişkili olmadığını belirlemişlerdir. Ayrıca 500 ve 1000 kişilik örneklemeler üzerinde YSA'nın sınıflandırma oranlarını inceleyen Cui vd. (2016) da örneklem büyüklüğünün artmasıyla NSO ve ÖSO'ların artıp azalmaları arasında herhangi bir sistematiklik tespit edememişlerdir. Teorik olarak YSA ideal cevap örüntüleri ve nitelik profilleri kullanılarak eğitildiğinden örneklem büyüklüğünden etkilenmemesi beklenir. Mevcut araştırma ve diğer çalışmalardan elde edilen sonuçlar bu yargıyı destekler niteliktedir.

Mevcut arařtırmada madde sayısı (15, 30 ve 45) arttıkça YSA'nın NSO ve ÖSO'larının azaldığı görölmektedir. İlgili alanyazında bu sonucu destekleyen ve karşı çıkan çalışmalar bulunmaktadır. Örneğın, 10, 20 ve 50 maddenin YSA'ya ait sınıflandırma oranlarına etkisini inceleyen Paulsen (2019), mevcut arařtırmadaki bu sonuç ile benzer bir sonuç elde etse de çalışmalarında 20 ve 40 maddenin etkisini inceleyen Cui vd. (2016) madde sayısı arttıkça NSO ve ÖSO'ların arttığını belirlemişlerdir. Buradan hareketle YSA'ların sınıflandırma oranlarının madde sayısına göre nasıl değıştiğinin belirlenmesi için daha fazla çalışmaya ihtiyaç olduğu düşünölmektedir.

Mevcut arařtırmada kayıp veri oranı (0; 0,05 ve 0,10) arttıkça YSA'nın NSO ve ÖSO'larının az miktarda azaldığı görölmektedir. İdeal tepki örüntöleri ve nitelik profilleriyle eğitilen YSA'nın örneklem büyüklüğündeki değışimden az miktarda etkilenmesi beklenir. Buna paralel olarak YSA'nın kayıp veri oranından etkilenmemesi veya az miktarda etkilenmesi de beklenmektedir. Mevcut arařtırma bulgusu bu durumu destekler niteliktedir. Alanyazında kayıp veri oranının YSA'nın NSO ve ÖSO'larına etkisini inceleyen herhangi bir çalışmaya ulaşılamadığından bu sonucun farklı veri setlerinde yapılacak çalışmalarla incelenmesi gerektiğı düşünölmektedir.

YSA'ya ait NSO ve ÖSO'ların farklı nitelik ve madde sayısı koşullarındaki değışimlerinin genellenememesinin koşullara göre değışen ağ mimarisinden kaynaklanabileceğı düşünölmektedir. Arařtırma kapsamında her bir koşul için deneme sürecine bir gizil katmandan başlanmış ve başlangıç nöron sayısı olarak her bir koşuldaki nitelik sayısı (çıkış katmanı nöron sayısı) seçilmiştir. Ağın bir gizil katmana sahip olması durumunda uygun nöron sayısı belirlendikten sonra benzer süreç ağın iki gizil katmana sahip olduğu durum için izlenmiştir. Ağın en düşük hata değerlerine sahip olduğu gizil katman ve nöron sayısı ağın mimarisini belirlemiştir. Bu nedenle her bir koşul için YSA farklı bir ağ yapısına sahip olmuştur. Ayrıca YSA'da gizil katman sayısı arttıkça girdi ve çıktı katmanları arasındaki ilişki doğrusallıktan uzaklaşmaktadır. Bu durumun elde edilen sonuçların genellenmesini zorlaştırdığı düşünölmektedir.

5.1.1.3. DINO modeline ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Durumları

Bu araştırmada nitelik sayısı (3, 5 ve 7) arttıkça DINO modelinin NSO ve ÖSO'larının azaldığı belirlenmiştir. Bu sonuç ilgili alanyazın tarafından da desteklenmektedir. Örneğin, DINO modeli temel alarak üretilen simülasyon veri setlerinde Chiu ve Douglas (2013) üç ve dört niteliğin, Cui vd. (2016) üç ve altı niteliğin DINO modelin sınıflandırma oranına etkisini benzer şekilde tespit etmişlerdir. Ayrıca DINA model temel alınarak üretilen simülasyon veri setlerinde yapılan çalışmalarda (Chiu vd., 2017; Cui vd., 2012; Cui vd., 2015; Paulsen, 2019 ve Shu vd., 2013) da benzer sonuçlar elde edilmiştir.

Mevcut araştırmada örneklem büyüklüğü (30, 100 ve 500) arttıkça DINO modelinin NSO ve ÖSO'ları arttığı tespit edilmiştir. Bu durum, Chiu vd. (2017), Paulsen (2019) ve Shu vd. (2013) tarafından küçük örneklemle (20, 50, 150, 200, vb.) yapılan çalışmaların sonuçlarıyla benzerlik göstermektedir. Bu çalışmalara göre örneklem büyüklüğü arttıkça DINA modelinin NSO ve ÖSO'ları artmaktadır. Ayrıca çeşitli koşullar altında DINA ve DINO modelin ve iki farklı YSA'nın sınıflandırma oranlarını inceleyen Cui vd. (2016) büyük örneklem koşullarında (500 ve 1000) çalıştıkları için DINO modeline ait NSO'ların genel olarak değişmediğini, ÖSO'ların ise ya değişmediklerini ya da çok az arttığını belirlemişlerdir.

Mevcut araştırmada madde sayısı (15, 30 ve 45) arttıkça DINO modelinin NSO ve ÖSO'ları artmaktadır. Bu sonuç, DINO temelli veri setlerinde madde sayısı koşullarını 20 ve 40 olarak ele alan Chiu ve Douglas (2013) ve Cui vd. (2016) çalışmalarının sonuçlarıyla benzerlik göstermektedir. Ayrıca POBT ile DINA modelinin sınıflandırma oranlarını 30 ve 50 madde koşullarında karşılaştıran Chiu vd. (2017) ve 10, 20 ve 50 madde koşullarında karşılaştıran Paulsen (2019) de benzer bulguları elde etmişlerdir.

Mevcut araştırmada kayıp veri oranı (0; 0,05 ve 0,10) arttıkça DINO modelinin NSO ve ÖSO'ları azalmaktadır. Alanyazında 0 ile 0,40 arasındaki kayıp veri oranının sınıflandırmaya etkisinin incelendiği görülmektedir (bkz. Ayers vd., 2009; Başokçu vd., 2016; Dai, 2017; Kalkan ve Başokçu, 2019; Sünbül, 2018; Xu & von Davier, 2006). Ek olarak alanyazında kayıp veri oranı varlığında DINO modelinin sınıflandırma oranlarını inceleyen herhangi bir çalışmaya ulaşılamazken Başokçu vd. (2016), Dai (2017), Kalkan ve Başokçu (2019) ve Sünbül'ün (2018) çalışmalarında DINA modelinin sınıflandırma oranını inceledikleri görülmektedir. Bu araştırmacıardan sadece Dai (2017) ve Sünbül

(2018), çeşitli kayıp veri oranlarında iki yönlü kayıp veri atama yönteminin DINA modelinin sınıflandırma oranlarına etkisini incelemişlerdir. Dai (2017) çalışmasında kayıp veri oranı olarak 0,5; 0,10; 0,20 ve 0,30'u ele alırken Sünbül (2018) 0,5; 0,10 ve 0,15'i ele almıştır. Simülasyon verilerinin üretilmesinde temel alınan modellerin ve araştırma kapsamında incelenen kayıp veri oranının değişmesine rağmen bu çalışmalarda da kayıp veri oranı arttıkça DINA modelinin sınıflandırma oranının azaldığı belirlenmiştir.

Sonuç olarak, değişen koşullar altında DINO modeline ait bulguların DINO ve DINA model çalışmalarıyla benzerlik gösterdiği görülmektedir. DINA ve DINO modelin ikili (duality) ilişkisi (bkz. Köhn & Chiu, 2016) göz önüne alındığında bunun beklenen bir durum olduğu söylenebilir. Ek olarak, mevcut araştırmada POBT'ye ait NSO ve ÖSO'ların değişimleri, örneklem büyüklüğü koşulunda YSA'ya ait NSO ve ÖSO'ların değişimleri ile, nitelik ve madde sayısı koşullarında ise DINO modeline ait NSO ve ÖSO'ların değişimleri ile benzerlik göstermektedir. POBT'ye YSA'ya ve DINO modeline ait NSO ve ÖSO'ların değişimleri ise sadece kayıp veri oranı koşulunda benzerlik göstermektedir. POBT ve DINO modeline ait NSO ve ÖSO'lardaki değişimin nitelik ve madde sayısı koşullarındaki benzerliği alanyazındaki çalışmalarla da paralellik göstermektedir (bkz. Chiu & Douglas, 2013; Chiu vd., 2017; Paulsen, 2019).

5.1.2. Simülasyon Verisinde POBT'ye, YSA'ya ve DINO Modeline ait NSO ve ÖSO'ların Nitelik Sayılarına (3, 5 ve 7), Örneklem Büyüklüklerine (30, 100 ve 500), Madde Sayılarına (15, 30 ve 45) ve Kayıp Veri Oranlarına (0; 0,05 ve 0,10) Göre Karşılaştırılması

İkinci alt problemde POBT'ye, YSA'ya ve DINO modeline ait NSO ve ÖSO'ların nitelik sayısı, örneklem büyüklüğü, madde sayısı ve kayıp veri oranı koşullarında birlikte nasıl değiştikleri incelenmiştir. Yapılan faktöriyel ANOVA'lar sonucunda tüm koşullar için etkileşim etkileri istatistiksel olarak anlamlı bulunmuştur. Ek olarak, etkileşim etkisinin anlamlı çıktığı durumlarda basit etkiler analiz edilmiştir.

Yapılan analizler sonucunda tüm koşullarda POBT'nin, DINO modelinden biraz düşük fakat karşılaştırılabilir sınıflandırma oranlarına sahipken YSA'nın her zaman POBT ve DINO modelinden daha düşük oranlara sahip olduğu belirlenmiştir. Bu sonuca paralel olarak POBT, YSA ve DINA modelinin sınıflandırma oranları üzerinde çalışan McCoy ve Willse (2014) ve Paulsen (2019) de POBT ve DINO modelinin sınıflandırma oranlarının

birbirlerine benzerken YSA'nın her zaman POBT ve DINO modelinden daha düşük sınıflandırma oranlarına sahip olduğunu tespit etmişlerdir.

Shu vd. (2013) madde ayırt ediciliğinin düşük olduğu (s ve g parametre değerleri 0,2 ile 0,4 arasındadır) ve karmaşık yapıli Q matris yapısının bulunduđu simülasyon veri setlerinde YSA, MCMC-YSA, JMLE-YSA ve DINA modelinin sınıflandırma oranlarını karşılaştırmışlardır. Shu ve diğerlerinin 2013 yılında yaptıkları bu çalışmada DINA modelin dört niteliğın bulunduđu durumlarda örneklem büyüklüğü 50 ve daha az olduğunda, altı niteliğın bulunduđu durumlarda ise örneklem büyüklüğü 150 ve daha az olduğunda kestirim yapamadığı görülmektedir. Hesaplama kapasitesindeki artış ve maddelere ait parametre kestirimlerinde EM algoritmasının kullanılmasıyla birlikte DINA modelinin 25 ve 30 kişinin bulunduđu küçük örneklemelerde de kestirim yapabildiğı belirlenmiştir (bkz. Chiu vd., 2017; Paulsen, 2019). Mevcut araştırmada da en küçük örneklem büyüklüğü olan 30 kişilik koşullarda DINO modelin kestirim yaptığı görülmektedir.

Hesaplama kapasitesindeki artış ve EM algoritmasının kullanımına ek olarak DINO modelinin küçük örneklemelerde POBT ve YSA'dan yüksek sınıflandırma oranlarına sahip olmasının nedenin simülasyon veri yapısının telafisel modellerden DINO model temel alınarak üretilmesinden kaynaklandığı düşünülmektedir. Ek olarak, veri üretiminde Q matris yapısını oluşturan nitelik ve madde dağılımlarının dengeli olmasına dikkat edilmiştir. Veri üretiminde DINO modelin temel alınması ve madde-nitelik dağılımlarının dengeli olmasının DINO modelin iyi sınıflandırma yapabilmesinde yardımcı olduğu düşünülmektedir. Alanyazında bu durumla benzer ve bu durumdan farklı çalışmalar bulunmaktadır. Mevcut araştırma sonucuna paralel olarak Chiu ve Douglas (2013), veri yapısına en uygun parametrik model ile maksimum olasılık tahmininin daha iyi kestirim yaptığını bu nedenle DINO temelli bazı küçük simülasyon örneklemelerinde DINO modelinin POBT'den daha iyi kestirimlerde bulunabileceğini belirtmişlerdir. Chiu ve Douglas'ın (2013) yaptıkları çalışmada elde edilen bulgular da bu durumu destekler niteliktedir. POBT ile DINA modeline ait bu sonuçlara benzer şekilde Cui vd. (2016) ve Shu vd. (2013) de çalışmalarında DINA temelli simülasyon veri setlerinde DINA modelinin YSA'dan yüksek sınıflandırma oranlarına sahip olduğunu belirlemişlerdir.

Alanyazında POBT'nin ve YSA'nın DINO veya DINA modelinden daha yüksek ve düşük sınıflandırma oranlarına sahip olmasının koşullara göre değiştiğini ifade eden çalışmalar bulunmaktadır. Örneğın DINA temelli veri setlerinde yaptıkları çalışmada McCoy ve

Willse (2014) bazı koşullarda DINA modelinin sınıflandırma oranlarının POBT'den az bir farkla, YSA'dan ise büyük bir farkla yüksek olduğunu tespit etmişlerdir. Chiu vd. (2017) 100 veya daha küçük örneklerde POBT'nin DINA modelinden daha yüksek sınıflandırma oranlarına sahipken 500 kişilik örnekte DINA modelinin POBT'den daha yüksek sınıflandırma oranlarına sahip olduğunu belirtmişlerdir. Paulsen (2019) ise POBT'nin örneklem büyüklüğü 25 iken madde sayısının az ve madde ayırt ediciliğinin düşük olduğu koşullarda DINA modelinden daha iyi sınıflandırma oranlarına sahip olduğunu belirlemiştir. Son olarak, Ma, de la Torre ve Xu (2020) 30 ve 50 örneklerde POBT'nin DINA modelinden, 200 ve 500 örneklerde ise DINA modelinin POBT'den daha yüksek sınıflandırma oranlarına sahip olduğunu tespit etmişlerdir.

Alanyazında POBT ve YSA'nın DINA ve DINO modeline göre daha yüksek sınıflandırma oranlarına sahip olduğunu belirten çalışmalar da bulunmaktadır. Örneğin, Akbay (2016) çalışmasında DINA ve DINO temelli 250, 500 ve 1000 kişilik veri setleri oluşturarak POBT, DINA ve DINO modelinin çeşitli koşullar altında sınıflandırma oranlarını incelemiştir. Akbay (2016) DINO temelli veri seti üzerinde POBT'nin DINO modelinden, DINA temelli veri seti üzerinde ise POBT'nin DINA modelinden daha yüksek sınıflandırma oranlarına sahip olduğunu tespit etmiştir. Guo vd. (2017) ise DINA temelli 5000 kişilik örneklem üzerinde yaptıkları çalışmalarında YSA'nın DINA modelinden daha yüksek sınıflandırma oranına sahip olduğunu belirlemişlerdir.

Sonuç olarak, POBT (Chiu & Douglas, 2013) ve YSA'nın (Shu vd., 2013) 2013 yılında bilişsel tanılamada kullanılmaya başlandıkları göz önüne alındığında birbirinden farklı sonuçlar elde eden çalışmaların bulunmasının olası olduğu düşünülmektedir. Bu nedenle POBT ve YSA'ların sınıflandırma oranlarının farklı koşullarda ve farklı BTM'lerle incelenmesi gerektiği düşünülmektedir.

5.1.3. Gerçek Veri Setinde (PISA 2015 İPÇ) POBT'nin, YSA'nın ve DINO Modelinin Sınıflandırma Oranlarının Benzerliğinin Nitelik Sayılarına (3, 7 ve 11) ve Örneklem Büyüklüklerine (30, 100 ve 500) Göre Durumları

Mevcut araştırmada gerçek veri setine ait ilk Q matrisleri, PISA uygulamasına ait raporlarda (bkz. OECD, 2017a; 2017b) verilen bilgiler doğrultusunda oluşturulmuştur. Bu raporlar (bkz. OECD, 2017a; 2017b) ve Tablo 5 temel alınarak PISA 2015 İPÇ yapısını temsil eden üç farklı Q matrisi oluşturulabilmektedir. PISA 2015 İPÇ uygulamasında A3

hücrelerini ölçmeye yönelik herhangi bir madde olmadığı için 12 nitelikli olarak oluşturulması gereken Q matrisi 11 nitelikli olarak oluşturulmuştur. Daha sonra PVAF yöntemi ile Q matrisi geçerlemesi yapılmıştır. Her iki Q matris türünden elde edilen bulgular beraber yorumlanmıştır.

Mevcut araştırmada geçerlemesi yapılan Q matrislerinin analizi sonucu elde edilen NSOB ve ÖSOB'lar teknik raporlardan elde edilen Q matrislerinin analizi sonucu elde edilen NSOB ve ÖSOB'lardan genel olarak daha yüksektir. Bu bölümde her iki Q matris türünden elde edilen ortak bulgular tartışılmıştır. POBT ile DINO model arasındaki NSOB ve ÖSOB'ların benzerliği POBT ile YSA ve YSA ile DINO arasındaki benzerlikten daha fazla olduğu görülmektedir. Ek olarak nitelik sayısı (3, 7 ve 11) arttıkça POBT ile YSA ve POBT ile DINO modeline ait NSOB'ların genel olarak önce azalıp sonra artarken ÖSOB'ların genel olarak azaldığı belirlenmiştir.

Mevcut araştırmada örneklem büyüklüğünün (30, 100 ve 500) artmasıyla POBT, YSA ve DINO modeline ait NSOB ve ÖSOB'ların artması, azalması veya sabit kalmasıyla ilgili herhangi bir sistematikliğe rastlanmamıştır. POBT ve YSA'ya ait NSO ve ÖSO'larının teorik olarak örneklem büyüklüğünden etkilenmedikleri göz önüne alındığında bu beklenen bir durumdur. Alanyazında POBT ve YSA ile DINO veya DINA model sınıflandırma oranlarının benzerliğini gerçek veri üzerinde inceleyen herhangi bir çalışmaya ulaşılamamıştır. Alanyazında sadece Chiu ve Douglas (2013) tarafından POBT ile HODINA'nın, Chiu vd. (2017) tarafından GPOBT ile G-DINA'nın ve Lim ve Drasgow (2017) tarafından uzman kanısı ile POBT'nin sınıflandırma oranlarının benzerliklerini tek bir örneklem üzerinde hesapladıkları görülmektedir.

Sonuç olarak, mevcut araştırmada simülasyon veri setlerinde NSO ve ÖSO'ların, gerçek veri setlerinde ise NSOB ve ÖSOB'ların düşük hesaplandığı ve NSO'ların ÖSO'lardan, NSOB'ların ise ÖSOB'lardan daha yüksek olduğu belirlenmiştir. POBT ile DINO modelinin sınıflandırma oranları arasındaki benzerlik hem simülasyon hem de gerçek veri setlerinde gözlemlenmektedir. Ayrıca hem simülasyon hem de gerçek veri setlerinde örneklem büyüklüğünün artmasıyla POBT ve YSA'nın sınıflandırma oranlarına ve bu oranların benzerliklerinin artış veya azalışlarına ait herhangi bir sistematikliğe rastlanmamıştır. POBT, YSA ve DINO modelinin sınıflandırma oranlarının ve bu oranların benzerliklerinin nitelik sayısı koşullarındaki değişimleri simülasyon ve gerçek veri setlerinde farklılaşmaktadır. Örneğin, nitelik sayısı arttıkça POBT ve DINO modeline ait NSO ve ÖSO'lar azalırken YSA'ya ait NSO'lar artmakta, POBT ile YSA'ya ve POBT ile

DINO modeline ait NSOB'lar ise önce azalıp sonra artmaktadır. Bu duruma ek olarak nitelik sayısı arttıkça POBT ve DINO modeline ait ÖSO'lar azalırken YSA'ya ait ÖSO'lar önce azalıp sonra artmakta, POBT, YSA ve DINO modeline ait ÖSOB'lar ise azalmaktadır.

5.2. Öneriler

Bu başlık altında kullanıcılara ve araştırmacılara yönelik önerilere yer verilmiştir.

5.2.1. Kullanıcılar için öneriler

Bir tanı testinde her bir nitelik yeterli güvenirliği sağlamak için yeterli sayıda madde ile ölçülmelidir. Bu nedenle testi oluşturan maksimum nitelik ve madde sayısı dikkatle belirlenmelidir. Öğrencilerin yorulması veya zaman sınırlaması düşünüldüğünde bir testte çok sayıda niteliğin güvenirlik derecesi kabul edilebilir düzeyin üstünde bir şekilde ölçülmesi gerçekçi değildir. Nitekim mevcut araştırmada 11 niteliğin bulunduğu veri setinde POBT, YSA ve DINO modeline ait ÖSOB'ların sıfıra yakın olduğunu görülmektedir. Bu nedenle karmaşık yapıların ölçülmeye ve değerlendirilmeye çalışıldığı günümüzde yapıların mümkün olduğunca az nitelikle ifade edilmesi önerilmektedir.

Mevcut araştırmada POBT ve DINO modelinin sınıflandırma oranlarının birbirlerine benzer olduğu görülmektedir. Öğrencilerin doğru nitelik sınıflarına yerleştirilmesinde POBT ve DINO modeli sonuçlarının birlikte değerlendirilmesinin sınıflandırma güvenirliğini arttıracakı düşünülmektedir. Bu nedenle küçük örneklemelerde bilişsel tanılama yapmak isteyen uygulayıcıların POBT ve DINO modelini birlikte kullanmaları önerilmektedir.

5.2.2. Araştırmacılar için öneriler

Bu araştırmada simülasyon veri setleri telafisel modellerden DINO modeli ve karmaşık Q matris yapısı temel alınarak üretilmiştir. Simülasyon veri setlerinde nitelikler arası ilişki orta düzeyde tutulmuş ve çok değişkenli normal dağılım kullanılarak dağılımları belirlenmiştir. Ayrıca simülasyon veri setlerindeki maddeler orta düzey ayırt ediciliğe sahiptir. Simülasyon verisine ait bu özellikler değiştirilerek POBT, YSA ve DINO modelinin sınıflandırma oranlarının incelenmesi önerilmektedir. Örneğin, telafisel

modellerden NIDO gibi modeller veya telafisel olmayan modeller temel alınarak veri setleri üretilebilir. Basit ve karmaşık yapılı Q matrisleri bir arada kullanılarak bu çalışma tekrarlanabilir. Nitelikler arası ilişkiler düşük veya yüksek tutulurken araştırma örnekleminde her bir nitelik örüntüsünün eşit temsil edilmesine izin veren tekdüze dağılım seçilebilir. Son olarak maddelerin ayırt edicilikleri düşük, orta ve yüksek olarak değiştirilebilir.

Mevcut araştırmada verilerin üretilmesi ve analizinde R-Studio programı kullanılmış, DINO modeli analizi için CDM, POBT analizi için NPCD, YSA analizi için neuralnet, kayıp veri oluşturma ve atama için missForest ve TestDataImputation paketlerinden yararlanılmıştır. POBT, YSA ve DINO modelinin sınıflandırma oranları farklı programlarda veya farklı paketler yardımıyla incelenmesi önerilmektedir. Örneğin, POBT, YSA ve DINO modelinin sınıflandırma oranları farklı algoritmalara sahip Mplus, Ox, Python gibi veri üretme ve analiz programları kullanılarak karşılaştırılabilir. Ayrıca verilerin üretilmesi ve DINO modeli analizi için GDINA paketi (Ma, de la Torre, Sorrel & Jiang, 2020), YSA analizi için AMORE (Limas vd., 2015), kayıp veri oluşturmak için Amelia (Honaker, King & Blackwell, 2019) ve kayıp veri atama için mice (van Buuren vd., 2020) veya mi (Gelman vd., 2015) paketleri kullanılabilir.

Mevcut araştırmada POBT ve YSA'nın sınıflandırma oranları DINO modelinin sınıflandırma oranları ile karşılaştırılmıştır. POBT ve YSA'nın sınıflandırma oranlarının farklı BTM'lerle karşılaştırılması önerilmektedir. Ayrıca POBT ve YSA maddelerin ikili puanlandığı koşullarda uygulanabilmektedir. Niteliklerin ve maddelerin çoklu puanlandığı durumlarda kullanılabilecek parametrik olmayan yöntemlerin geliştirilmesine ihtiyaç olduğu düşünülmektedir. Bu duruma ek olarak POBT ve YSA sonsal dağılım hesaplanmasına izin vermemektedir. Bu POBT'ye ve YSA'ya ait sınıflandırma oranlarının gerçek veri setlerinde hesaplanamamasına sebep olmaktadır. Bu nedenle gerçek veri setlerinde sınıflandırma oranlarının hesaplanabileceği farklı parametrik olmayan yöntemlere ihtiyaç olduğu düşünülmektedir.

Mevcut araştırmada POBT analizi için ağırlıklandırılmış Hamming uzaklık ölçütü kullanılmıştır. Bu ölçüt yerine Hamming, cezalandırılmış Hamming veya başka bir uzaklık ölçütü kullanılarak POBT'nin, YSA'nın ve DINO modelinin sınıflandırma oranlarının incelenmesi önerilmektedir. Ek olarak YSA yapısı ve eğitim sürecinin mevcut araştırmadan farklı şekillerde ele alınarak POBT'nin, YSA'nın ve DINO modelinin sınıflandırma oranlarının karşılaştırılması da önerilmektedir. Örneğin, farklı başlangıç

nöronu atama yöntemi kullanılabilir. Mevcut araştırmada YSA eğitimi için ağırlık geri takibi ile esnek geri yayılım algoritması kullanılmıştır. Alternatif olarak geri yayılım, ağırlık takibi olmayan veya modifiye edilmiş küresel yakınsak esnek yayılım algoritmaları kullanılabilir. Mevcut araştırmada YSA eğitiminde geleneksel geri yayılım algoritması kullanılmadığı için ağırlık öğrenme oranı parametresi için özel bir değer kullanılmamıştır. Geleneksel geri yayılım algoritması kullanılarak ağırlık öğrenme hızındaki değişimin sınıflandırma oranlarındaki etkisi incelenebilir. Ayrıca mevcut araştırmada AMORE paketi kullanılmadığı için ağırlık momentum parametresi için özel bir değer kullanılmamıştır. AMORE paketi kullanılarak ağırlık momentumundaki değişimin sınıflandırma oranlarına etkisi incelenebilir. Ek olarak lojistik fonksiyon yerine hiperbolik tanjant fonksiyonu seçilebilir veya çapraz entropi yerine hata karelerinin toplamı ile hata miktarı değerlendirilebilir.

Mevcut araştırmada simülasyon faktörleri olarak nitelik sayısı, örneklem büyüklüğü, madde sayısı ve kayıp veri oranı belirlenmiştir. Bu faktörlerin dışında Q matrisinin yanlış belirlenme oranının sınıflandırma oranları üzerindeki etkisi de incelenebilir. Ayrıca mevcut araştırma kapsamında ele alınan faktörlerin düzeyleri değiştirilerek POBT'nin, YSA'nın ve DINO modelinin sınıflandırma oranlarının incelenmesi önerilmektedir.

Alanyazında kayıp veri oranının POBT ve YSA'nın sınıflandırma oranlarına etkisini inceleyen herhangi bir çalışmaya ulaşılamadığından bu bulgunun farklı veri setlerinde yapılacak çalışmalarla incelenmesi önerilmektedir. Mevcut araştırmada YSA'ya ait sınıflandırma oranlarının nitelik ve madde sayısı koşullarındaki değişimleriyle ilgili genellenebilir sonuçlara ulaşılamamıştır. YSA'ya ait sınıflandırma oranlarının nitelik ve madde sayısı koşullarında nasıl değiştiklerini inceleyen daha fazla araştırmanın yapılması önerilmektedir. Ek olarak, mevcut araştırmada POBT, YSA ve DINO modelinin sınıflandırma oranlarının ve bu oranların benzerliklerinin nitelik sayısı koşullarındaki değişimleri, simülasyon ve gerçek veri setlerinde farklılaşmaktadır. Bu durumu inceleyen daha fazla çalışma yapılması önerilmektedir.

POBT, YSA ve DINO modeline ait sınıflandırma oranları incelendiğinde PISA 2015 İPÇ yapısının üç nitelikli Q matrisi ile incelenmesinin daha uygun olduğu görülmektedir. Bu üç nitelik OECD (2017a; 2017b) teknik raporlarında “ortak bir anlayışın oluşturulması ve sürdürülmesi”, “problemi çözmek için uygun eylemlerin belirlenmesi” ve “takım organizasyonunun oluşturulması ve sürdürülmesi” yeterlik alanları olarak ifade

edilmektedir. PISA 2015 İPÇ yapısını farklı yöntemlerle incelemek isteyen araştırmacılara İPÇ yapısını bu üç alt yeterlik alanına göre incelemeleri önerilmektedir.



KAYNAKLAR

- Akbay, L. (2016). Relative efficiency of the nonparametric approach on attribute classification for small sample cases. *Journal of European Education* 6(1), 43-59. doi: 10.18656/jee.20599
- Al-Shareef, A. J., & Abbod, M. F. (2010). *Neural networks initial weights optimization*. Paper presented at the 12th International Conference on Modelling and Simulation, Cambridge, UK. Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5481067>
- American Management Association (2010). *Critical skills survey: Executive summary*. Retrieved from <http://www.p21.org/storage/documents/Critical%20Skills%20Survey%20Executive%20Summary.pdf>
- Anastasiadis A. D., Magoulas, G. D., & Vrahatis, M. N. (2005) New globally convergent training scheme based on the resilient propagation algorithm. *Neurocomputing*, 64, 253-270. <https://doi.org/10.1016/j.neucom.2004.11.016>
- Arı, A. & Berberle, E. (2017). Yapay sinir ağları ile tahmin ve sınıflandırma problemlerinin çözümü için ara yüz tasarımı. *ACTA INFOLOGICA*, 1(2). 55-73.
- Ayers, E., Nugent, R., & Dean, N. (2008). *Skill set profile clustering based on student capability vectors computed from online tutoring data*. Proceedings of the Educational Data Mining -EDM2008: 1st International Conference on Educational Data Mining. Retrieved from <https://www.semanticscholar.org/paper/Skill-set-profile-clustering-based-on-student-from-Ayers-Nugent/6525d161295b327081fc85a53cc5cf36f681e6cb>

- Ayers, E., Nugent, R., & Dean, N. (2009). *A Comparison of student skill knowledge estimates*. Proceedings of the Educational Data Mining-EDM2009: 2nd International Conference on Educational Data Mining. Retrieved from <https://eric.ed.gov/?id=ED539062>
- Başokçu, T. O. (2011). *Bağıl ve mutlak değerlendirme ile DINA modele göre yapılan sınıflandırmaların geçerliğinin karşılaştırılması*. Doktora Tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Başokçu, T. O., Kalkan, Ö. K., & Öğretmen, T. (2016). *DINA modele dayalı madde parametre kestiriminde kayıp veri ele alma yöntemleri etkisinin incelenmesi*. V. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresi'nde sunulmuş bildiri, Akdeniz Üniversitesi, Antalya. http://epod2016.akdeniz.edu.tr/_dinamik/333/53.pdf sayfasından erişilmiştir.
- Bernaards, C. A., & Sijtsma, K. (2000). Influence of imputation and EM methods on factor analysis when item nonresponse in questionnaire data is nonignorable. *Multivariate Behavioral Research*, 35(3), 321–364. https://doi.org/10.1207/S15327906MBR3503_03
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Briggs, D. C., & Circi, R. (2017) Challenges to the use of artificial neural networks for diagnostic classifications with student test data. *International Journal of Testing*, 17(4), 302-321. <https://doi.org/10.1080/15305058.2017.1297816>
- Borysewicz, P., & Steinhauer, E. (2014). *PISA 2015 collaborative problem-solving*. Paper presented at the Innovative Assessment of Collaboration Conference, Arlington, Virginia. Retrieved from https://www.ets.org/research/events/collaboration_conference_2014
- Chang, C.-J., Chang, M.-H., Chiu, B.-C., Liu, C.-C., Chiang, S.-H. F., Wen, C.-T., ... Chen, W. (2017). An analysis of student collaborative problem-solving activities mediated by collaborative simulations. *Computers & Education*, 114, 222-235. <https://doi.org/10.1016/j.compedu.2017.07.008>
- Chen, J., & de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, 37(6), 419-437. <http://dx.doi.org/10.1177/0146621613479818>

- Chiu, C. (2008). *Cluster analysis for cognitive diagnosis: Theory and applications*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/>
- Chiu, C.-Y., & Douglas, J. (2013). A nonparametric approach to cognitive diagnosis by proximity to ideal response patterns. *Journal of Classification*, 30, 225-250. <https://doi.org/10.1007/s00357-013-9132-9>
- Chiu, C.-Y., Douglas, J., & Li, X. (2009). Cluster analysis for cognitive diagnosis: Theory and applications. *Psychometrika*, 74(4), 633–665. <https://doi.org/10.1007/s11336-009-9125-0>
- Chiu, C.-Y. & Köhn, H.-F. (2015). Consistency of cluster analysis for cognitive diagnosis: The DINO model and the DINA model revisited. *Applied Psychological Measurement*, 39(6), 1–15. <https://doi.org/10.1177/0146621615577087>
- Chiu, C.-Y., & Köhn, H.-F. (2016). Consistency of cluster analysis for cognitive diagnosis: The reduced reparameterized unified model and the general diagnostic model. *Psychometrika*, 81(3), 585–610. <https://doi.org/10.1007/s11336-016-9499-8>
- Chiu, C.-Y., & Köhn, H.-F. (2019). Consistency theory for the general nonparametric classification method. *Psychometrika*, 84(3), 830–845. <https://doi.org/10.1007/s11336-019-09660-x>
- Chiu, C.-Y., Sun, Y., & Bian, Y. (2017). Cognitive diagnosis for small educational programs: The general nonparametric classification method. *Psychometrika*, 83(2), 355–375. <https://doi.org/10.1007/s11336-017-9595-4>
- Clinch, J.J. & Keselman, H.J. (1982). Parametric alternatives to the analysis of variance. *Journal of Educational Statistics*, 7(3), 207–214. <https://doi.org/10.2307/1164645>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Cui, Y., Gierl, M., & Chang, H.-H. (2012). Estimating Classification Consistency and Accuracy for Cognitive Diagnostic Assessment. *Journal of Educational Measurement*, 49(1), 19–38. <https://doi.org/10.1111/j.1745-3984.2011.00158.x>

- Cui, Y., Gierl, M., & Guo, Q. (2016). Statistical classification for cognitive diagnostic assessment: An artificial neural network approach. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 36(6), 1065-1082. <https://doi.org/10.1080/01443410.2015.1062078>
- Cui, Y., Gierl, M., & Guo, Q. (2017). The rule space and attribute hierarchy methods. In A. A. Rupp & J. P. Leighton (Eds.), *The handbook of cognition and assessment frameworks, methodologies, and applications* (pp. 354-378). UK: John Wiley & Sons.
- Cui, Y., & Leighton, J. P. (2009). The hierarchy consistency index: Evaluating person fit for cognitive diagnostic assessment. *Journal of Educational Measurement*, 46(4), 429-449. <https://doi.org/10.1111/j.1745-3984.2009.00091.x>
- Cui, Y., Leighton, J. P., Gierl, M. J., & Hunka, S. (2006, April). *A person-fit statistic for the attribute hierarchy method: The hierarchy consistency index*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.
- Cunningham, K. M., & Ogilvie, J. C. (1972). Evaluation of hierarchical grouping techniques: A preliminary study. *Computer Journal*, 15, 209-213.
- Dai, S. (2017). *Investigation of missing responses in implementation of cognitive diagnostic models*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/>
- Dai, S., Wang, X., & Svetina, D. (2019, March). *Test Data Imputation* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/TestDataImputation/index.html>
- de Boeck, P. & Scalise, K. (2019). Collaborative problem solving: Processing actions, time, and performance. *Frontiers in Psychology*, 10(1280), 1-9. <https://doi.org/10.3389/fpsyg.2019.01280>
- de la Torre, J. (2008). An empirically-based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement*, 45(4), 343-362. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- de la Torre, J., & Chiu, C.-Y. (2015). A general method of empirical Q-matrix validation. *Psychometrika*, 81(2), 253-273. <https://doi.org/10.1007/s11336-015-9467-8>

- de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3), 333-353. <https://doi.org/10.1007/BF02295640>
- de la Torre, J., & Douglas, J. A. (2008). Model evaluation and multiple strategies in cognitive diagnosis: An analysis of fraction subtraction data. *Psychometrika*, 73(4), 595. <https://doi.org/10.1007/s11336-008-9063-2>
- de la Torre, J., & Minchen, N. (2014). Cognitively diagnostic assessments and the cognitive diagnosis model framework. *Psicologi'a Educativa*, 20(2), 89-97. <https://doi.org/10.1016/j.pse.2014.11.001>
- de la Torre, J., Hong, Y., & Deng, W. (2010). Factors affecting the item parameter estimation and classification accuracy of the DINA model. *Journal of Educational Measurement*, 47(2), 227-249. <https://doi.org/10.1111/j.1745-3984.2010.00110.x>
- Boeck, P. & Scalise, K. (2019). Collaborative problem solving: Processing actions, time, and performance. *Frontiers in Psychology*, 10(1280), 1-9. <https://doi.org/10.3389/fpsyg.2019.01280>
- DiBello, L. V., & Stout, W. (2007). Guest editors' introduction and overview: IRT-based cognitive diagnostic models and related methods. *Journal of Educational Measurement*, 44(4), 285-291. <https://doi.org/10.1111/j.1745-3984.2007.00039.x>
- DiBello, L. V., Roussos, L. A., & Stout, W. (2007). Review of cognitively diagnostic assessment and a summary of psychometric models. *Handbook of statistics*, 26, 979-1030. [https://doi.org/10.1016/S0169-7161\(06\)26031-0](https://doi.org/10.1016/S0169-7161(06)26031-0)
- Dijkstra, T., & Henseler, J. (2011). Linear indices in nonlinear structural equation models: Best fitting indices and other composites. *Quality & Quantity*, 45(6), 1505-1518. <https://doi.org/10.1007/s11135-010-9359-z>
- Elkonca, F., Karakaya, İ., Şata, M., & Ceyhan, G. (2018). Dmf kaynaklarının belirlenmesinde örtük sınıf dmf yaklaşımının kullanılması: PISA-2015 Türkiye örneği. 6. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresinde sunulmuştur, Prizren, Sırbistan.
- Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods*, 3(3), 380-396. <https://doi.org/10.1037/1082-989X.3.3.380>

- Erkuş, A. (2013). *Davranış bilimleri için bilimsel araştırma süreci*. (4. Baskı). Ankara: Seçkin Yayıncılık.
- ETS (2016). PISA 2015 released field trial cognitive items–CPS. Retrieved from https://www.pi.ac.cy/keea/pisa2015/ReleasedMaterial/CPS_eng_ell.pdf
- Fisher, G. (1973). Linear logistic test model as an instrument in educational research. *Psychologica*, 37(6), 359-374.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. New York, NY: McGraw-Hill.
- Fritsch, S., Guenther, F., Wright, M. N., Suling, M., & Mueller, S. M. (2019, February). *neuralnet* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf>
- Garson, G. D. (1998). *Neural networks: An introductory guide for social scientists*. London: SAGE.
- Gelman, A., Hill, J., Su, Y.-S., Yajima, M., Pittau, M., Goodrich, B., ... Kropko, J. (2015, April). *mi* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/mi/mi.pdf>
- Gierl, M. J. (2007). Making diagnostic inferences about cognitive attributes using the rule-space model and attribute hierarchy method. *Journal of Educational Measurement*, 44(4), 325–340. <https://doi.org/10.1111/j.1745-3984.2007.00042.x>
- Gierl, M. J., Cui, Y., & Hunka, S. (2007, April). *Using connectionist models to evaluate examinees' response patterns on tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.
- Gierl, M. J., Cui, Y., & Hunka, S. (2008). Using connectionist models to evaluate examinees' response patterns on tests. *Journal of Modern Applied Statistical Methods*, 7, 234–245. Doi: 10.22237/jmasm/1209615480
- Gierl, M. J., Cui, Y. & Zhou, J. (2009). Reliability and attribute-based scoring in cognitive diagnostic assessment. *Journal of Educational Measurement*, 46(3) 293–313. <https://doi.org/10.1111/j.1745-3984.2009.00082.x>

- Gierl, M. J., Leighton, J. P., & Hunka, S. M. (2007). Using the attribute hierarchy method to make diagnostic inferences about respondents' cognitive skills. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 242–274). Cambridge, UK: Cambridge University.
- Gierl, M. J., Wang, C., & Zhou, J. (2008). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills in algebra on the SAT[®]. *Journal of Technology, Learning, and Assessment*, 6(6), 1-50.
- Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, 49, 732–764.
- Guo, Q., Cutumisu, M., Cui, Y. (2017). *A neural network approach to estimate student skill mastery in cognitive diagnostic assessments*. Poster presented at the 10th International Conference on Educational Data Mining. Retrieved from <https://era.library.ualberta.ca/items/6863ba3d-2fcf-44d1-b345-060bc72ebc8a>
- Günther, F., & Fritsch, S. (2010). Neuralnet: Training of neural networks. *The R Journal*, 2(1), 30-38. Haung, W.-C. (2017). Exploring Science-Scenario CPS Behaviors of Taiwanese Students at Age 15 in Human-Agent System with Multiple-Choice Items. (Master dissertation). Retrieved from <https://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi/login?o=dnclcdr&s=id=%22105NTCT0629003%22.&searchmode=basic>
- Haung, W.-C. (2017). Exploring Science-Scenario CPS Behaviors of Taiwanese Students at Age 15 in Human-Agent System with Multiple-Choice Items. (Master dissertation). Retrieved from <https://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi/login?o=dnclcdr&s=id=%22105NTCT0629003%22.&searchmode=basic>
- He, Q., von Davier, M., Greiff, S., Steinhauer, E. W., & Borysewicz, P. B. (2017). Collaborative problem-solving measures in the Programme for International Student Assessment (PISA). In A. A. von Davier, M. Zhu & P. C. Kyllonen (Eds.), *Methodology of educational measurement and assessment: Innovative assessment of collaboration* (pp. 95-112). Princeton: Springer.
- Henson, R. A. & Templin, J. (2007, April). *Large-scale language assessment using cognitive diagnosis models*. Paper presented at the annual meeting of the National Council on Measurement in Education (NCME), Chicago, IL.

- Herborn, K., Mustafic, M., & Greiff, S. (2017). Mapping an experiment-based assessment of collaborative behavior onto collaborative problem solving in PISA 2015: A cluster analysis approach for collaborator profiles. *Journal of Educational Measurement Spring*, 54(1), 103–122. <https://doi.org/10.1111/jedm.12135>
- Herborn, K., Stadler, M., Mustafić, M., & Greiff, S. (2020). The assessment of collaborative problem solving in PISA 2015: Can computer agents replace humans? *Computers in Human Behavior*, 104, 1-9. <https://doi.org/10.1016/j.chb.2018.07.035>
- Honaker, J., Gary King, G., & Blackwell, M. (2019, November). *Amelia* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/Amelia/index.html>
- Hubert, L. J., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2, 193-218.
- Huebner, A., & Wang, C. (2011). A note on comparing examinee classification methods for cognitive diagnosis models. *Educational and Psychological Measurement*, 71(2), 407-419. <https://doi.org/10.1177/0013164410388832>
- IBM Corp. (2019). *IBM SPSS Statistics for Windows*. Version 22.0. Armonk, NY: IBM Corp.
- Intrator, O., & Intrator, I. (2001). Interpreting neural-network results: A simulation study. *Computational Statistics & Data Analysis*, 37(3), 373-393. [https://doi.org/10.1016/S0167-9473\(01\)00016-0](https://doi.org/10.1016/S0167-9473(01)00016-0)
- Johnson, M. S., & Sinharay, S. (2018). Measures of agreement to assess attribute-level classification accuracy and consistency for cognitive diagnostic assessments. *Journal of Educational Measurement*, 55(4), 635–664. <https://doi.org/10.1111/jedm.12196>
- Johnson, M. S., & Sinharay, S. (2020). The reliability of the posterior probability of skill attainment in diagnostic classification models. *Journal of Educational and Behavioral Statistics*, 45(1), 5–31. <https://doi.org/10.3102/1076998619864550>
- Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), 258–272. <https://doi.org/10.1177/01466210122032064>

- Kalkan, Ö. K., & Başokçu, T. O. (2019). The effect of the item–attribute relation on the DINA model estimations in the presence of missing data. *PAU Journal of Education*, 46, 290-306. <https://doi.org/10.9779/pauefd.546797>
- Kato, K. (2009). *Improving efficiency of cognitive diagnosis by using diagnostic items and adaptive testing*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/>
- Kavzoglu, T., & Mather, P. M. (2003) The use of backpropagating artificial neural networks in land cover classification. *International Journal of Remote Sensing*, 24(23), 4907-4938. <https://doi.org/10.1080/0143116031000114851>
- Kirk, R. (2014). *Experimental design: Procedures for the behavioral sciences*. Thousand Oaks: SAGE.
- Köhn, H.-F., & Chiu, C.-Y. (2017). A procedure for assessing the completeness of the Q-matrices of cognitively diagnostic tests. *Psychometrika*, 82(1), 112–132. <https://doi.org/10.1007/s11336-016-9536-7>
- Köhn, H.-F., Chiu, C.-Y., & Brusco M. J. (2015). Heuristic cognitive diagnosis when the Q-matrix is unknown. *British Journal of Mathematical and Statistical Psychology*, 68(2), 268–291. <https://doi.org/10.1111/bmsp.12044>
- Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2012). The impact of model misspecification on parameter estimation and item-fit assessment in log-linear diagnostic classification models. *Journal of Educational Measurement*, 49(1), 59-81. <https://doi.org/10.1111/j.1745-3984.2011.00160.x>
- Kuo, B.-C., Liao, C.-H., Pai, K.-C., Shih, S.-C., Li, C.-H., & Mok, M. M. C. (2019). Computer-based collaborative problem-solving assessment in Taiwan. *Educational Psychology*, 40(9), 1164-1185, <https://doi.org/10.1080/01443410.2018.1549317>
- Lamb, R. L., Annetta, L., Vallett, D. B., & Troy D. S. (2014). Cognitive diagnostic like approaches using neural-network analysis of serious educational videogames. *Computers & Education*, 70, 92–104. <https://doi.org/10.1016/j.compedu.2013.08.008>

- Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoaka's rule-space approach. *Journal of Educational Measurement*, 41(3), 205–237. <https://doi.org/10.1111/j.1745-3984.2004.tb01163.x>
- Leong, J. P. (2020). PISA 2015 collaborative problem solving assessment : cross-group measurement invariance in Singapore. (Master dissertation). Retrieved from <https://repository.nie.edu.sg/handle/10497/22344?mode=full>
- Li, C.-H., & Liu, Z.-Y. (2017). Collaborative Problem-Solving Behavior of 15-Year-Old Taiwanese Students in Science Education. *EURASIA Journal of Mathematics Science and Technology Education*, 13(10), 6677-6695. <https://doi.org/10.12973/ejmste/78189>
- Lim, H. J. & Lee, S. Y. (2018), Student-, School-, and ICT-Factors Predicting Computer-based Collaborative Problem Solving: Focusing on Analyses of Multi-level Models. *Journal of The Korean Association of Information Education*, 22(4), 457-471. <https://doi.org/10.14352/jkaie.2018.22.4.457>
- Lim, Y. S., & Drasgow, F. (2017): Nonparametric calibration of item-by-attribute matrix in cognitive diagnosis. *Multivariate Behavioral Research*, 52(5), 562-575. <https://doi.org/10.1080/00273171.2017.1341829>
- Lim, H. J. & Lee, S. Y. (2018), Student-, School-, and ICT-Factors Predicting Computer-based Collaborative Problem Solving: Focusing on Analyses of Multi-level Models. *Journal of The Korean Association of Information Education*, 22(4), 457-471. <https://doi.org/10.14352/jkaie.2018.22.4.457>
- Limas, M. C., Mere, J. B. O., Marcos, A. G., de Pison Ascacibar, F. J. M., Espinoza, A. V. P., Elias, F. A., & Ramos, J. M. P. (2015, February). AMORE [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/AMORE/index.html>
- Liu, R., Huggins-Manley, A. C., & Bulut, O. (2018). Retrofitting Diagnostic Classification Models to Responses from IRT-Based Assessment Forms. *Educational and Psychological Measurement*, 78(3), 357-383. <https://doi.org/10.1177/0013164416685599>
- Liu, R., Qian, H., Luo, X., & Woo, A. (2017). Relative diagnostic profile: A subscore reporting framework. *Educational and Psychological Measurement*, 78(6), 1072-1088. <https://doi.org/10.1177/0013164417740170>

- Leong, J. P. (2020). PISA 2015 collaborative problem solving assessment : cross-group measurement invariance in Singapore. (Master dissertation). Retrieved from <https://repository.nie.edu.sg/handle/10497/22344?mode=full>
- Ma, C., de la Torre, J., & Xu, G. (2020). Bridging Parametric and Nonparametric Methods in Cognitive Diagnosis. Retrieved from arXiv:2006.15409
- Ma, W. (2017). *A sequential cognitive diagnosis model for graded response: Model development, q-matrix validation, and model comparison*. New Brunswick: Rutgers The State University of New Jersey.
- Ma, W., de la Torre, J., Sorrel, M. & Jiang, Z. (2020, May). *GDINA* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/GDINA/index.html>
- Madison, M. J., & Bradshaw, L. P. (2015). The effects of Q-matrix design on classification accuracy in the log-linear cognitive diagnosis model. *Educational and Psychological Measurement*, 75(3), 491–511. <https://doi.org/10.1177/0013164414539162>
- Matzavela, V., Chrysafiadi, K., & Alepis, E. (2017). *Questionnaires and artificial neural networks: A literature review on modern techniques in education*. Paper presented at the IEEE Global Engineering Education Conference (EDUCON). Athens, Greece. Retrieved from <https://ieeexplore.ieee.org/document/7943077>
- McCoy, T., & Willse, J. (2014, April). *Accuracy of neural network versus nonparametric approaches in diagnostic classification*. Paper presented at the National Council on Measurement in Education.
- McGlohen, M. (2004). *The application of cognitive diagnosis and computerized adaptive testing to a large-scale assessment*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/>
- Millî Eğitim Bakanlığı, (2018). *MEB 2023 Eğitim Vizyonu*. http://2023vizyonu.meb.gov.tr/doc/2023_EGITIM_VIZYONU.pdf adresinden erişilmiştir.
- Muthén, L. K., & Muthén, B. O. (1998-2011). *Mplus User's Guide*. Sixth Edition. Los Angeles, CA: Muthén & Muthén.

- No Child Left Behind Act of 2001, Pub. L. No. 107-110. Retrieved from <http://thomas.loc.gov>
- Nouri, J., Åkerfeldt, A., Fors, U., & Selander, S. (2017). Assessing collaborative problem solving skills in technology-enhanced learning environments—the PISA framework and modes of communication. *International Journal of Emerging Technologies in Learning*, 12(4), 163-174. <https://doi.org/10.3991/ijet.v12i04.6737>
- Nordin, N. M., & Osman, K. (2018). Students' generated animation: An innovative approach to inculcate collaborative problem solving (CPS) skills in learning physics. *Journal of Education in Science, Environment and Health (JESEH)*, 4(2), 206-226. <https://doi.org/10.21891/jeseh.436758>
- Nouri, J., Åkerfeldt, A., Fors, U., & Selander, S. (2017). Assessing collaborative problem solving skills in technology-enhanced learning environments—the PISA framework and modes of communication. *International Journal of Emerging Technologies in Learning*, 12(4), 163-174. <https://doi.org/10.3991/ijet.v12i04.6737>
- OECD (2013). *PISA 2015 draft collaborative problem-solving framework*. Retrieved from https://www.oecd.org/callsfortenders/Annex%20ID_PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf
- OECD (2017a). *PISA 2015 collaborative problem-solving framework*. Retrieved from <https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>
- OECD (2017b). *PISA 2015 results (Volume V): Collaborative problem solving*. Retrieved from [https://www.oecd.org/publications/pisa-2015-results-volume-v-9789264285521-en.htm#:~:text=PISA%202015%20Results%20\(Volume%20V\)%3A%20Collaborative%20Problem%20Solving%2C%20is,try%20to%20solve%20a%20problem](https://www.oecd.org/publications/pisa-2015-results-volume-v-9789264285521-en.htm#:~:text=PISA%202015%20Results%20(Volume%20V)%3A%20Collaborative%20Problem%20Solving%2C%20is,try%20to%20solve%20a%20problem)
- Özkan, E., & Öner, D. (2019). Investigation of the development of the Van Hiele levels of geometric thinking in a computer-supported collaborative learning (CSCL) environment. *Mersin University Journal of the Faculty of Education*, 15(2): 473-490. <https://doi.org/10.17860/mersinefd.522491>
- Paulsen, J. (2019). *Examining cognitive diagnostic modeling in small sample contexts*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/>

- Riedmiller M. (1994). *Rprop - Description and Implementation Details*. Germany: University of Karlsruhe.
- Riedmiller, M., & Braun, H. (1993). *A direct adaptive method for faster backpropagation learning: The RPROP algorithm*. Proceedings of the IEEE International Conference on Neural Networks (ICNN). Retrieved from <https://ieeexplore.ieee.org/document/298623>
- Robitzsch, A., Kiefer, T., George, A. C., & Uenlue, A. (2019, September). CDM [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/CDM/index.html>
- Rosen, Y., & Tager, M. (2013). *Computer-based assessment of collaborative problem-solving skills: Human-to-agent versus human-to-human approach*. Boston, MA: Pearson Education.
- Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Foundations*. Cambridge, MA: MIT.
- Rupp, A. A. (2013). Clustering and Classification. In T. D. Little (Ed.), *The Oxford Handbook of Quantitative Methods, Volume 2: Statistical Analysis* (pp. 517-550). USA: Oxford University.
- Rupp, A. A., & Templin, J. (2008a). The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Educational and Psychological Measurement*, 68(1), 78-96. <https://doi.org/10.1177/0013164407301545>
- Rupp, A. A., & Templin, J. L. (2008b). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives*, 6(4), 219-262. <https://doi.org/10.1080/15366360802490866>
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement*. New York, NY: The Guilford.
- Rutherford, A. (2011). *ANOVA and ANCOVA: A GLM Approach*. UK: John Wiley & Sons.

- Sarkar, D., Andrews, F., Wright, K., Klepeis, N., & Murrell, P. (2020, April). *lattice* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/lattice/lattice.pdf>
- Schrepp, M. (2005). About the connection between knowledge structures and latent class models. *Methodology*, 1(3), 93–103. <https://doi.org/10.1027/1614-2241.1.3.93>
- Sessoms, J., & Henson, R. A. (2018). Applications of diagnostic classification models: A literature review and critical commentary. *Measurement: Interdisciplinary Research and Perspectives*, 16(1), 1-17. <https://doi.org/10.1080/15366367.2018.1435104>
- Shu, Z., Henson, R., & Willse, J. (2013). Using neural network analysis to define methods of DINA model estimation for small sample sizes. *Journal of Classification*, 30, 173-194. <https://doi.org/10.1007/s00357-013-9134-7>
- Shuying, S. (2016). *Nonparametric diagnostic classification analysis for testlet based tests*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/>
- Sparks, J. N. (1963). Expository notes on the problem of making multiple comparisons in a completely randomized design. *Journal of Experimental Education*, 31(4), 343-349. <https://doi.org/10.1080/00220973.1963.11010788>
- Stadler, M., Herborn, K., Mustafić, M., & Greiff, S. (2018). Computer-Based Collaborative Problem Solving in PISA 2015 and the Role of Personality. *Journal of Intelligence*, 7(15), 1-15. <https://doi.org/10.3390/jintelligence7030015>
- Stadler, M., Herborn, K., Mustafić, M., & Greiff, S. (2020). The assessment of collaborative problem solving in PISA 2015: An investigation of the validity of the PISA 2015 CPS tasks. *Computers & Education*, 157, 1-11. <https://doi.org/10.1016/j.compedu.2020.103964>
- Stekhoven, D. J. (2016, August). *missForest* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/missForest/missForest.pdf>
- Sünbül, S. (2018). The impact of different missing data handling methods on DINA model. *International Journal of Evaluation and Research in Education (IJERE)*, 7(1), 77-86.

- Sünbül, S. (2013). *Bilişsel tanı modellerinde parametre kestirimini ve sınıflama tutarlılığını etkileyen faktörlerin incelenmesi*. Doktora Tezi, Mersin Üniversitesi Eğitim Bilimleri Enstitüsü, Mersin.
- Şata, M., Çorbacı, E. C., & Karakaya, İ. (2018). PISA 2015 Uygulamasında İşbirlikli Problem Çözme Alt Testlerinden Xandar (Zamara) Ünitesinin Cinsiyete Göre Ölçme Değişmezliğinin Test Edilmesi. 27th International Congress on Educational Sciences kongresinde sunulmuştur, Antalya, Türkiye.
- Şata, M., Karaduman, B., & Karakaya, İ. (2018). PISA 2015 Sonuçlarına Göre İşbirlikli Problem Çözme Performansının Öğrenci Özellikleri Açısından İncelenmesi: Türkiye Örnekleme. II. UBEK kongresinde sunulmuştur, Aksaray, Türkiye.
- Tan, W.Y. (1982). Sampling distributions and robustness of t, F and variance ratio in two samples and ANOVA models with respect to departure from normality. *Communications in Statistics: Theory and Methods*, 11(7), 731-750. <https://doi.org/10.1080/03610928208828267>
- Tatsuoka, K. (1990). Architecture of Knowledge Structures and Cognitive Diagnosis: A Statistical Pattern Classification Approach. In P. Nichols, S. Chipman, & R. Brennan (Eds.), *Cognitively Diagnostic Assessments* (pp. 327–359). Hillsdale, NJ: Erlbaum.
- Tatsuoka, K. K. (2009). *Cognitive assessment: An introduction to the rule-space method*. Florence, KY: Routledge.
- Templin, J., & Bradshaw, L. (2014). Hierarchical diagnostic classification models: A family of models for estimating and testing attribute hierarchies. *Psychometrika*, 79(2), 317–339. <https://doi.org/10.1007/s11336-013-9362-0>
- Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, 11(3), 287-305. <https://doi.org/10.1037/1082-989X.11.3.287>
- Toker, T. (2010). *Cognitive diagnostic assessment of TIMSS-2007 mathematics achievement items for 8th graders in Turkey*. (Yüksek lisans tezi). <https://tez.yok.gov.tr> sayfasından erişilmiştir.

- Tomarken, A.J. & Serlin, R.C. (1986). Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. *Psychological Bulletin*, 99(1), 90–99. <https://doi.org/10.1037/0033-2909.99.1.90>
- Xu, X., & von Davier, M. (2006). *Cognitive diagnosis for NAEP proficiency data*. Princeton, NJ: ETS.
- van Buuren, S., Groothuis-Oudshoorn, K., Vink, G., Schouten, R., Robitzsch, A., Doove, L., ... Gray, B. (2020, May). *Mice* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/mice/mice.pdf>
- von Davier, D., & Yamamoto, (2008). *Reporting test outcomes with models for cognitive diagnosis*. Princeton, NJ: ETS. Retrieved from <https://doi.org/10.1002/j.2333-8504.2006.tb02034.x>.
- Wang, M.-J. (2015). Computer-based Assessment of Collaborative Problem-Solving for intermediate Elementary Students with One on One, Human-to-Agent Approach. (Master dissertation). Retrieved from <http://ntcuir.ntcu.edu.tw/handle/987654321/385>
- Wang, C., & Gierl, M. J. (2011). Using the attribute hierarchy method to make diagnostic inferences about examinees' cognitive skills in critical reading. *Journal of Educational Measurement*, 48(2), 165–187.
- Wang, M.-J. (2015). Computer-based Assessment of Collaborative Problem-Solving for intermediate Elementary Students with One on One, Human-to-Agent Approach. (Master dissertation). Retrieved from <http://ntcuir.ntcu.edu.tw/handle/987654321/385>
- Wang, S., & Douglas, J. (2015). Consistency of nonparametric classification in cognitive diagnosis. *Psychometrika*, 80(1), 85–100. <https://doi.org/10.1007/S11336-013-9372-Y>
- Wang, W., Song, L., Chen, P., Meng, Y., & Ding, S. (2015). Attribute-level and pattern-level classification consistency and accuracy indices for diagnostic assessment. *Journal of Educational Measurement*, 52(4), 457–476. <https://doi.org/10.1111/jedm.12096>
- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., ... Dunnington, D. (2020, June). *ggplot2* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/ggplot2/index.html>

- Willse, J., Henson, R., & Templin, J. (2007, April). *Using sum scores or IRT in Place of cognitive diagnosis models: Can existing or more familiar models do the job?* Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, Illinois.
- Yavuz, E., & Atar, H. Y. (2020). An examination of Turkish students' PISA 2015 collaborative problem-solving competencies. *International Journal of Assessment Tools in Education*, 7(4), 588–606. <https://doi.org/10.21449/ijate.682103>
- Zheng, Y., Chiu, C.-Y., & Douglas, J. A. (2019, November). *NPCD* [software package in R]. Retrieved from <https://cran.r-project.org/web/packages/NPCD/NPCD.pdf>



EKLER



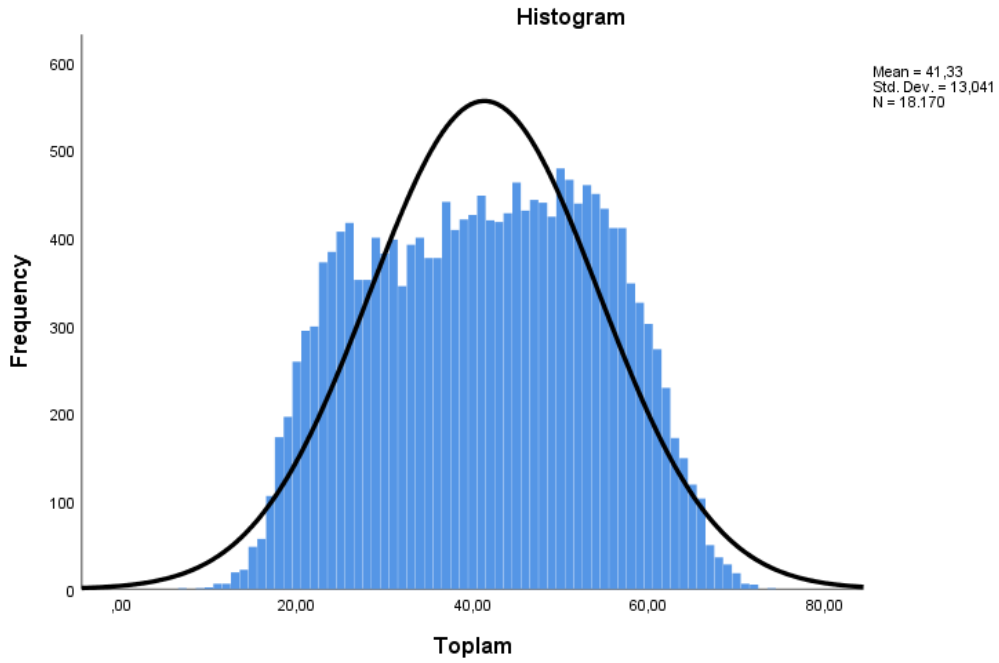
EK-1. Gerçek Veri Setlerine Ait Betimsel Bulgular

Tablo 34

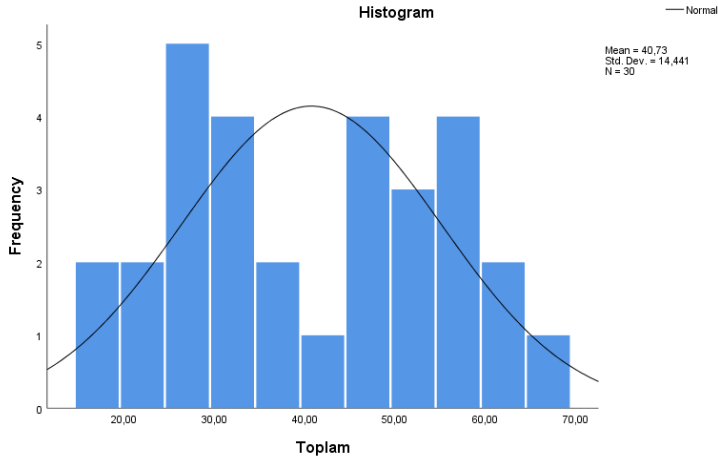
Gerçek Veri Setlerine Ait Betimsel Bulgular

Ülkeler	<i>f</i>	500 Kişilik	100 Kişilik	30 Kişilik
Avustralya	796	18	1	
Avusturya	414	15	3	1
Belçika	520	17	2	1
Bulgaristan	305	12	1	
Kanada	1316	34	6	1
Şili	368	8	1	
Çin Taipei	551	9	1	
Kolombiya	588	11	3	1
Kosta Rika	288	9	4	2
Danimarka	405	9	2	2
Estonya	364	11	1	1
Finlandiya	371	15	4	2
Fransa	350	8	1	
Almanya	366	10	4	2
Yunanistan	381	13	3	
Macaristan	309	8	1	1
İzlanda	195	7	2	
İsrail	367	15	4	
İtalya	78	2	2	
Letonya	325	14	5	1
Litvanya	70	1		
Lüksemburg	294	12	4	2
Makao	38			
Meksika	320	5	2	
Karadağ	331	11	3	1
Hollanda	362	6	2	1
Yeni Zelanda	241	3	1	1
Norveç	322	4		
Peru	228	5	1	
Portekiz	369	9	1	1

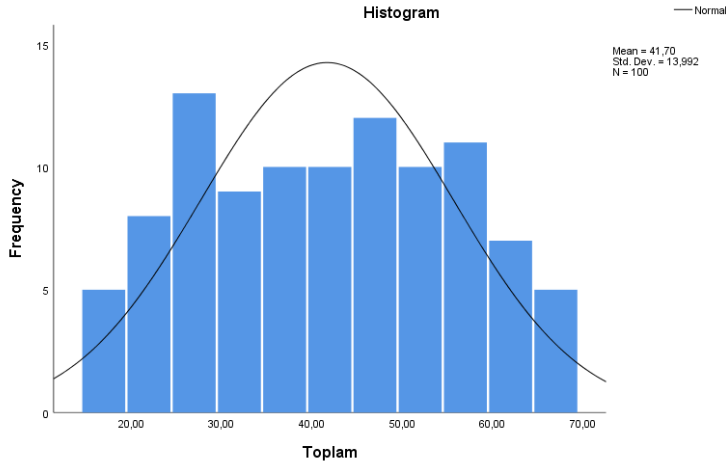
Rusya Federasyonu	338	17	1	
Singapur	412	17	6	1
Slovak cumhuriyeti	325	7	2	
İspanya	364	11	2	2
İsveç	274	9	2	
Birleşik Arap Emirlikleri	965	30	5	2
Birleşik Krallık	949	23	3	
Amerika Birleşik Devletleri	331	12		
Uruguay	318	9	2	
Çin	671	21	5	3
İspanya (Bölgeler)	1781	38	6	1
ABD (Massachusetts)	106	4	1	
ABD (Kuzey Carolina)	104	1		
Toplam	18170	500	100	30



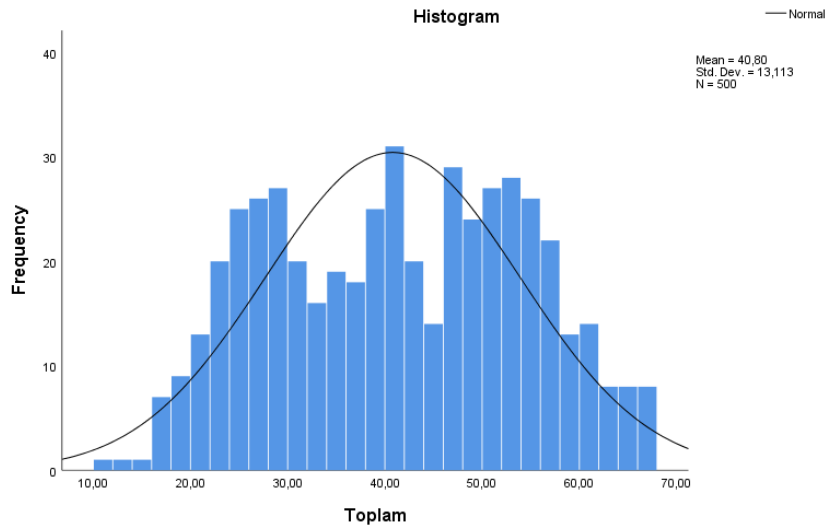
Şekil 16. Çalışma evreninin dağılımı



Şekil 17. Örneklem büyüklüğü 30 olan grubun dağılımı



Şekil 18. Örneklem büyüklüğü 100 olan grubun dağılımı



Şekil 19. Örneklem büyüklüğü 500 olan grubun dağılımı

EK 2. Simülasyon Veri Setlerinin Üretilmesinde Kullanılan Q Matrisleri

Tablo 35

Simülasyon Veri Setlerinin Üretilmesinde Kullanılan Q Matrisleri

3 nitelikli Q matrisi			5 nitelikli Q matrisi					7 nitelikli Q matrisi						
1	0	0	1	0	0	0	0	1	0	0	0	0	0	0
0	1	0	0	1	0	0	0	0	1	0	0	0	0	0
0	0	1	0	0	1	0	0	0	0	1	0	0	0	0
1	0	0	0	0	0	1	0	0	0	0	1	0	0	0
0	1	0	0	0	0	0	1	0	0	0	0	1	0	0
0	0	1	0	1	0	0	0	0	0	0	0	0	1	0
0	1	0	0	0	0	0	1	0	0	0	0	0	0	1
1	1	0	1	1	0	0	0	1	1	0	0	0	0	0
0	1	1	0	0	1	1	0	0	1	1	0	0	0	0
1	0	1	1	1	0	0	0	0	0	1	1	0	0	0
0	1	1	0	1	1	0	0	0	0	0	1	1	0	0
1	0	1	0	0	1	1	0	0	0	0	0	1	1	0
1	1	0	0	0	0	1	1	0	0	0	0	0	1	1
1	1	1	1	0	0	1	1	1	0	0	1	0	0	1
1	1	1	1	0	1	0	1	1	1	1	0	0	0	0

EK 3. Simülasyon Veri Setlerinin Yapısı ve Koşulların Belirlenmesinde Dikkate Alınan Çalışmalar

Tablo 36

Simülasyon Veri Setlerinin Yapısı ve Koşulların Belirlenmesinde Dikkate Alınan Çalışmalar

Çalışmalar	s parametres	g parametres	Nitelik dağılımı	Madde sayısı	Nitelik sayısı	Örneklem büyüklüğü
Shu, Henson ve Willse, 2013	U[0,1; 0,2] U[0,2; 0,40]	U[0,1; 0,2] U[0,2; 0,40]	MVN	60	4, 6	20; 50, 150, 200, 500, 1000
Cui, Gierl ve Guo, 2016	U[0; 0,20] U[0; 0,40]	U[0; 0,20] U[0; 0,40]	Tekdüze	20, 40	3, 6	500, 1000
Guo, Cutumisu ve Cui, 2017	U[0; 0,10]	U[0; 0,10]	*	28	7	5000
Chiu ve Douglas, 2013	U[0; 0,1] U[0; 0,3] U[0; 0,4] U[0; 0,5]	U[0; 0,1] U[0; 0,3] U[0; 0,4] U[0; 0,5]	Tekdüze ve MVN	20, 40	3, 4, 8	1000
Wang ve Douglas, 2015	U[0; 0,1] U[0; 0,3] U[0; 0,5]	U[0; 0,1] U[0; 0,3] U[0; 0,5]	Tekdüze ve MVN	20, 40	3	1000
Akbay, 2016	U[0; 0,15] U[0; 0,25]	U[0; 0,15] U[0; 0,25]	Tekdüze	20	6	250, 500, 1000
Wang, 2016	U[0; 0,1] U[0; 0,3] U[0; 0,5]	U[0; 0,1] U[0; 0,3] U[0; 0,5]	Tekdüze ve MVN	20, 40	3	1000
Lim & Drasgow, 2017	U[0; 0,1] U[0; 0,3] U[0; 0,5] U [0,6; 1]	U[0; 0,1] U[0; 0,3] U[0; 0,5] U [0,6; 1]	Tekdüze ve MVN	20, 40 60	3, 5, 7	250, 1000, 5000
McCoy ve Willse, 2014	U[0; 0,10] U[0,15;0,25] U[0,35;0,45]	U[0; 0,10] U[0,15;0,25] U[0,35;0,45]	MVN	20, 50	4, 8	20, 50, 100, 500, 2000
Paulsen, 2019	U[0; 0,15] U[0,25;0,40]	U[0; 0,15] U[0,25;0,40]	MVN	10, 20, 50	4, 6	25, 50, 150, 1000
Chiu, Sun ve Bian, 2018	U[0; 0,1] U[0; 0,2] U[0; 0,3]	U[0; 0,1] U[0; 0,2] U[0; 0,3]	Tekdüze ve MVN	30, 50	3, 5	30, 50, 100 ve 500
Shuying,	U [0; 0,2]	U [0; 0,2]	MVN	24	4	50, 100,

2016						500, 1000, 10000
Liu & Cheng, 2018	U[0; 0,05] U[0; 0,1] U[0; 0,2]	U[0; 0,05] U[0; 0,1] U[0; 0,2]	MVN	30	6	10, 20, 30, 50, 100 ve 150
Chiu, 2008; Chiu, Douglas ve Li, 2009	U[0; 0,15] U[0; 0,30]	U[0; 0,15] U[0; 0,30]	MVN	20, 40 80	3, 4	100, 500
Chiu and Kohn, 2015b	U[0; ,15] U[0; ,30]	U[0; ,15] U[0; ,30]	Tekdüze ve MVN	20, 40, 80	3, 4	100, 500
Chiu and Kohn, 2016	U[0,80;1] U[0; 0,20] U[0,35; 0,8] N[0; 1] logN[1; 0,30]	U[0,80;1] U[0; 0,20] U[0,35; 0,8] N[0; 1] logN[1; 0,30]	MVN	20,40, 80,100, 120	3, 4	500
Sünbül, 2013	U[0,05; 0,1] U[0,2; 0,30]	U[0,05; 0,1] U[0,2; 0,30]	MVN	15, 30, 45	3, 4	200, 500, 1000, 5000

* Simülasyon verisinin dağılımı hakkında herhangi bir bilgi verilmemiş.

MVN: çok değişkenli normal (multivariate normal) dağılım

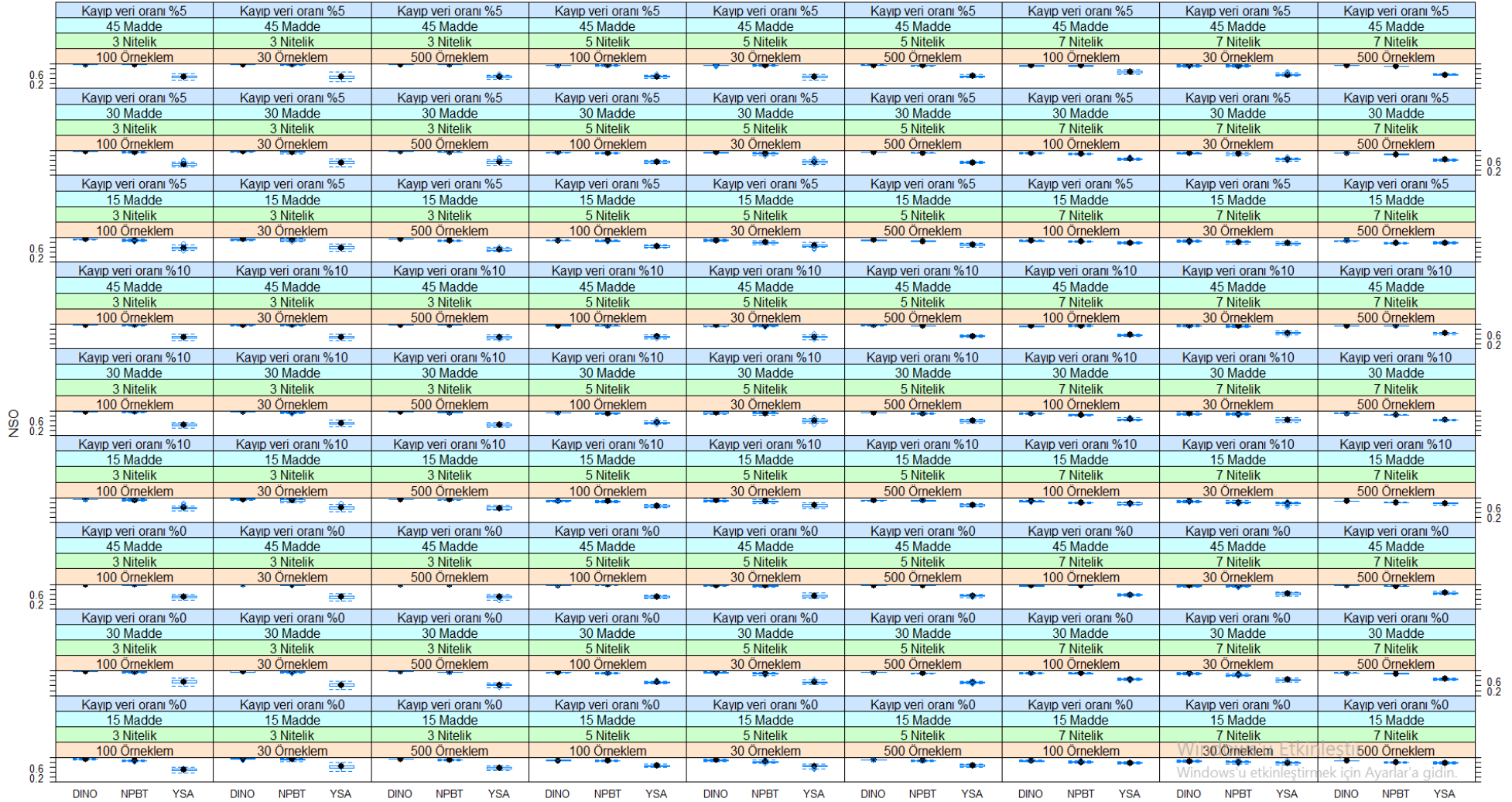
Tablo 37

Kayıp Veri Türü, Miktarı ve Kayıp Veri İle Baş Etme Yollarının Belirlenmesinde Dikkate Alınan Çalışmalar

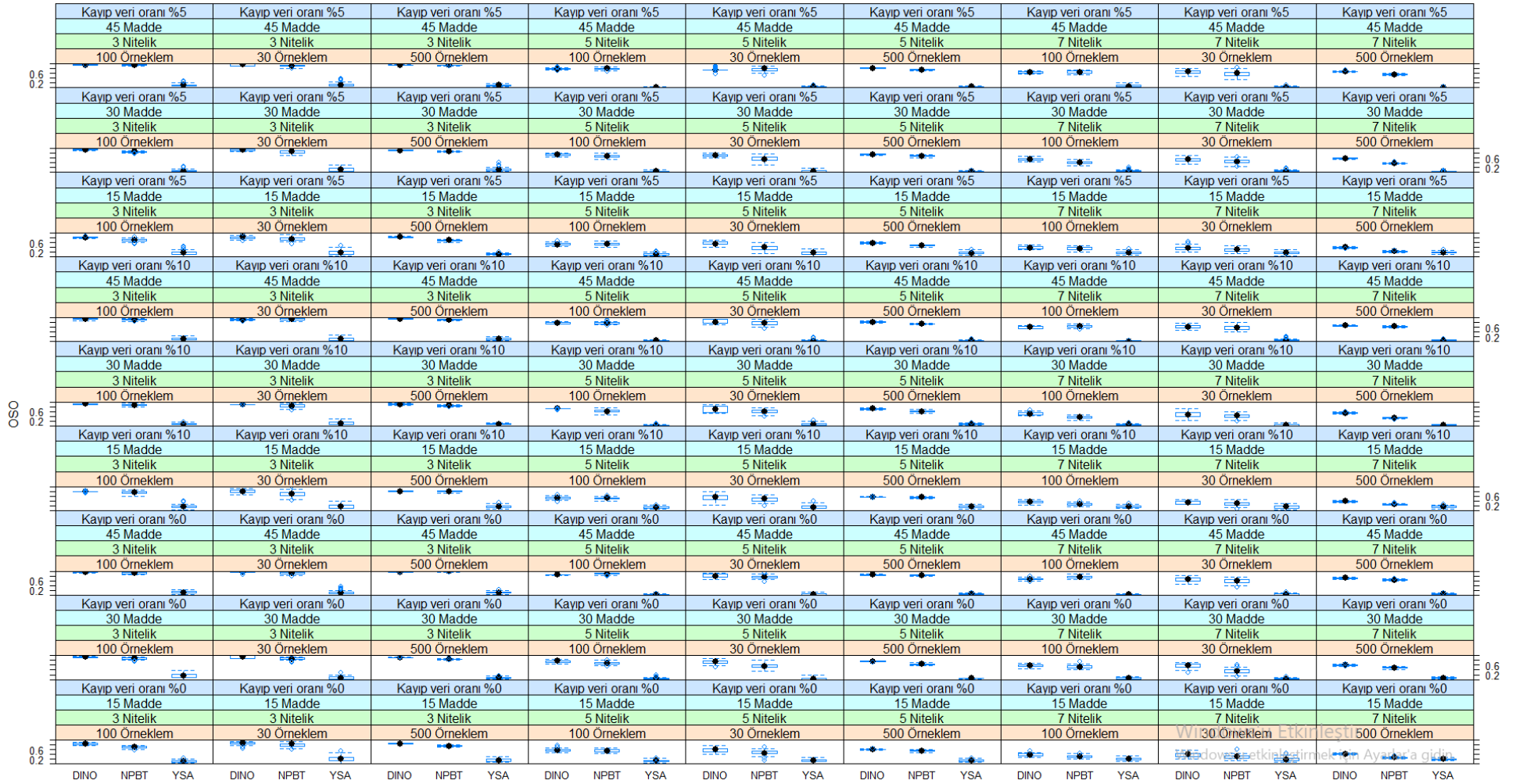
Çalışmalar	Kayıp veri türü	Kayıp veri oranı (%)	Kayıp veri ile baş etme yolu
Xu and von Davier, 2006	Rastgele (missing at random, MAR)	0,10, 25 ve 50	*
Ayers, Nugent, and Dean, 2009	Tamamıyla rastgele (completely at random, MCAR)	0, 10, 20, 30 ve 40	Ortalama ve medyan atama
Başokçu, Kalkan, Öğretmen, 2016	Rastgele	0, 5, 10 ve 15	MCMC, EM ve regresyon
Dai, 2017	Rastgele, Rastgele olmayan (missing not at random, MNAR) ve Karışık desen (the MIXED mechanism)	0, 5, 10, 20 ve 30	Kayıp veriyi yanlış cevap kabul etme, kişi ortalaması atama, iki yönlü atama ve EM
Sünbül, 2018	Rastgele ve Tamamıyla rastgele	0, 5 10 ve 15	Kayıp veriyi yanlış cevap kabul etme, kişi ortalaması atama, iki yönlü atama ve EM
Kalkan, Başokçu, 2019	Rastgele	0, 5 10 ve 15	Çoklu atama yöntemi (multiple imputation, MI)

* Kayıp veri ile baş etme yolu hakkında herhangi bir bilgi verilmemiş.

EK-4. Değişen Koşullar Altında POBT, YSA ve DINO Modelinin NSO ve ÖSO'ları



Şekil 20. POBT, YSA ve DINO modelin değişen koşullar altında NSO değerleri



Şekil 21. POBT, YSA ve DINO modelin değişen koşullar altında ÖSO değerleri

EK-5. Değişen koşullarda POBT'ye ait NSO ve ÖSO'lar

Tablo 38

Değişen Koşullarda POBT'ye Ait NSO ve ÖSO'lar

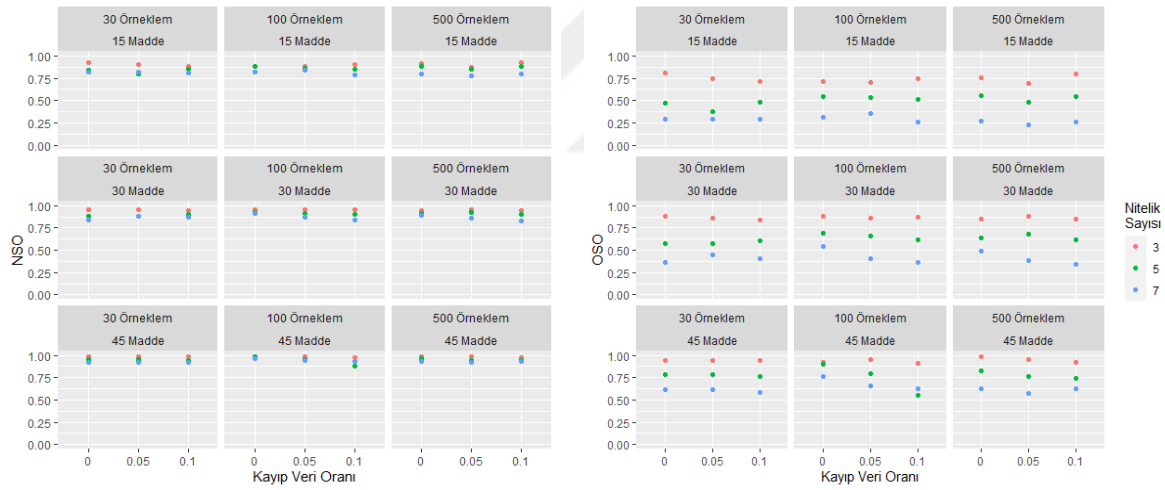
Ni	N	M	%	NSO _{Ni1}	NSO _{Ni2}	NSO _{Ni3}	NSO _{Ni4}	NSO _{Ni5}	NSO _{Ni6}	NSO _{Ni7}	NSO	ÖSO
3	30	15	0	0,889	0,959	0,942					0,930	0,810
				0,064	0,036	0,045					0,028	0,071
				0,05	0,853	0,953	0,903				0,903	0,751
			0,10	0,063	0,039	0,055					0,029	0,070
				0,845	0,951	0,852					0,883	0,707
				0,068	0,041	0,073					0,038	0,094
		30	0	0,964	0,975	0,929					0,956	0,879
				0,029	0,034	0,047					0,022	0,058
				0,05	0,918	0,967	0,961				0,949	0,862
			0,10	0,052	0,034	0,036					0,025	0,067
				0,935	0,948	0,939					0,941	0,840
				0,045	0,041	0,047					0,024	0,060
	45	0	0	0,968	0,994	0,974					0,978	0,938
				0,032	0,015	0,029					0,016	0,046
				0,05	0,964	0,994	0,978				0,979	0,939
			0,10	0,036	0,013	0,029					0,014	0,042
				0,974	0,986	0,977					0,979	0,943
				0,026	0,021	0,027					0,015	0,041
		15	0	0,878	0,872	0,880					0,877	0,708
				0,037	0,035	0,045					0,023	0,050
				0,05	0,888	0,863	0,875				0,875	0,699
			0,10	0,034	0,040	0,033					0,026	0,055
				0,928	0,931	0,848					0,902	0,752
				0,023	0,027	0,040					0,021	0,051
	100	30	0	0,924	0,961	0,985					0,957	0,883
				0,029	0,021	0,014					0,015	0,038
				0,05	0,965	0,965	0,918				0,949	0,861
			0,10	0,020	0,020	0,026					0,011	0,030
				0,931	0,984	0,947					0,954	0,873
				0,026	0,012	0,023					0,012	0,033
		45	0	0,952	0,994	0,969					0,971	0,919
				0,020	0,008	0,018					0,010	0,026
				0,05	0,983	0,989	0,975				0,983	0,950
			0,10	0,014	0,010	0,015					0,007	0,020

			0,10	0,951	0,985	0,968			0,968	0,910
				0,021	0,013	0,017			0,011	0,029
500	15	0	0,879	0,930	0,912				0,907	0,761
			0,015	0,012	0,013				0,008	0,021
		0,05	0,847	0,883	0,876				0,869	0,692
			0,029	0,014	0,015				0,012	0,027
		0,10	0,945	0,923	0,902				0,924	0,801
			0,011	0,013	0,014				0,008	0,020
	30	0	0,951	0,930	0,953				0,945	0,854
			0,010	0,011	0,013				0,006	0,017
		0,05	0,946	0,967	0,950				0,954	0,875
			0,013	0,007	0,010				0,006	0,015
		0,10	0,955	0,966	0,912				0,944	0,849
			0,011	0,008	0,019				0,008	0,021
	45	0	0,984	0,998	0,995				0,992	0,977
			0,006	0,002	0,003				0,002	0,007
		0,05	0,976	0,996	0,973				0,982	0,947
			0,007	0,003	0,007				0,003	0,009
		0,10	0,958	0,989	0,963				0,970	0,915
			0,010	0,005	0,008				0,005	0,013
5	30	15	0	0,847	0,861	0,786	0,838	0,880	0,842	0,471
				0,063	0,069	0,073	0,065	0,068	0,033	0,091
		0,05	0,695	0,836	0,851	0,759	0,841		0,796	0,380
			0,078	0,065	0,061	0,066	0,069		0,036	0,102
		0,10	0,812	0,863	0,867	0,913	0,781		0,847	0,477
			0,068	0,065	0,064	0,051	0,077		0,032	0,093
	30	0	0,846	0,895	0,850	0,886	0,937		0,883	0,573
			0,064	0,055	0,064	0,070	0,046		0,028	0,087
		0,05	0,928	0,859	0,839	0,841	0,958		0,885	0,572
			0,052	0,071	0,070	0,068	0,038		0,031	0,101
		0,10	0,933	0,908	0,852	0,856	0,944		0,899	0,608
			0,043	0,052	0,061	0,065	0,043		0,025	0,087
	45	0	0,962	0,945	0,965	0,952	0,908		0,946	0,775
			0,035	0,049	0,030	0,036	0,055		0,017	0,067
		0,05	0,948	0,962	0,971	0,898	0,964		0,948	0,783
			0,041	0,038	0,030	0,056	0,031		0,020	0,076
		0,10	0,938	0,956	0,921	0,934	0,970		0,944	0,765
			0,045	0,043	0,053	0,052	0,030		0,022	0,086
100	15	0	0,923	0,807	0,882	0,806	0,987		0,881	0,554

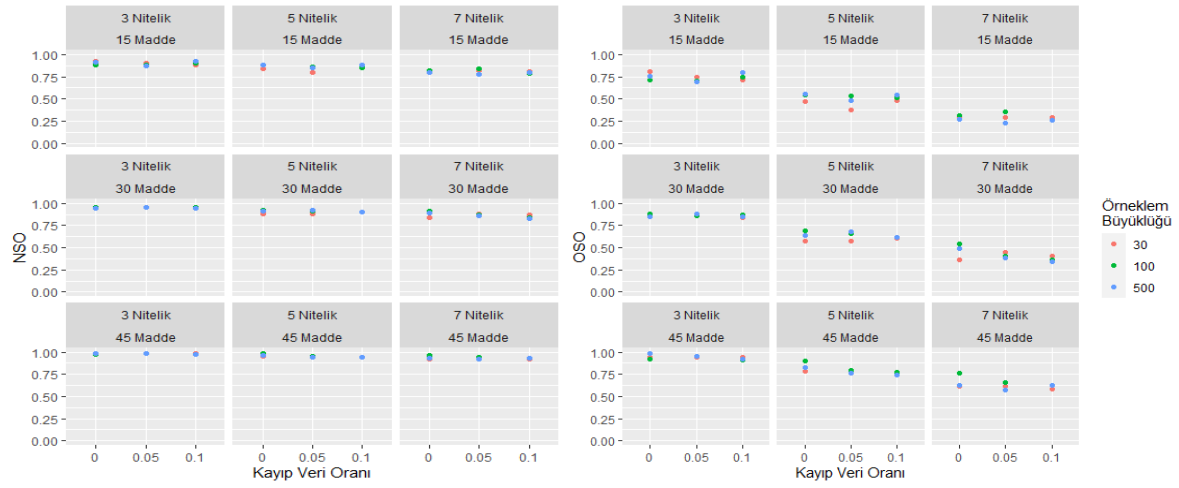
				0,025	0,040	0,033	0,040	0,010					0,016	0,051	
				0,05	0,887	0,882	0,825	0,815	0,900					0,862	0,526
					0,029	0,030	0,037	0,040	0,034					0,017	0,052
				0,10	0,772	0,895	0,784	0,916	0,894					0,853	0,510
					0,041	0,031	0,043	0,028	0,037					0,017	0,047
		30	0		0,862	0,969	0,892	0,941	0,944					0,922	0,686
					0,045	0,017	0,035	0,022	0,026					0,015	0,055
				0,05	0,856	0,957	0,868	0,930	0,939					0,910	0,665
					0,035	0,021	0,041	0,037	0,026					0,015	0,052
				0,10	0,878	0,928	0,865	0,893	0,940					0,901	0,619
					0,044	0,026	0,036	0,033	0,026					0,017	0,058
		45	0		0,967	0,986	0,985	0,972	0,976					0,977	0,897
					0,019	0,012	0,014	0,018	0,018					0,009	0,037
				0,05	0,943	0,980	0,969	0,895	0,976					0,952	0,794
					0,022	0,014	0,017	0,031	0,017					0,010	0,042
				0,10	0,954	0,958	0,912	0,936	0,956					0,943	0,767
					0,020	0,021	0,028	0,026	0,021					0,011	0,042
500	15	0		0,904	0,928	0,859	0,813	0,887					0,878	0,560	
					0,015	0,011	0,015	0,017	0,032					0,010	0,030
				0,05	0,822	0,903	0,807	0,811	0,904					0,849	0,478
					0,016	0,015	0,016	0,018	0,019					0,008	0,024
				0,10	0,905	0,879	0,797	0,927	0,870					0,876	0,552
					0,014	0,014	0,018	0,012	0,015					0,007	0,021
		30	0		0,926	0,937	0,862	0,882	0,937					0,909	0,643
					0,029	0,012	0,016	0,015	0,014					0,008	0,025
				0,05	0,940	0,944	0,855	0,915	0,930					0,917	0,684
					0,023	0,010	0,016	0,012	0,011					0,008	0,026
				0,10	0,852	0,914	0,909	0,894	0,923					0,898	0,621
					0,026	0,013	0,015	0,035	0,013					0,010	0,033
		45	0		0,932	0,964	0,939	0,985	0,983					0,961	0,828
					0,011	0,010	0,010	0,005	0,005					0,004	0,016
				0,05	0,907	0,972	0,917	0,941	0,964					0,940	0,757
					0,014	0,008	0,012	0,010	0,009					0,005	0,020
				0,10	0,930	0,936	0,944	0,926	0,946					0,936	0,742
					0,011	0,013	0,009	0,011	0,012					0,005	0,019
7	30	15	0		0,792	0,834	0,826	0,870	0,791	0,795	0,804		0,816	0,285	
					0,075	0,078	0,061	0,051	0,076	0,074	0,073		0,032	0,085	
				0,05	0,828	0,843	0,820	0,807	0,771	0,824	0,832		0,818	0,286	
					0,068	0,069	0,067	0,066	0,080	0,070	0,074		0,027	0,079	

		0,023	0,043	0,012	0,015	0,020	0,024	0,012	0,009	0,029
	0,05	0,793	0,859	0,862	0,943	0,881	0,815	0,854	0,858	0,384
		0,018	0,016	0,018	0,011	0,015	0,017	0,017	0,007	0,021
	0,10	0,790	0,826	0,842	0,818	0,789	0,886	0,873	0,832	0,336
		0,023	0,025	0,027	0,025	0,022	0,026	0,021	0,009	0,024
45	0	0,944	0,955	0,956	0,951	0,935	0,851	0,921	0,930	0,630
		0,010	0,009	0,010	0,010	0,012	0,019	0,012	0,005	0,025
	0,05	0,926	0,934	0,927	0,924	0,958	0,925	0,817	0,916	0,565
		0,011	0,011	0,013	0,012	0,009	0,013	0,015	0,005	0,023
	0,10	0,903	0,893	0,957	0,937	0,965	0,948	0,905	0,930	0,630
		0,014	0,013	0,009	0,012	0,009	0,010	0,013	0,005	0,023

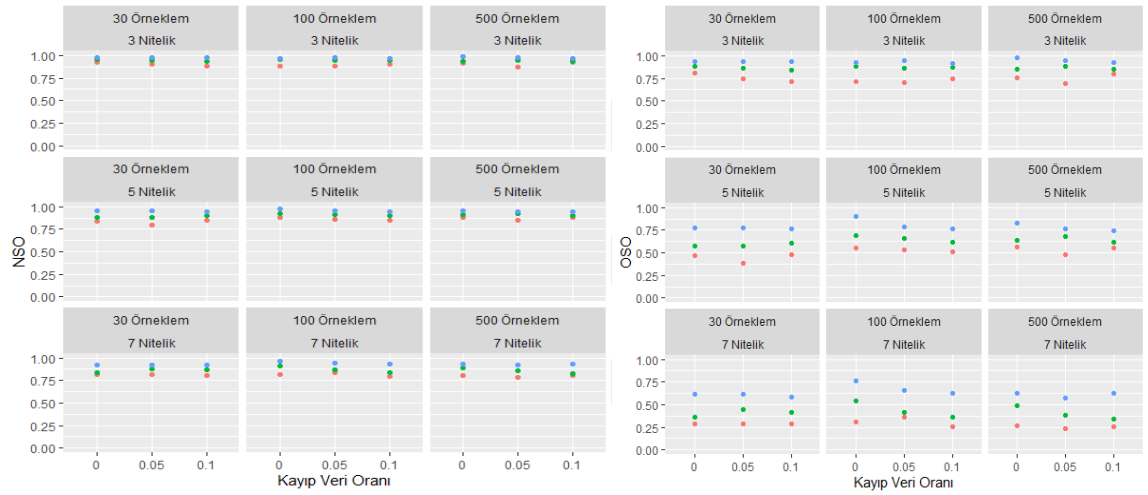
Not: NSO_{Ni1}: Birinci niteliğe; NSO_{Ni2}: İkinci niteliğe; NSO_{Ni3}: Üçüncü niteliğe; NSO_{Ni4}: Dördüncü niteliğe; NSO_{Ni5}: Beşinci niteliğe; NSO_{Ni6}: Altıncı niteliğe; NSO_{Ni7}: Yedinci niteliğe ait sınıflandırma oranı ve NSO: Genel nitelik düzeyi sınıflandırma oranı. Her bir hücrenin altındaki değer o hücredeki sınıflandırma oranlarının standart sapmalarını göstermektedir.



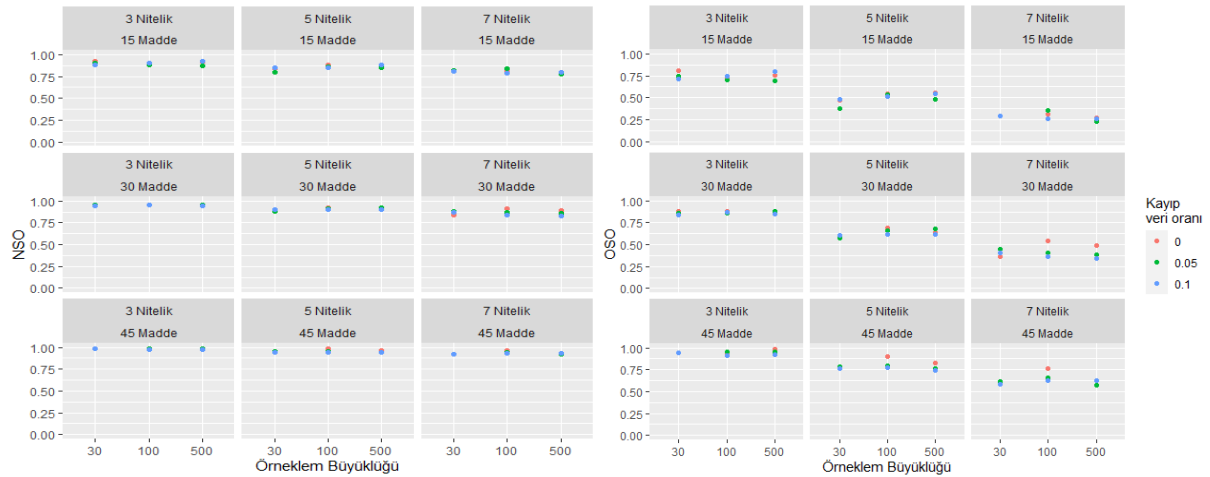
Şekil 22. Nitelik sayılarına (3, 5 ve 7) göre POBT'nin NSO ve ÖSO'ları



Şekil 23. Örneklem büyüklüğüne (30, 100 ve 500) göre POBT'nin NSO ve ÖSO'ları



Şekil 24. Madde sayısına (15, 30 ve 45) göre POBT'nin NSO ve ÖSO'ları



Şekil 25. Kayıp veri oranına (0, 0,05 ve 0,10) göre POBT'nin NSO ve ÖSO'ları

EK-6. Değişen koşullarda YSA'ya ait NSO ve ÖSO'lar

Tablo 39

Değişen Koşullarda YSA'ya Ait NSO ve ÖSO'lar

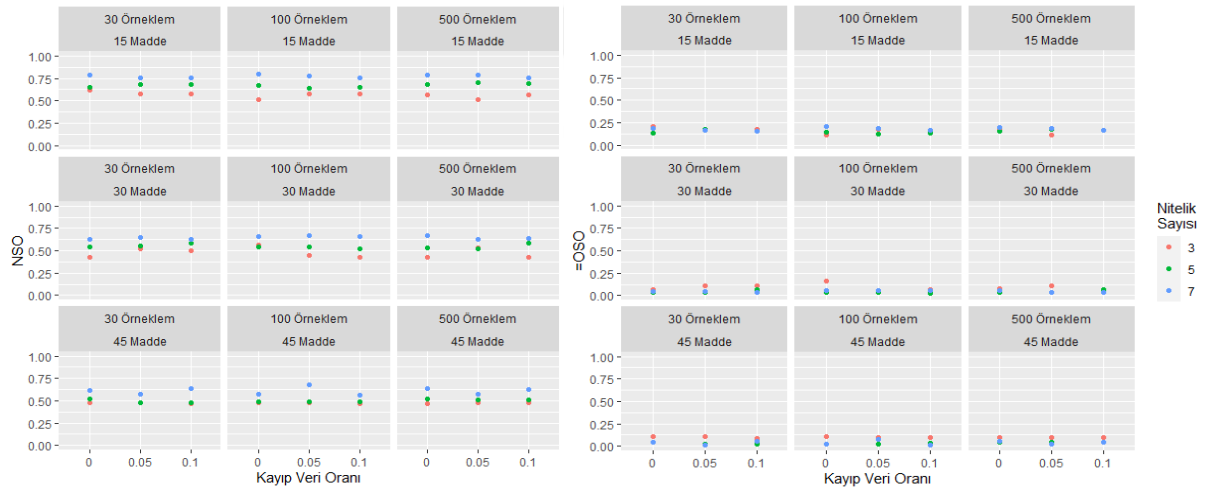
Ni	N	M	%	NSO _{Ni1}	NSO _{Ni2}	NSO _{Ni3}	NSO _{Ni4}	NSO _{Ni5}	NSO _{Ni6}	NSO _{Ni7}	NSO	ÖSO
3	30	15	0	0,608	0,631	0,629					0,622	0,214
				0,109	0,103	0,123					0,072	0,109
				0,05	0,575	0,583	0,583				0,580	0,178
			0,10	0,099	0,111	0,110					0,071	0,094
				0,573	0,592	0,566					0,577	0,167
				0,102	0,114	0,106					0,062	0,088
		30	0	0,355	0,460	0,473					0,429	0,074
				0,113	0,112	0,104					0,070	0,055
				0,05	0,453	0,529	0,557				0,513	0,093
			0,10	0,107	0,112	0,109					0,069	0,067
				0,495	0,522	0,472					0,496	0,111
				0,090	0,115	0,097					0,057	0,065
		45	0	0,415	0,539	0,485					0,480	0,107
				0,123	0,109	0,109					0,065	0,070
				0,05	0,437	0,509	0,482				0,476	0,111
			0,10	0,117	0,120	0,110					0,075	0,072
				0,407	0,524	0,469					0,467	0,092
				0,102	0,117	0,109					0,069	0,062
100	15	15	0	0,510	0,516	0,500					0,508	0,115
				0,081	0,079	0,077					0,049	0,042
				0,05	0,569	0,585	0,583				0,579	0,168
			0,10	0,072	0,075	0,080					0,051	0,068
				0,569	0,582	0,576					0,576	0,160
				0,075	0,083	0,079					0,053	0,067
		30	0	0,490	0,568	0,623					0,560	0,161
				0,086	0,099	0,093					0,065	0,085
				0,05	0,381	0,459	0,502				0,447	0,065
			0,10	0,074	0,074	0,082					0,047	0,039
				0,364	0,465	0,458					0,429	0,073
				0,072	0,085	0,078					0,050	0,038
		45	0	0,436	0,535	0,480					0,484	0,108
				0,074	0,080	0,083					0,048	0,047
				0,05	0,426	0,526	0,481				0,478	0,104
			0,10	0,076	0,077	0,082					0,054	0,046

500	15	0	0,10	0,432	0,515	0,477			0,475	0,099	
				0,081	0,082	0,078			0,052	0,042	
				0,565	0,584	0,576			0,575	0,163	
				0,063	0,067	0,065			0,047	0,057	
			0,05	0,508	0,529	0,498			0,512	0,114	
				0,059	0,066	0,061			0,038	0,032	
		30	0	0,10	0,579	0,580	0,565			0,574	0,162
					0,056	0,060	0,068			0,046	0,054
					0,359	0,465	0,467			0,430	0,076
					0,064	0,067	0,065			0,044	0,028
				0,05	0,453	0,546	0,590			0,530	0,115
					0,070	0,074	0,075			0,054	0,062
	45	0	0,10	0,368	0,471	0,457			0,432	0,073	
				0,064	0,078	0,065			0,047	0,031	
				0,431	0,513	0,474			0,472	0,100	
				0,067	0,075	0,074			0,051	0,036	
			0,05	0,435	0,524	0,487			0,482	0,099	
				0,055	0,071	0,070			0,041	0,031	
		30	0	0,10	0,434	0,527	0,475			0,479	0,101
					0,060	0,074	0,063			0,041	0,035
					0,606	0,602	0,622	0,704	0,735	0,654	0,127
					0,095	0,092	0,105	0,101	0,084	0,047	0,064
				0,05	0,643	0,612	0,692	0,730	0,744	0,684	0,165
					0,106	0,097	0,105	0,098	0,096	0,052	0,077
5	30	15	0,10	0,654	0,602	0,702	0,734	0,712	0,681	0,148	
				0,095	0,099	0,104	0,102	0,096	0,053	0,071	
				0,413	0,542	0,591	0,520	0,612	0,536	0,040	
				0,097	0,089	0,102	0,100	0,091	0,043	0,035	
			0,05	0,429	0,534	0,623	0,541	0,618	0,549	0,044	
				0,082	0,096	0,100	0,102	0,101	0,044	0,041	
	45	0	0,10	0,488	0,587	0,654	0,577	0,646	0,591	0,072	
				0,098	0,110	0,106	0,107	0,100	0,050	0,051	
				0,477	0,554	0,506	0,480	0,598	0,523	0,041	
				0,103	0,091	0,090	0,104	0,105	0,049	0,035	
			0,05	0,428	0,535	0,483	0,424	0,549	0,484	0,024	
				0,099	0,111	0,098	0,096	0,100	0,050	0,027	
100	15	0	0,10	0,433	0,505	0,467	0,441	0,566	0,482	0,022	
				0,100	0,102	0,087	0,099	0,084	0,044	0,029	
				0,633	0,600	0,667	0,708	0,731	0,668	0,136	

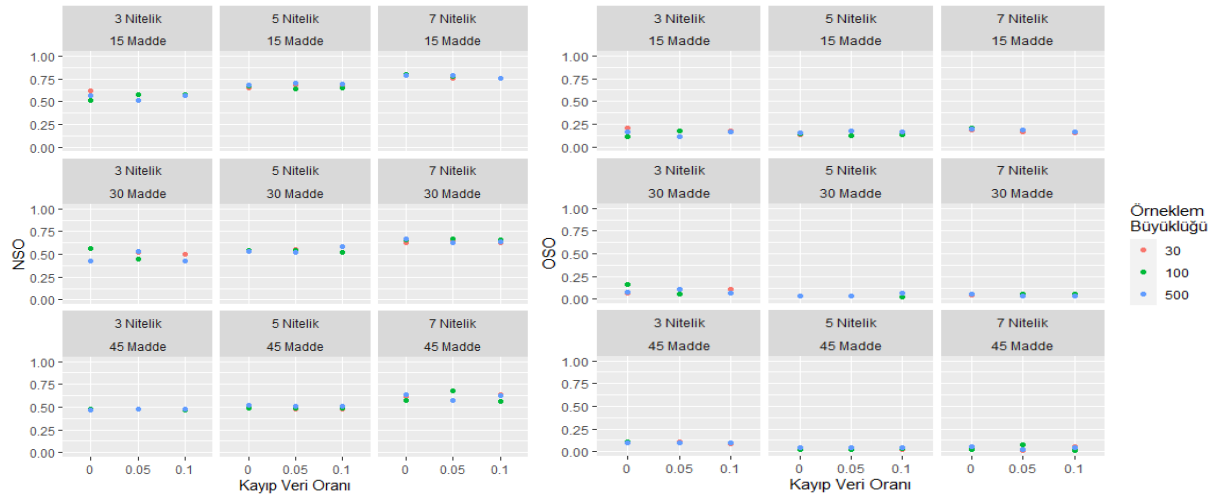
				0,063	0,053	0,070	0,069	0,054			0,029	0,040
			0,05	0,609	0,579	0,639	0,697	0,694			0,644	0,119
				0,063	0,058	0,081	0,061	0,059			0,033	0,041
			0,10	0,618	0,590	0,652	0,696	0,714			0,654	0,127
				0,066	0,060	0,073	0,066	0,055			0,031	0,038
	30	0		0,414	0,533	0,610	0,532	0,610			0,540	0,044
				0,058	0,067	0,059	0,062	0,059			0,027	0,023
		0,05		0,420	0,537	0,616	0,514	0,612			0,540	0,041
				0,052	0,064	0,059	0,066	0,059			0,029	0,024
		0,10		0,415	0,525	0,585	0,507	0,588			0,524	0,033
				0,051	0,072	0,061	0,063	0,069			0,031	0,019
	45	0		0,438	0,518	0,494	0,434	0,555			0,488	0,025
				0,055	0,057	0,058	0,057	0,068			0,028	0,015
		0,05		0,448	0,506	0,485	0,438	0,559			0,487	0,025
				0,054	0,060	0,066	0,060	0,060			0,030	0,014
		0,10		0,443	0,522	0,477	0,429	0,554			0,485	0,028
				0,052	0,075	0,070	0,066	0,072			0,038	0,017
500	15	0		0,647	0,606	0,690	0,728	0,730			0,680	0,152
				0,045	0,039	0,052	0,058	0,036			0,027	0,033
		0,05		0,655	0,632	0,719	0,746	0,745			0,699	0,169
				0,048	0,053	0,069	0,068	0,058			0,036	0,050
		0,10		0,639	0,628	0,710	0,741	0,720			0,688	0,156
				0,051	0,050	0,072	0,060	0,053			0,032	0,042
	30	0		0,410	0,533	0,605	0,519	0,606			0,535	0,041
				0,036	0,050	0,042	0,037	0,047			0,023	0,011
		0,05		0,411	0,514	0,582	0,512	0,602			0,524	0,036
				0,038	0,047	0,049	0,047	0,042			0,022	0,010
		0,10		0,490	0,592	0,658	0,583	0,651			0,595	0,069
				0,040	0,072	0,048	0,043	0,054			0,032	0,027
	45	0		0,471	0,554	0,513	0,472	0,602			0,522	0,041
				0,041	0,049	0,048	0,042	0,049			0,027	0,015
		0,05		0,465	0,542	0,507	0,467	0,588			0,514	0,040
				0,045	0,047	0,049	0,046	0,045			0,028	0,017
		0,10		0,454	0,539	0,500	0,475	0,584			0,511	0,037
				0,045	0,044	0,043	0,048	0,047			0,026	0,016
7	30	15	0	0,712	0,723	0,810	0,832	0,871	0,785	0,782	0,788	0,193
				0,080	0,090	0,077	0,071	0,071	0,075	0,075	0,029	0,081
		0,05		0,698	0,679	0,791	0,785	0,852	0,763	0,760	0,761	0,160
				0,081	0,093	0,088	0,090	0,077	0,092	0,070	0,033	0,067

		0,047	0,054	0,046	0,040	0,049	0,050	0,053	0,024	0,020
	0,05	0,493	0,562	0,699	0,604	0,675	0,675	0,687	0,628	0,038
		0,027	0,034	0,034	0,035	0,037	0,029	0,028	0,015	0,010
	0,10	0,506	0,574	0,711	0,614	0,691	0,677	0,682	0,636	0,041
		0,028	0,030	0,036	0,032	0,029	0,034	0,029	0,014	0,012
45	0	0,596	0,620	0,653	0,594	0,650	0,671	0,715	0,643	0,051
		0,052	0,048	0,051	0,049	0,047	0,046	0,041	0,024	0,019
	0,05	0,518	0,532	0,577	0,514	0,581	0,605	0,633	0,566	0,017
		0,036	0,036	0,031	0,038	0,037	0,037	0,032	0,015	0,007
	0,10	0,587	0,608	0,640	0,596	0,646	0,664	0,692	0,633	0,044
		0,046	0,041	0,044	0,039	0,043	0,042	0,047	0,021	0,016

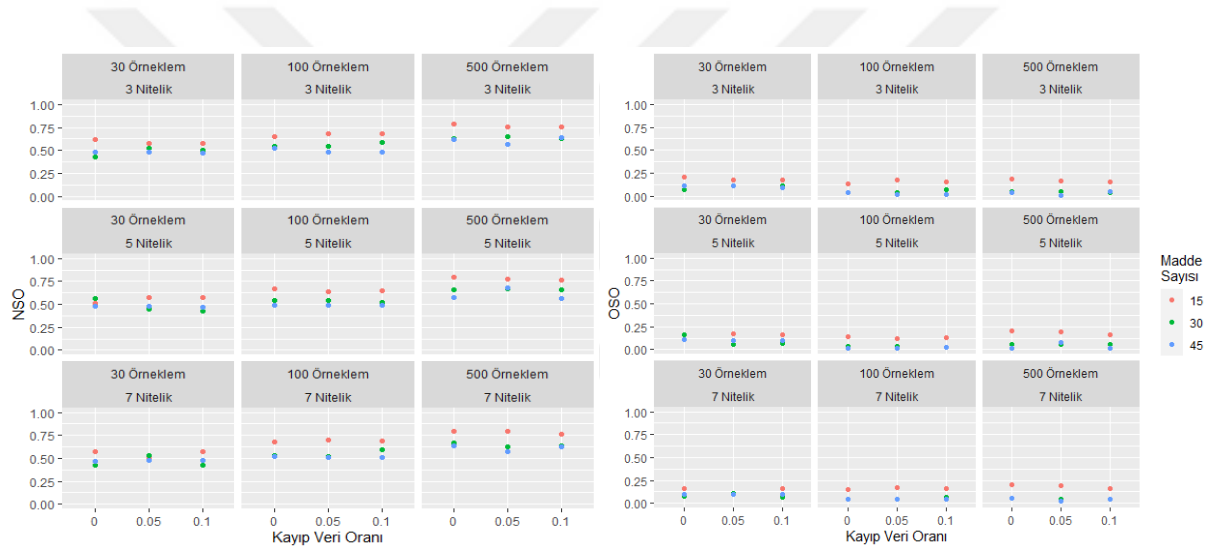
Not: NSO_{Ni1}: Birinci niteliğe; NSO_{Ni2}: İkinci niteliğe; NSO_{Ni3}: Üçüncü niteliğe; NSO_{Ni4}: Dördüncü niteliğe; NSO_{Ni5}: Beşinci niteliğe; NSO_{Ni6}: Altıncı niteliğe; NSO_{Ni7}: Yedinci niteliğe ait sınıflandırma oranı ve NSO: Genel nitelik düzeyi sınıflandırma oranı. Her bir hücrenin altındaki değer o hücredeki sınıflandırma oranlarının standart sapmalarını göstermektedir.



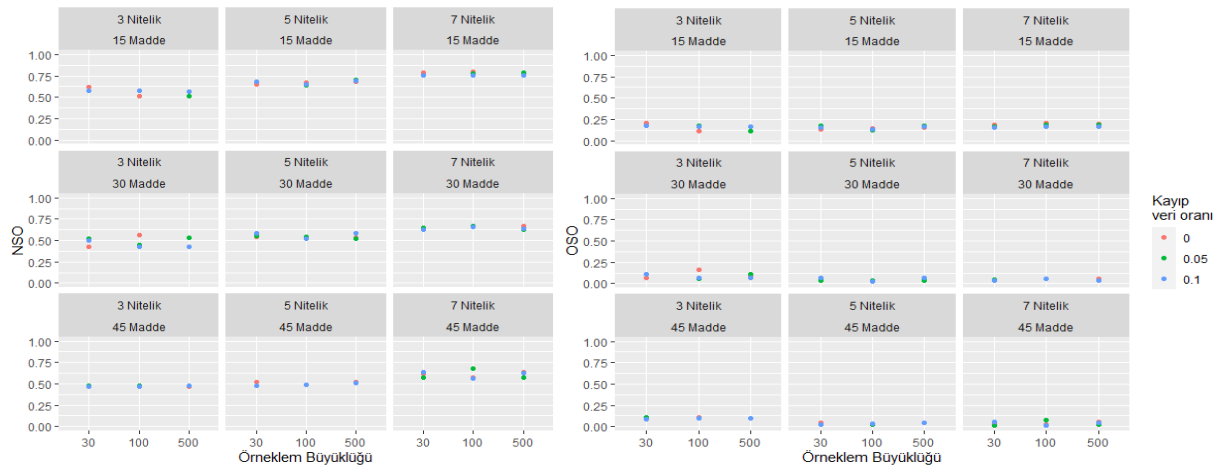
Şekil 26. Nitelik sayılarına (3, 5 ve 7) göre YSA'nın NSO ve ÖSO'ları



Şekil 27. Örneklem büyüklüğüne (30, 100 ve 500) göre YSA'nın NSO ve ÖSO'ları



Şekil 28. Madde sayısına (15, 30 ve 45) göre YSA'nın NSO ve ÖSO'ları



Şekil 29. Kayıp veri oranına (0, 0,05 ve 0,10) göre YSA'nın NSO ve ÖSO'ları

EK-7. Değişen koşullarda DINO modeline ait NSO ve ÖSO'lar

Tablo 40

Değişen Koşullarda DINO Modeline Ait NSO ve ÖSO'lar

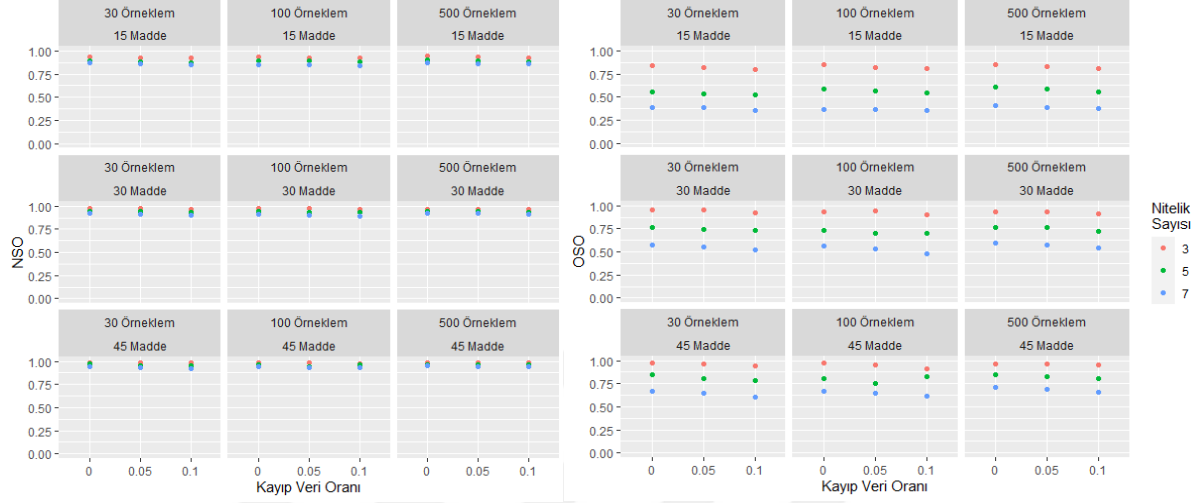
Ni	N	M	%	NSO _{Ni1}	NSO _{Ni2}	NSO _{Ni3}	NSO _{Ni4}	NSO _{Ni5}	NSO _{Ni6}	NSO _{Ni7}	NSO	ÖSO
3	30	15	0	0,944	0,972	0,919					0,945	0,850
				0,047	0,035	0,083					0,030	0,084
				0,912	0,956	0,931					0,933	0,819
			0,05	0,064	0,045	0,044					0,031	0,089
				0,908	0,960	0,916					0,928	0,806
				0,061	0,032	0,056					0,025	0,069
		30	0	0,934	1,000	1,000					0,978	0,934
				0,055	0,000	0,000					0,018	0,055
				0,962	0,992	0,982					0,979	0,938
			0,10	0,026	0,014	0,023					0,012	0,037
				0,923	0,983	0,994					0,967	0,900
				0,025	0,026	0,013					0,007	0,020
		45	0	0,967	1,000	0,999					0,989	0,966
				0,003	0,000	0,011					0,004	0,011
				0,983	1,000	0,968					0,984	0,951
			0,05	0,017	0,000	0,033					0,017	0,050
				0,945	0,977	0,985					0,969	0,911
				0,026	0,020	0,021					0,012	0,034
100	15	15	0	0,932	0,970	0,920					0,941	0,840
				0,018	0,013	0,026					0,012	0,031
				0,919	0,965	0,915					0,933	0,818
			0,05	0,033	0,009	0,022					0,011	0,028
				0,881	0,941	0,956					0,926	0,799
				0,009	0,003	0,014					0,005	0,011
		30	0	0,963	0,986	0,994					0,981	0,948
				0,016	0,008	0,008					0,007	0,021
				0,969	0,993	0,986					0,983	0,953
			0,05	0,017	0,006	0,013					0,008	0,021
				0,949	0,987	0,978					0,972	0,920
				0,014	0,012	0,007					0,005	0,013
		45	0	0,981	0,993	0,992					0,989	0,966
				0,011	0,008	0,007					0,005	0,016
				0,980	0,995	0,987					0,987	0,962
			0,05	0,011	0,006	0,010					0,003	0,008
				0,969	0,983	0,983					0,978	0,937
				0,015	0,014	0,018					0,009	0,025
500	15	15	0	0,940	0,987	0,915					0,947	0,852

				0,010	0,005	0,009			0,005	0,014
			0,05	0,927	0,978	0,914			0,940	0,833
				0,013	0,005	0,008			0,006	0,016
			0,10	0,918	0,970	0,899			0,929	0,805
				0,010	0,010	0,008			0,005	0,013
		30	0	0,948	0,986	0,989			0,974	0,925
				0,006	0,004	0,004			0,002	0,008
			0,05	0,956	0,982	0,984			0,974	0,927
				0,009	0,006	0,004			0,004	0,012
			0,10	0,950	0,976	0,977			0,968	0,910
				0,013	0,006	0,008			0,006	0,016
		45	0	0,974	0,996	0,992			0,988	0,963
				0,005	0,002	0,003			0,002	0,006
			0,05	0,969	0,996	0,990			0,985	0,956
				0,007	0,003	0,004			0,003	0,009
			0,10	0,968	0,993	0,985			0,982	0,948
				0,007	0,004	0,004			0,003	0,008
5	30	15	0	0,812	0,891	0,898	0,905	0,967	0,895	0,588
				0,089	0,058	0,067	0,065	0,031	0,025	0,080
			0,05	0,815	0,888	0,865	0,908	0,950	0,885	0,568
				0,060	0,061	0,067	0,043	0,041	0,025	0,083
			0,10	0,816	0,891	0,873	0,895	0,926	0,880	0,539
				0,115	0,057	0,055	0,055	0,045	0,027	0,112
		30	0	0,849	0,977	0,954	0,924	0,975	0,936	0,728
				0,076	0,027	0,037	0,053	0,024	0,019	0,077
			0,05	0,833	0,972	0,962	0,904	0,971	0,928	0,704
				0,080	0,023	0,042	0,037	0,026	0,014	0,052
			0,10	0,862	0,974	0,915	0,910	0,971	0,926	0,703
				0,073	0,028	0,058	0,063	0,039	0,038	0,125
		45	0	0,906	0,984	0,991	0,914	0,992	0,957	0,803
				0,067	0,023	0,016	0,072	0,014	0,017	0,082
			0,05	0,853	0,970	0,965	0,929	0,996	0,943	0,746
				0,042	0,013	0,021	0,030	0,015	0,011	0,046
			0,10	0,911	0,999	0,974	0,931	0,988	0,961	0,829
				0,077	0,005	0,035	0,048	0,017	0,019	0,075
	100	15	0	0,815	0,885	0,882	0,891	0,955	0,886	0,562
				0,048	0,027	0,031	0,033	0,024	0,014	0,048
			0,05	0,815	0,881	0,861	0,881	0,946	0,877	0,527
				0,043	0,038	0,036	0,030	0,022	0,014	0,047
			0,10	0,803	0,867	0,867	0,881	0,925	0,869	0,521
				0,041	0,040	0,036	0,035	0,034	0,016	0,053
		30	0	0,881	0,977	0,952	0,925	0,978	0,943	0,764

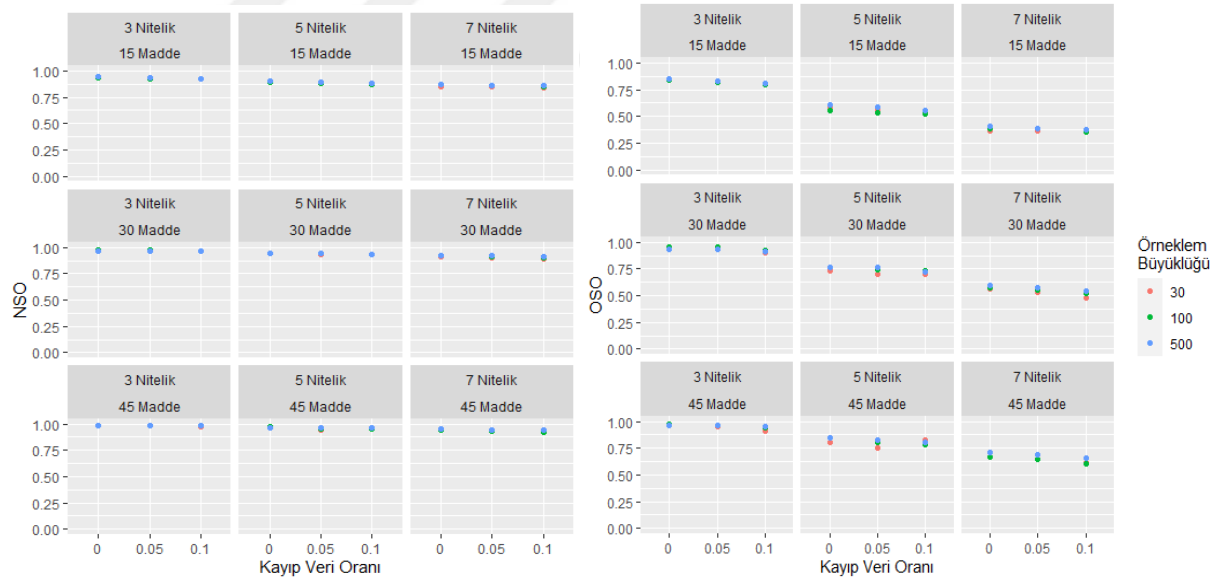
				0,032	0,016	0,023	0,022	0,014		0,011	0,045		
			0,05	0,887	0,972	0,947	0,923	0,964		0,939	0,745		
				0,031	0,021	0,018	0,020	0,020		0,010	0,045		
			0,10	0,897	0,940	0,959	0,853	0,989		0,928	0,728		
				0,013	0,006	0,010	0,015	0,006		0,002	0,012		
		45	0	0,933	0,982	0,965	0,962	0,992		0,967	0,848		
				0,029	0,015	0,021	0,012	0,004		0,003	0,015		
			0,05	0,926	0,973	0,949	0,941	0,982		0,954	0,801		
				0,024	0,012	0,019	0,020	0,015		0,007	0,034		
			0,10	0,912	0,967	0,945	0,936	0,975		0,947	0,777		
				0,025	0,016	0,017	0,020	0,016		0,010	0,035		
		500	15	0	0,834	0,906	0,881	0,922	0,962	0,901	0,607		
				0,011	0,009	0,005	0,007	0,004		0,003	0,009		
			0,05	0,814	0,893	0,881	0,910	0,956		0,891	0,586		
				0,018	0,014	0,012	0,012	0,008		0,007	0,027		
			0,10	0,809	0,891	0,874	0,892	0,944		0,882	0,559		
				0,015	0,010	0,018	0,013	0,007		0,006	0,010		
			30	0	0,875	0,976	0,964	0,930	0,977	0,944	0,765		
				0,013	0,004	0,006	0,009	0,009		0,003	0,015		
			0,05	0,875	0,972	0,956	0,931	0,978		0,942	0,756		
				0,014	0,004	0,012	0,010	0,008		0,004	0,016		
			0,10	0,864	0,966	0,942	0,916	0,969		0,931	0,718		
				0,015	0,008	0,010	0,014	0,008		0,005	0,023		
			45	0	0,940	0,983	0,964	0,950	0,988	0,965	0,847		
				0,009	0,004	0,012	0,007	0,004		0,004	0,017		
			0,05	0,929	0,980	0,960	0,946	0,988		0,961	0,828		
				0,012	0,005	0,009	0,008	0,003		0,003	0,016		
			0,10	0,928	0,976	0,953	0,940	0,980		0,955	0,809		
				0,012	0,007	0,010	0,009	0,006		0,005	0,018		
7	30	15	0	0,782	0,803	0,860	0,874	0,978	0,881	0,803	0,854	0,375	
				0,076	0,072	0,074	0,059	0,026	0,040	0,080	0,025	0,082	
			0,05	0,773	0,788	0,876	0,880	0,948	0,893	0,821	0,854	0,367	
				0,069	0,072	0,063	0,058	0,036	0,061	0,053	0,025	0,081	
			0,10	0,747	0,782	0,860	0,864	0,947	0,867	0,831	0,843	0,347	
				0,064	0,073	0,068	0,050	0,038	0,057	0,067	0,024	0,066	
			30	0	0,820	0,863	0,941	0,902	0,947	0,975	0,916	0,909	0,557
				0,080	0,073	0,039	0,059	0,042	0,028	0,053	0,024	0,091	
			0,05	0,808	0,881	0,941	0,885	0,937	0,960	0,915	0,904	0,525	
				0,068	0,062	0,041	0,065	0,043	0,035	0,061	0,021	0,081	
			0,10	0,810	0,846	0,910	0,888	0,928	0,951	0,907	0,892	0,484	
				0,079	0,068	0,054	0,052	0,041	0,044	0,049	0,026	0,107	
			45	0	0,920	0,914	0,948	0,923	0,948	0,958	0,953	0,938	0,665

			0,049	0,049	0,041	0,050	0,043	0,037	0,047	0,020	0,087
			0,05	0,911	0,922	0,940	0,919	0,943	0,954	0,946	0,648
				0,053	0,049	0,047	0,054	0,051	0,038	0,038	0,019
			0,10	0,885	0,929	0,938	0,915	0,944	0,958	0,931	0,929
				0,056	0,053	0,042	0,052	0,042	0,040	0,051	0,018
100	15	0	0,793	0,811	0,894	0,898	0,955	0,885	0,843	0,868	0,393
			0,044	0,040	0,032	0,028	0,019	0,038	0,038	0,015	0,057
		0,05	0,795	0,820	0,887	0,888	0,944	0,885	0,824	0,863	0,388
			0,039	0,029	0,030	0,043	0,019	0,036	0,031	0,016	0,063
		0,10	0,985	0,962	0,962	0,850	0,782	0,805	0,875	0,979	0,876
			0,019	0,016	0,023	0,017	0,043	0,043	0,035	0,020	0,037
	30	0	0,827	0,863	0,954	0,909	0,950	0,978	0,936	0,917	0,575
			0,039	0,044	0,020	0,028	0,025	0,016	0,029	0,013	0,047
		0,05	0,823	0,864	0,941	0,901	0,939	0,966	0,927	0,909	0,547
			0,044	0,042	0,023	0,029	0,026	0,019	0,031	0,013	0,052
		0,10	0,824	0,861	0,929	0,891	0,931	0,952	0,912	0,900	0,517
			0,034	0,041	0,030	0,031	0,026	0,024	0,032	0,013	0,046
500	45	0	0,904	0,922	0,954	0,932	0,951	0,963	0,955	0,940	0,674
			0,029	0,026	0,022	0,023	0,023	0,019	0,016	0,009	0,042
		0,05	0,902	0,918	0,944	0,922	0,947	0,956	0,943	0,933	0,647
			0,025	0,024	0,025	0,032	0,021	0,020	0,023	0,008	0,038
		0,10	0,891	0,909	0,937	0,913	0,932	0,928	0,946	0,922	0,602
			0,021	0,020	0,038	0,021	0,029	0,015	0,016	0,010	0,046
	15	0	0,800	0,815	0,893	0,899	0,960	0,902	0,847	0,874	0,415
			0,021	0,019	0,013	0,012	0,010	0,014	0,017	0,006	0,021
		0,05	0,797	0,813	0,880	0,888	0,947	0,893	0,837	0,865	0,395
			0,022	0,016	0,017	0,013	0,011	0,012	0,018	0,007	0,023
		0,10	0,793	0,807	0,874	0,878	0,935	0,884	0,829	0,857	0,380
			0,019	0,018	0,016	0,017	0,011	0,014	0,016	0,006	0,022
	30	0	0,845	0,882	0,954	0,909	0,952	0,977	0,944	0,923	0,603
			0,018	0,017	0,008	0,014	0,008	0,006	0,014	0,006	0,026
		0,05	0,843	0,870	0,945	0,901	0,950	0,970	0,932	0,916	0,579
			0,013	0,021	0,007	0,013	0,010	0,006	0,011	0,003	0,017
		0,10	0,830	0,866	0,937	0,894	0,937	0,961	0,926	0,907	0,541
			0,020	0,018	0,010	0,011	0,011	0,009	0,015	0,005	0,022
	45	0	0,917	0,931	0,957	0,938	0,961	0,969	0,967	0,949	0,712
			0,012	0,012	0,009	0,015	0,008	0,008	0,009	0,003	0,018
		0,05	0,913	0,925	0,954	0,938	0,952	0,965	0,959	0,944	0,688
			0,013	0,017	0,009	0,016	0,009	0,009	0,009	0,005	0,020
		0,10	0,906	0,920	0,944	0,928	0,946	0,959	0,953	0,937	0,658
			0,013	0,014	0,010	0,013	0,010	0,009	0,009	0,004	0,017

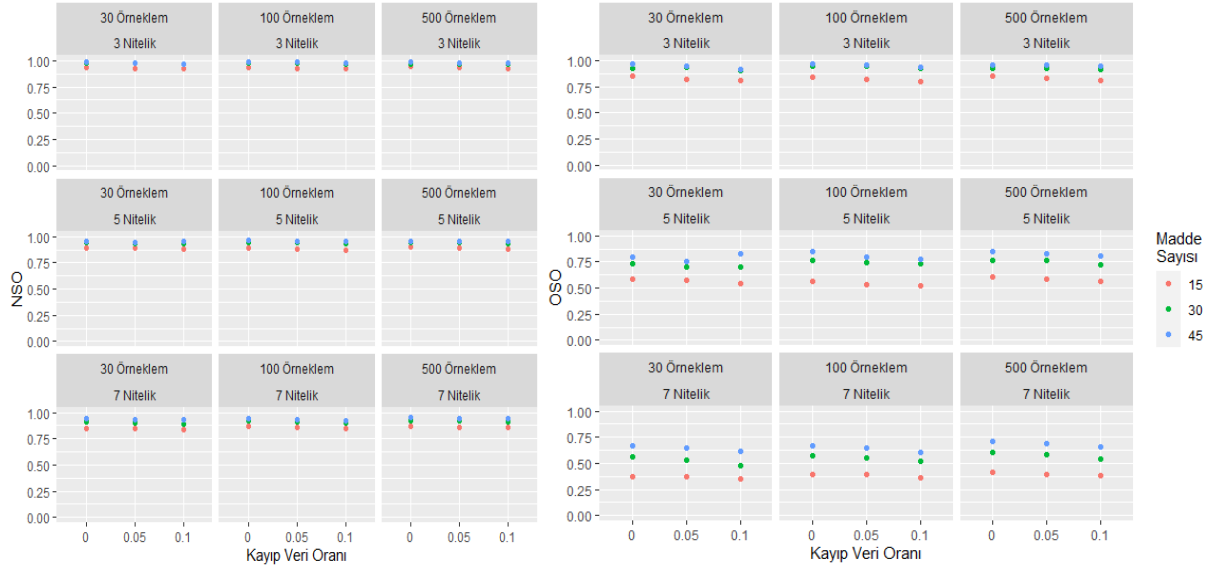
Not: NSO_{Ni1} : Birinci niteliğe; NSO_{Ni2} : İkinci niteliğe; NSO_{Ni3} : Üçüncü niteliğe; NSO_{Ni4} : Dördüncü niteliğe; NSO_{Ni5} : Beşinci niteliğe; NSO_{Ni6} : Altıncı niteliğe; NSO_{Ni7} : Yedinci niteliğe ait sınıflandırma oranı ve NSO : Genel nitelik düzeyi sınıflandırma oranı. Her bir hücrenin altındaki değer o hücredeki sınıflandırma oranlarının standart sapmalarını göstermektedir.



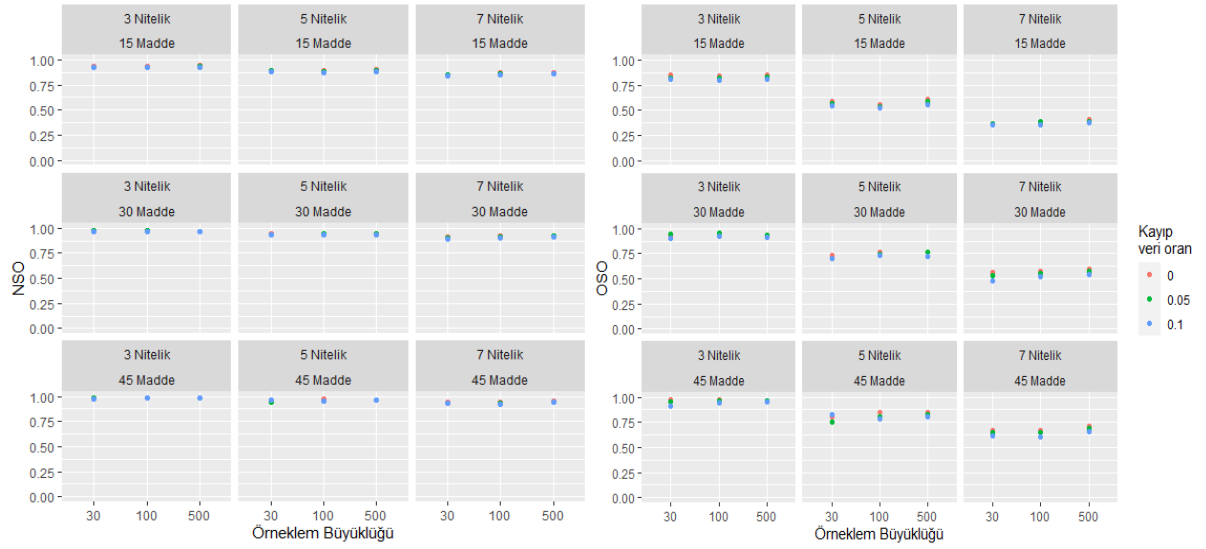
Şekil 30. Nitelik sayılarına (3, 5 ve 7) göre DINO modelin NSO ve ÖSO'ları



Şekil 31. Örneklem büyüklüğüne (30, 100 ve 500) göre DINO modelin ÖSO'ları



Şekil 32. Madde sayısına (15, 30 ve 45) göre DINO modelin NSO ve ÖSO'ları



Şekil 33. Kayıp veri oranına (0; 0,05 ve 0,10) göre DINO modelin NSO ve ÖSO'ları

EK 8. Değişen Koşullar Altında POBT ve DINO Modelinin Uyum İndeksleri

Tablo 41

Değişen Koşullar Altında POBT ve DINO Modelinin Uyum İndeksleri

Ni	N	M	%	POBT			DINO			
				AIC	BIC	mean.rmsea	loglikelihood	deviance	AIC	BIC
3	30	15	0	404,523	456,368	0,172	-229,863	459,726	533,726	585,570
			.05	406,883	458,728	0,163	-223,239	446,479	520,479	572,323
			.10	420,354	472,198	0,163	-223,364	446,728	520,728	572,573
		30	0	913,853	1007,734	0,209	-416,444	832,888	966,889	1060,769
			.05	900,125	994,005	0,204	-419,256	838,512	972,512	1066,392
			.10	877,088	970,968	0,208	-409,676	819,353	953,353	1047,233
		45	0	1377,042	1512,958	0,201	-642,048	1284,096	1478,096	1614,012
			.05	1351,610	1487,526	0,240	-642,722	1285,443	1479,443	1615,359
			.10	1316,992	1452,908	0,216	-633,196	1266,391	1460,391	1596,308
	100	15	0	1435,936	1532,327	0,102	-820,093	1640,186	1714,186	1810,577
			.05	1452,361	1548,753	0,097	-824,831	1649,662	1723,662	1820,053
			.10	1294,177	1390,568	0,092	-838,758	1677,515	1751,516	1847,907
		30	0	2878,282	3052,828	0,119	-1527,462	3054,925	3188,925	3363,471
			.05	2729,411	2903,958	0,116	-1520,665	3041,330	3175,330	3349,876
			.10	2776,766	2951,313	0,119	-1500,168	3000,337	3134,337	3308,883
		45	0	4444,659	4697,360	0,121	-2230,088	4460,176	4654,177	4906,878
			.05	4238,101	4490,802	0,123	-2223,910	4447,819	4641,819	4894,521
			.10	4277,134	4529,836	0,125	-2199,461	4398,922	4592,922	4845,623
	500	15	0	6914,918	7070,859	0,042	-4169,469	8338,938	8412,938	8568,878
			.05	6582,786	6738,727	0,044	-4099,754	8199,509	8273,509	8429,449
			.10	5883,812	6039,753	0,047	-4070,229	8140,459	8214,459	8370,399
		30	0	14081,715	14364,093	0,053	-7781,566	15563,133	15697,133	15979,512
			.05	13180,590	13462,969	0,055	-7745,914	15491,828	15625,828	15908,207
			.10	14139,571	14421,950	0,059	-7702,725	15405,449	15539,450	15821,828
		45	0	19810,015	20218,832	0,055	-11263,087	22526,173	22720,173	23128,990
			.05	20361,894	20770,711	0,056	-11179,865	22359,729	22553,729	22962,546
			.10	20026,675	20435,492	0,061	-11065,229	22130,457	22324,457	22733,274
5	30	15	0	445,524	530,997	0,240	-227,058	454,117	576,117	661,590
			.05	478,146	563,619	0,226	-225,661	451,323	573,323	658,796
			.10	426,755	512,228	0,239	-224,455	448,910	570,910	656,383
		30	0	979,123	1106,632	0,302	-435,942	871,885	1053,885	1181,394
			.05	909,824	1037,333	0,316	-434,590	869,179	1051,180	1178,688
			.10	929,890	1057,399	0,297	-438,223	876,446	1058,446	1185,955
		45	0	1381,512	1551,057	0,347	-647,436	1294,872	1536,872	1706,417
			.05	1334,211	1503,756	0,351	-651,782	1303,565	1545,565	1715,110
			.10	1322,806	1492,350	0,304	-617,275	1234,549	1476,549	1646,094

7	100	15	0	1229,765	1388,681	0,149	-842,889	1685,778	1807,778	1966,693
			.05	1258,459	1417,374	0,149	-840,344	1680,688	1802,688	1961,603
			.10	1186,624	1345,539	0,141	-826,150	1652,300	1774,300	1933,216
		30	0	2627,350	2864,421	0,215	-1575,793	3151,586	3333,586	3570,657
			.05	2678,179	2915,250	0,212	-1562,832	3125,665	3307,665	3544,735
			.10	2705,561	2942,632	0,211	-1563,984	3127,968	3309,968	3547,039
		45	0	3856,557	4171,783	0,237	-2303,396	4606,792	4848,792	5164,018
			.05	4271,909	4587,135	0,227	-2294,701	4589,401	4831,401	5146,627
			.10	4017,356	4332,582	0,223	-2270,429	4540,859	4782,859	5098,084
	500	15	0	6076,245	6333,336	0,072	-4343,373	8686,746	8808,746	9065,837
			.05	5529,541	5786,632	0,066	-4301,156	8602,312	8724,313	8981,404
			.10	5224,984	5482,076	0,062	-4266,010	8532,020	8654,020	8911,111
		30	0	13906,078	14289,608	0,102	-8102,518	16205,036	16387,036	16770,565
			.05	11940,851	12324,380	0,101	-8063,312	16126,624	16308,624	16692,153
			.10	13243,066	13626,595	0,101	-8079,624	16159,249	16341,249	16724,778
		45	0	20221,575	20731,543	0,113	-11762,049	23524,097	23766,097	24276,065
			.05	22002,281	22512,249	0,114	-11654,627	23309,254	23551,253	24061,221
			.10	20577,386	21087,353	0,115	-11620,327	23240,654	23482,654	23992,622
	7	30	15	0	617,938	837,926	0,283	-223,877	447,754	761,754
			.05	592,660	812,648	0,282	-223,362	446,724	760,723	980,711
			.10	600,753	820,741	0,257	-217,913	435,825	749,825	969,813
		30	0	1131,077	1393,101	0,372	-444,531	889,061	1263,061	1525,085
			.05	1061,851	1323,875	0,362	-437,972	875,944	1249,945	1511,968
			.10	1072,949	1334,973	0,356	-436,207	872,415	1246,415	1508,439
		45	0	1534,798	1838,858	0,404	-650,093	1300,186	1734,185	2038,245
			.05	1503,452	1807,512	0,389	-646,046	1292,092	1726,092	2030,152
			.10	1562,344	1866,403	0,395	-644,218	1288,436	1722,436	2026,496
	100	15	0	1318,417	1727,429	0,221	-852,536	1705,072	2019,072	2428,083
			.05	1263,782	1672,793	0,212	-837,957	1675,913	1989,913	2398,925
			.10	1315,540	1724,552	0,206	-830,064	1660,127	1974,127	2383,139
		30	0	2693,854	3181,021	0,284	-1604,485	3208,969	3582,969	4070,136
			.05	2880,331	3367,498	0,281	-1604,773	3209,545	3583,545	4070,712
			.10	3088,618	3575,785	0,270	-1597,048	3194,096	3568,096	4055,263
		45	0	4198,363	4763,685	0,328	-2353,706	4707,412	5141,412	5706,734
			.05	4232,214	4797,536	0,318	-2332,840	4665,679	5099,679	5665,001
			.10	4263,259	4828,580	0,301	-2316,066	4632,132	5066,133	5631,455
	500	15	0	5561,806	6223,500	0,097	-4479,702	8959,405	9273,405	9935,098
			.05	5729,590	6391,284	0,093	-4432,368	8864,735	9178,735	9840,429
			.10	5907,193	6568,887	0,090	-4369,212	8738,424	9052,424	9714,117
		30	0	12429,830	13217,962	0,154	-8319,330	16638,661	17012,661	17800,793
			.05	13830,807	14618,939	0,150	-8291,401	16582,803	16956,803	17744,935
			.10	14720,973	15509,105	0,146	-8253,325	16506,650	16880,650	17668,781

45	0	20537,840	21452,409	0,185	-12106,614	24213,227	24647,227	25561,797
	.05	20402,936	21317,506	0,181	-12053,881	24107,761	24541,761	25456,331
	.10	20317,345	21231,915	0,175	-11970,842	23941,683	24375,684	25290,254

NOT¹: POBT'nin ortalama RMSEA deęerleri hesaplatılmıř fakat kořulların bazı replikasyonlarında ortalama RMSEA deęerleri hesaplatılamadıęı iin kořul genelinde ortalamaları alınamamıřtır. Bu nedenle bu tabloda POBT'nin deęiřen kořullarda ortalama RMSEA deęerlerinin ortalamalarına yer verilememiřtir.

NOT²: Hata fonksiyonu negatif log-olabilirlik fonksiyonuna eřit olmadıęından YSA'ya ait AIC ve BIC deęerleri hesaplanamamıřtır.





GAZİLİ OLMAK AYRICALIKTIR..