

AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
TEZLİ YÜKSEK LİSANS PROGRAMI

DEĞİŞEN MADDE FONKSİYONU BELİRLEMEDE KULLANILAN
TEKNİKLERİN I. TİP HATA VE İSTATİSTİKSEL GÜÇ ORANLARININ
KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Ayşe BİLİCİOĞLU

Antalya, Mayıs 2019

AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME
TEZLİ YÜKSEK LİSANS PROGRAMI

DEĞİŞEN MADDE FONKSİYONU BELİRLEMEDE KULLANILAN
TEKNİKLERİN I. TİP HATA VE İSTATİSTİKSEL GÜÇ ORANLARININ
KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Ayşe BİLİCİOĞLU

Danışman: Doç. Dr. Bayram BIÇAK

Antalya, Mayıs 2019

DOĞRULUK BEYANI

Yüksek lisans tezi olarak sunduğum bu çalışmayı, bilimsel ahlak ve geleneklere aykırı düşecek bir yol ve yardıma başvurmaksızın yazdığımı, yararlandığım eserlerin kaynakçalarda gösterilenlerden oluştuğunu ve bu eserleri her kullarışımda alıntı yaparak yararlandığımı belirtir; bunu onurumla doğrularım. Enstitü tarafından belli bir zamana bağılı olmaksızın, tezimle ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara katlanacağımı bildiririm.

.../.../.....

Ayşe BİLİCİOĞLU

AKDENİZ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Ayşe BİLİCİOĞLU' nun bu çalışması 27. 05. 2019 tarihinde jürimiz tarafından Eğitim Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme Tezli Yüksek Lisans Programında Yüksek Lisans Tezi olarak oy birliği/ oy çokluğu ile kabul edilmiştir.

İMZA

Başkan : Dr. Öğr. Üyesi Hüseyin SELVİ

Mersin Üniversitesi, Tıp Fakültesi, Tıp Eğitimi Anabilim Dalı

Üye : Dr. Öğr. Üyesi Hakan KOĞAR

Akdeniz Üniversitesi, Eğitim Fakültesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı

Üye (Danışman) : Doç. Dr. Bayram BIÇAK

Akdeniz Üniversitesi, Eğitim Fakültesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı

YÜKSEK LİSANS TEZİNİN ADI: Değişen Madde Fonksiyonu Belirlemede Kullanılan Tekniklerin I. Tip Hata Ve İstatistiksel Güç Oranlarının Karşılaştırılması

ONAY: Bu tez, Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından uygun görülmüş ve Enstitü Yönetim Kurulunun tarihli ve sayılı kararıyla kabul edilmiştir.

Doç. Dr. Ramazan KARATAŞ

Enstitü Müdürü

TEŐEKKÜR

Çalıőmam boyunca bilgisini ve desteęini asla esirgemeyen, yardımlarına ve rehberlięine minnettar olduęum, deęerli hocam Doç. Dr. Bayram BIÇAK' a,

Yüksek lisans eęitimim boyunca yetişmemde çok büyük emeęi olan, bilimsel disiplinini ve bilgi birikimini her zaman örnek aldıęım, öęrencilerini her zaman destekleyen ve yardımcı olmak için elinden geleni yapan kıymetli hocam Doç. Dr. İbrahim Alper KÖSE' ye,

Görüş ve önerileriyle bu araőtırmaya çok deęerli katkılarda bulunan, süreç boyunca sorduęum soruları özen ve ilgi ile yanıtlayan, güzel sohbetleri ile desteklerini asla esirgemeyen Dr. Öğr. Üyesi Hakan KOĞAR'a ve Dr. Öğr. Üyesi Güçlü ŐEKERCİOęLU'na,

Abant İzzet Baysal Üniversitesi ve Akdeniz Üniversitesi Eęitimde Ölçme ve Deęerlendirme Anabilim Dalı'ndaki tüm hocalarıma,

Her daim dostluklarını ve akademik desteklerini yanımda hissettięim arkadaşlarım Gizem BERBER' e, Yusuf AYDIN' a ve Umay Hazar DENİZ' e,

Son olarak, bugünlere gelmemde en büyük emeęi veren, aldıęım tüm kararlarda, her türlü sıkıntıda ve sevincimde yanımda olan, güvenlerini ve sevgilerini tüm kalbimle hissettięim babam Cihan'a, annem Őadiye' ye ve kardeőim Cihat' a ne kadar teşekkür etsem azdır.

Ayőe BİLİCİOęLU

ÖZET

DEĞİŞEN MADDE FONKSİYONU BELİRLEMEDE KULLANILAN TEKNİKLERİN I. TIP HATA VE İSTATİSTİKSEL GÜÇ ORANLARININ KARŞILAŞTIRILMASI

Ayşe BİLİCİOĞLU

Yüksek Lisans, Eğitim Bilimleri Anabilim Dalı

Eğitimde Ölçme ve Değerlendirme Programı

Tez Danışmanı: Doç. Dr. Bayram BIÇAK

Mayıs 2019, 67 Sayfa

Bu araştırmada, iki kategorili puanlanan maddelerde, KTK ve MTK' ye dayalı DMF belirleme tekniklerinin belirlenen koşullar altında I. Tip hata ve istatistiksel güç oranlarının karşılaştırılması amaçlanmıştır. Bu amaç doğrultusunda, Mantel- Haenszel (MH), Lojistik Regresyon (LR), Lord' un χ^2 si ve Raju' nun Alan Ölçümleri teknikleri kullanılarak simülasyon çalışması yapılmıştır.

Temel araştırma modelindeki araştırmada kullanılan simülasyon deseninde, 2PL madde tepki modeli, odak ve referans grubun yetenek dağılımları ile DMF türü sabit tutulan koşullar; örneklem büyüklüğü (1800, 3000), örneklem büyüklüğü oranları (0.50, 1), test uzunluğu (20, 80) ve DMF içeren madde oranı (0, 0.05, 0.10) manipüle edilen koşullardır. Veriler, R programında I. Tip hata ve istatistiksel güç oranları için ayrı ayrı üretilmiştir. Odak ve referans grubundaki a ($\bar{x}= 0.8$, $SS= 0.4$) ve b (-2, 2) parametreleri aynı dağılımdan çekilmiştir. Belirli koşullar altında farklı DMF belirleme tekniklerinin I. Tip hata oranlarının belirlenmesi amacıyla, örneklem büyüklüğü $N_1= 1800$ ve $N_2= 3000$; örneklem büyüklüğü oranları $R/O_1= 0.50$ ve $R/O_2= 1$; test uzunluğu $TU_1= 20$ ve $TU_2= 80$ olacak şekilde değiştirilerek veriler üretilmiştir. İstatistiksel güç oranlarının belirlenmesi amacıyla ise örneklem büyüklüğü $N_1= 1800$ ve $N_2= 3000$; örneklem büyüklüğü oranları $R/O_1= 0.50$ ve $R/O_2= 1$; test uzunluğu $TU_1= 20$ ve $TU_2= 80$ olan koşullara ek olarak DMF içeren madde oranı için odak gruptaki belirli maddelerin b parametresine 0.50 eklenmiştir. DMF içeren maddelerin üretimi, odak grubun veri üretimindeki b parametrelerinin, referans grubun b parametresi üzerinde yapılan

manipölasyonlarla elde edilerek tek biçimli DMF gösteren veriler üretilmiştir. DMF analizlerinde, iki örneklem büyüklüğü ve örneklem büyüklüğü oranı, iki test uzunluğu ve üç DMF içeren madde oranına göre üretilen (2x2x2x3) toplam 24 veri seti için 25 tekrarlama yapılmıştır.

Verilerin analizi, R programındaki “difR” paketi kullanılarak yapılmıştır. MTK’ ye dayalı teknikler için yapılan analizlerde “lrm” paketi kullanılarak; 2PL madde tepki modeline uygun kestirimler yapılmıştır.

Araştırmanın sonucunda, I. Tip hata oranları için tüm tekniklerin, nominal alfa düzeyinin üzerinde şişirilmiş I. Tip hata oranları gösterdiği bulunmuştur. En yüksek hataya MH; en düşük hataya Raju’ nun Alan Ölçümleri tekniğinin sahip olduğuna ulaşılmıştır. Diğer yandan, en yüksek istatistiksel güç oranlarını da MH tekniği vermiştir. I. Tip hata ve istatistiksel güç oranlarına ait bulgular birlikte incelendiğinde, iki kurama ait tekniklerinde 1800 örneklem büyüklüğünde daha iyi performans gösterdiği bulunmuştur. Ek olarak, örneklem büyüklüğünün artması MTK’ den ziyade KTK’ ye dayalı teknikleri daha fazla etkilemektedir. Bununla birlikte, test uzunluğunun 80 madde olduğu ve örneklem büyüklüğünün eşit olmadığı koşullarda, tekniklerin I. Tip hatalarının daha düşük; istatistiksel güç oranlarının daha yüksek olduğuna ulaşılmıştır.

Anahtar Kelimeler: Klasik Test Kuramı, Madde Tepki Kuramı, Değişen Madde Fonksiyonu, I. Tip Hata, İstatistiksel Güç.

ABSTRACT

THE COMPARISON OF DIFFERENT TECHNIQUES USED TO DETERMINE DIFFERENTIAL ITEM FUNCTIONING IN TERMS OF TYPE 1 ERROR AND STATISTICAL POWER RATES

Ayşe BİLİCİOĞLU

Postgraduate, Department of Educational Sciences

Measurement and Evaluation in Education

Supervisor: Assoc. Prof. Dr. Bayram BIÇAK

May 2019, 67 Pages

In this study, it was aimed to compare type 1 error and statistical power rates under the conditions specified by DIF determination techniques based on CTT and IRT in items graded by dichotomous scored. For this purpose, a simulation study was conducted by using Mantel-Haenszel (MH), Logistic Regression (LD), Lord's Chi- Square, and Raju's Areas Measures.

In the simulation design used in basic research model, two- parameter item response model, focus and reference group's ability distribution, and DIF type were the fixed factors while sample size (1800, 3000), rates of sample size (0.50, 1), test length (20, 80) and item rate including DIF (0, 0.05, 0.10) were manipulated factors. The data were produced for Type 1 error and statistical power rates separately in R program. The parameters a ($\bar{x}= 0.8$, $SS= 0.4$) and b (-2, 2) in focus and reference groups were drawn from the same distribution. In order to determine Type 1 error rates by using DIF determination techniques under specific conditions, the data were produced by determining samples size as $N_1= 1800$ and $N_2= 3000$, the rates of sample size as $R/O_1= 0.50$ and $R/O_2= 1$, test length as $TU_1= 20$ and $TU_2= 80$. For the purpose of specifying statistical power rates, in addition to the same factors determined for Type 1 error, 0.50 was added to parameter b of certain items in the focus group for items including DIF rates. The items including DIF were produced by manipulating b parameter of reference group which was taken from focus group. As a result, data were produced showing uniform DIF. In DIF analyses, 25 replications were conducted for a total of 24 data set ($2 \times 2 \times 2 \times 3$) produced

according to two sample sizes, two rates of sample sizes, two test lengths, and three rates of items including DIF mentioned above.

The analysis of the data was carried out by using “difR” pack in R program. In the analyses conducted for techniques based on IRT, predictions were made in accordance with 2PL item response design by using “ltm” command.

In the current study, it was found that all techniques used for Type 1 error showed Type 1 error rates inflated above nominal alpha level. It was also demonstrated that MH had the highest error rate while Raju’s Areas Measures had the lowest error rate. On the other hand, MH technique produced the highest statistical power rates. The analysis of the findings of Type 1 error and statistical power rates illustrated that techniques based on both of the theories performed better in 1800 sample size. Furthermore, the increase in the sample size affected techniques based on CTT rather than IRT. Also, the findings demonstrated that the techniques’ Type 1 error rates were lower while their statistical power rates were higher under the conditions where test length was 80-item long and the sample sizes were not equal.

Keywords: Classical Test Theory, Item Response Theory, Differential Item Functioning, Type 1 Error, Statistical Power.

TABLULAR LİSTESİ

Tablo 2.1. DMF belirleme teknikleri	11
Tablo 2.2. Eşleştirilen Değişkenin j Düzeyinde Bir Maddeye Ait Kontincensi Tablosu.....	14
Tablo 2.3. MH DMF İstatistiğinin Yorumlanması.....	15
Tablo 2.4. Bradley' in I. Tip Hata Oranlarına İlişkin Sınıflandırması.....	21
Tablo 3.1. Araştırmanın manipüle edilen koşulları.....	28
Tablo 4.1. MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si Tekniklerinin 1800 Örneklem Büyüklüğü için I. Tip Hata Oranları.....	30
Tablo 4.2. MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si Tekniklerinin 3000 Örneklem Büyüklüğü için I. Tip Hata Oranları.....	31
Tablo 4.3. MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si Tekniklerinin 1800 Örneklem Büyüklüğü için İstatistiksel Güç Oranları.....	32
Tablo 4.4. MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si Tekniklerinin 3000 Örneklem Büyüklüğü için İstatistiksel Güç Oranları.....	34

ŞEKİLLER LİSTESİ

Şekil 2.1. DMF Göstermeyen Madde Grafiği	9
Şekil 2.2. Tek Biçimli DMF Gösteren Madde Grafiği	10
Şekil 2.3. Tek Biçimli Olmayan DMF Gösteren Madde Grafiği.....	10
Şekil 2.4. İki MKE Arasındaki Alan.....	18
Şekil 2.5. İki Farklı Grup İçin MKE'nin Grafiği.....	19



İÇİNDEKİLER

TEŞEKKÜR.....	i
ÖZET.....	ii
ABSTRACT.....	iv
TABLolar LİSTESİ.....	vi
ŞEKİLLER LİSTESİ.....	vii
İÇİNDEKİLER.....	viii

BÖLÜM I

GİRİŞ

1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı.....	4
1.3. Araştırmanın Önemi.....	5
1.4. Araştırmanın Varsayımları.....	6
1.5. Araştırmanın Sınırlılıkları.....	6

BÖLÜM II

KURAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

2.1. Madde Yanlılığı ve Değişen Madde Fonksiyonu.....	8
2.2. Değişen Madde Fonksiyonu.....	8
2.3. Klasik Test Kuramına Dayalı Teknikler.....	12
2.3.1. Mantel – Haenszel (MH).....	13
2.3.2. Lojistik Regresyon(LR).....	15
2.4. Madde Tepki Kuramına Dayalı Teknikler.....	17
2.4.1. Raju' nun Alan Ölçümleri.....	17
2.4.2. Lord' un χ^2 'si (Ki-Kare) Tekniğı.....	19
2.5. İlgili Araştırmalar.....	21

BÖLÜM III

YÖNTEM

3.1. Araştırmanın Modeli.....	25
3.2. Simülasyon Deseni.....	25
3.2.1. Sabit Koşullar.....	25
3.2.1.1. İki Parametrelili Madde Tepki Modeli.....	25
3.2.1.2. Odak ve Referans Grubun Yetenek Dağılımları.....	26
3.2.1.3. DMF Türü.....	26
3.2.2. Manipüle Edilen Koşullar.....	26
3.2.2.1. Örneklem Büyüklüğü ve Örneklem Büyüklüğü Oranları.....	26
3.2.2.2. Test Uzunluğu.....	26
3.2.2.3. DMF İçeren Madde Oranı.....	27
3.3. Verilerin Üretimi.....	27
3.4. Verilerin Analizi.....	29

BÖLÜM IV

BULGULAR

4.1. I.Tip Hata Oranlarına İlişkin Bulgular.....	30
4.1.1. 1. Alt Probleme İlişkin Bulgular.....	30
4.1.2. 2. Alt Probleme İlişkin Bulgular.....	31
4.2. İstatistiksel Güç Oranlarına İlişkin Bulgular.....	32
4.2.1. 3. Alt Probleme İlişkin Bulgular.....	32
4.2.2. 4. Alt Probleme İlişkin Bulgular.....	33

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

5.1. Sonuç ve Tartışma.....	36
5.2. Öneriler.....	42
5.2.1. Uygulamaya Dönük Öneriler.....	42
5.2.2. İleriye Dönük Araştırmalara İlişkin Öneriler.....	43
KAYNAKÇA.....	46

ÖZGEÇMİŞ.....	52
İNTİHAL RAPORU.....	54



I. GİRİŞ

Tezin bu bölümünde; ilgili literatür ile birlikte tezin konusu olarak ele alınan problem ve alt problemlere, araştırmanın amacına, araştırmanın önemine, sınırlılıklarına ve araştırmada yapılan varsayımlara ilişkin bilgiler yer almaktadır.

1.1. Problem Durumu

Bireylerin yetenekleri, ilgileri, kişilik özellikleri, tutum, uyum ve başarı gibi özelliklerinin ölçülmesi amacıyla çeşitli testler ve ölçekler kullanılmaktadır. Bu ölçme araçlarından elde edilen veriler ile bireylere ait ölçülen nitelik veya özelliklerin düzeyleri belirlenmektedir. Bu ölçümler doğrultusunda bireyler hakkında önemli kararlar alınmaktadır. Bu alınan kararların kişiler ve kurumlar açısından hayatiyeti mümkün olduğunca objektif ve doğru bir şekilde ölçme yapmanın önemi ortaya çıkmaktadır. Bireylere uygulanan ölçme araçlarından elde edilen ölçümlerin doğru ve hassas olması, geniş bilgi verebilmesi oldukça önemlidir.

Bir ölçme aracında bulunması gereken iki temel özellik vardır. Bu özellikler, güvenilirlik ve geçerliktir. Bir ölçme aracının olmazsa olmaz iki özelliğinden biri olan güvenilirlik ölçümlerin kararlılığına ve tutarlılığına işaret eder. Güvenirlik, bir araştırmanın bulgularının gerçeği yansıtıp yansıtmadığı ve bunun derecesi, araştırma farklı zaman veya farklı uygulayıcılar tarafından uygulandığında benzer sonuçlara ulaşılması ile ilgilidir. Güvenirlik bir testin karar vermek amacıyla kullanabilmesinde önemli bir özellik olmasına rağmen tek başına yeterli değildir. Ölçme aracının amaçlanan özelliği ölçüp ölçemediği geçerlik kavramı ile ortaya konulmaktadır. “Geçerlik, kanıtların ve teorilerin testlerin önerilen kullanımıyla elde edilen test puanlarının yorumlarını destekleme derecesini ifade eder.” (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999).

Teknik olarak güvenilirlik ve geçerlik açısından yeterli görülen bazı testlerin uygulama süreçleri, uygulanan kişilerin karakteristikleri ya da kültürel faktörler gibi nedenlerle yanlış ölçmelere neden olabilmektedir. Yetenek düzeyleri aynı olan ancak farklı alt gruplarda yer alan bireylerin bir maddeye veya bir teste verdikleri tepkiler farklılaşabilmektedir. Bu da maddenin veya testin farklı alt gruplarda yer alan bireylerde, farklı fonksiyon göstermesi durumuyla sıklıkla

karşılaşılmasına neden olmaktadır. Camilli ve Shepard (1994) yanlılığı, test puanlarının belirli bir gruba karşı sistematik hata içermesi olarak tanımlamaktadır. Aynı yetenek düzeyindeki gruplardan birine avantaj; diğerine dezavantaj sağlayan bu durum yalnızca ölçme aracında yer alan birkaç madde için söz konusu ise madde yanlılığı, testin genelinde var olan bir durum ise test yanlılığı olarak adlandırılmaktadır. Bu noktada madde ve test yanlılığı geçerliği etkileyen en önemli tehditlerin arasındadır (Clauser ve Mazor, 1998; Zumbo, 1999). Test puanları amaçlanan yapıyı gruplar veya durumlar arasında aynı şekilde yansıtamazsa (veya farklı gruplar için farklı yapılara işaret ederse) puan yorumları yanlı olur (Kaya, 2017). Madde yanlılığı ise test koşullarından ya da maddenin bazı özelliklerinden dolayı, bir maddeyi bir grubun doğru yanıtlama olasılığının diğer gruba göre daha az ya da çok olması durumudur (Zumbo, 1999). Bu durumda maddeyi doğru cevaplama olasılığı ölçülen yetenek düzeyinden çok, herhangi bir gruba ait olmaya bağlı olmaktadır. Bu durum ölçme aracının hatalı olduğunun bir göstergesi olarak kabul edilmektedir (Osterlind, 1983). Clauser ve Mazor (1998) bir testin tarafsızlığına ve geçerliğine ilişkin en büyük tehditlerden birinin madde ve test yanlılığı olduğunu söylerken, Camilli ve Shepard (1994) değişen madde fonksiyonunu (DMF) yanlılığın bir indeksi olarak tanımlamışlardır.

Madde yanlılığını tespit etme sürecinde izlenen yaklaşımlar dikkate alındığında bir maddenin yanlı olup olmadığının belirlenmesinin ilk aşaması, DMF belirlemek amacıyla gerekli istatistiksel analizlerin uygulanmasıdır. DMF belirlemek amacıyla yapılacak analizler için belirlenen alt gruplardan biri odak; diğeri referans grup olarak adlandırılır. Odak ve referans grubunda olan bireylerin performansları karşılaştırılır. Aynı yetenek düzeyinde olmasına rağmen farklı gruplarda yer alan bireylerin, bir maddeyi doğru cevaplama olasılıklarının farklı olması durumunda, maddenin DMF gösterdiği söylenmektedir. Fakat bu maddeye kesin olarak yanlıdır denilemez. Yanlı olan bir maddede DMF gözlenir, ancak DMF gösteren her madde yanlı değildir. Bu fark, alt gruplar arasındaki gerçek farktan da kaynaklı olabilir. Bu yüzden yanlılığın olup olmadığının belirlenmesinin ikinci aşamasına geçilir. DMF gösteren maddelerde uzman görüşü, içerik analizi gibi teknikler kullanılarak madde yanlılığının olup olmadığına karar verilir (Camilli ve Shepard, 1994).

DMF belirlemek amacıyla kullanılan pek çok teknik vardır. İki veya çok kategorili verilerde kullanılan farklı DMF belirleme teknikleri bulunmaktadır. Teknikler, genel olarak iki kuram altında toplanmaktadır. Klasik Test Kuramı (KTK) 'na dayalı olarak geliştirilen DMF belirleme

tekniklerinin sık kullanılanlarından bazıları; Mantel- Haenszel (MH), Varyans Analizi, Dönüştürülmüş Madde Güçlüğü ve Lojistik Regresyon (LR) olarak sıralanabilir. Madde Tepki Kuramı (MTK) 'na dayalı tekniklerin sık kullanılanlarından bazıları ise; Olabilirlik Oran Testi, Raju'nun Alan Ölçümleri ve Lord'un χ^2 'dir (Camilli ve Shepard, 1994; Hambleton, Swaminathan ve Rogers, 1991).

DMF' nin geçerlik üzerine etkisi göz önünde bulundurulduğunda bu teknikler ve tekniklerin hangi koşullarda nasıl çalıştıklarını ortaya koymak oldukça önemlidir. Literatür incelendiğinde, DMF üzerine yapılan araştırmaların bir kısmının gerçek veriler bir kısmının ise simülasyon veriler üzerinde yapıldığı görülmektedir. Yapılan bu araştırmalar “ Maddenin DMF içermesindeki temel neden nedir?”, “Bir madde DMF içeriyorsa ne yapılmalıdır?” gibi soruların yanıtını bulmak amacıyla yapılmaktadır. Araştırmacılar ise bu soruların yanıtı olarak, ilgili maddenin testten çıkarılarak yeni bir maddenin eklenmesini çözüm yolu olarak sunmaktadırlar. Ancak bu süreç, zaman alıcı olmak ile birlikte testin kapsam geçerliğine de etki etmesine neden olacaktır. Getirilen bir diğer öneri ise, DMF içeren maddenin düzeltilmesi şeklindedir. Bu öneri ise araştırmacıları, düzeltilmiş maddelerin bir gruba uygulanıp, yeniden DMF analizi yapılmasını gerektirmesi problemi ile yüz yüze bırakmaktadır (Ellis ve Raju, 2003). Bu nedenler ile DMF kaynaklarının belirlenmesi ve madde yanlılığı varlığının azaltılması adına yapılacak çalışmalar oldukça önemlidir.

Daha önce değinilen DMF' li maddenin yanlı olup olmadığının belirlenmesi amacıyla kullanılan tekniklerden en sık kullanılanı uzman görüşüdür. Ancak, araştırmacılar bazı durumlarda ilgili maddenin neden yüksek düzeyde DMF içerdiği sorusuna yanıt bulamadıklarını ifade etmektedirler. Bu eksikliğin giderilmesi amacıyla bir maddedeki DMF varlığının olası diğer nedenleri olan I. Tip hata ve istatistiksel güç oranları üzerine yapılan araştırmalar gerek duyulmuştur.

I. Tip hata bir testte yer alan madde gerçekte DMF içermezken, yapılan analiz ile madde DMF içeriyor şeklinde bir sonuca ulaşılmasıdır. İstatistiksel güç ise ilgili maddenin DMF içeriyorken, gerçekte olduğu gibi DMF içeriyor kararının verilmesidir.

Bu araştırmada, KTK ve MTK kuramları altında yer alan farklı tekniklerin karşılaştırılması planlanmıştır. Aynı tekniklerin farklı koşullar altında nasıl sonuçlar verdiğinin ortaya konulması, diğer tekniklerle kıyaslanması ve I.Tip hata ve istatistiksel güç oranlarının belirlenmesi, araştırmacıların kendi ölçme araçları için doğru tekniği seçmesi açısından oldukça önem arz etmektedir. Çünkü DMF' nin belirlenmesi amacıyla geliştirilen tekniklerin

performansı aynı düzeyde değildir. Bu tekniklerin farklı test uzunluğu, DMF içeren madde oranı, model türü (2PLM ve 3 PLM) ve örneklem büyüklüğü durumlarında, tekniklerin performanslarının nasıl farklılaştığının ve hangi durumlarda kullanıma uygun olduğunun ortaya konulması gerekmektedir. Kristjansson, Aylesworth, Mcdowell ve Zumbo (2005), madde yanlılığının, yapılan ölçmelerin geçerliğini tehdit eden güçlü bir etken olduğunu; bu yanlılıkların belirlenmesi amacıyla kullanılan teknikler üzerine mümkün olduğunca çalışılması gerektiğini belirtmişlerdir. Literatürde fark edilen MTK' ye dayalı teknikler üzerine yapılan çalışmaların eksikliğinin giderilmesi de amaçlanarak bu araştırmada yalnızca KTK' ye değil, aynı zamanda MTK' ye dayalı tekniklerin de kullanımı tercih edilmiştir. Ancak hem yukarıda belirtilen koşullarda yapılan bir karşılaştırma araştırmasında, hem de MTK gibi büyük örneklem isteyen bir kuramın altında gelişen tekniklerin kullanılacak olması sebebiyle gerçek veriler kullanılarak araştırmanın yürütülmesi oldukça güçtür. Simülasyon veri ile istenilen koşullarda, gerektiği kadar veri üretilmektedir. Bu nedenle araştırmada simülasyon veri kullanımı tercih edilmiştir.

Araştırmanın problem durumu, KTK ve MTK' ye dayalı dört DMF belirleme tekniğinin; test uzunluğu, DMF içeren madde oranı, örneklem büyüklüğü ve örneklem büyüklüğü oranlarının değiştiği farklı koşullar altında, I.Tip hata oranları ve istatistiksel güç oranlarının belirlenmesidir.

1.2. Araştırmanın Amacı

Bu araştırmanın genel amacı iki kategorili puanlanan maddelerde, farklı kuramlara ait DMF belirleme tekniklerinin belirlenen koşullar altında I. Tip hata ve istatistiksel güç oranlarının belirlenmesidir. Bu amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır.

1. Örneklem büyüklüğünün 1800; örneklem büyüklüğü oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF içeren madde oranının 0 olduğu koşullar altında MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si tekniklerinin I.Tip hata oranları nasıl değişmektedir?
2. Örneklem büyüklüğünün 3000; örneklem büyüklüğü oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF içeren madde oranının 0 olduğu koşullar altında MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si tekniklerinin I. Tip hata oranları nasıl değişmektedir?

3. Örneklem büyüklüğünün 1800; örneklem büyüklüğü oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF içeren madde oranının 0.05 ve 0.1 olduğu koşullar altında MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si tekniklerinin istatistiksel güç oranları nasıl değişmektedir?

4. Örneklem büyüklüğünün 3000; örneklem büyüklüğü oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF içeren madde oranının 0.05 ve 0.1 olduğu koşullar altında MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si tekniklerinin istatistiksel güç oranları nasıl değişmektedir?

1.3. Araştırmanın Önemi

Bir ölçme işleminin hatalardan tamamen arınık olmasının olanaksız olduğu bilinmektedir. Ölçme sürecinde kullanılan araca, ölçme işlemini yapan uygulamacıya ya da çevresel faktörlere bağlı olarak minimal düzeyde de olsa hatalar mevcuttur. Dolaylı ölçmenin sıklıkla yer aldığı sosyal bilimlerde ise bu türden hataların daha da büyük oranlarda gözlenebileceği şüphesizdir. Ölçme işlemine karışan bu hatalardan biri de yapı ve yordama geçerliğini tehdit eden madde yanlılığıdır. Dolayısıyla içerisinde yanlı madde barındıran bir ölçme aracı ile yapılan ölçümler sonucunda alınacak kararlar, hem geçerlik hem de güvenilirlik açısından sorunlu olmasına neden olacaktır. Madde yanlılığının belirlenmesinde ilk adım olarak DMF analizleri vardır. Ölçme aracı geliştirme, uyarılma ve yapılan ölçme işlemlerine dayalı bir şekilde karar verme aşamasında DMF analizlerinin kullanıldığı görülmektedir. Yapılan analizlerin sonucuna bağlı olarak testin geliştirilmesi sürecinde, yanlı maddelerin testten çıkarılması ya da düzeltilmesi yoluna gidilmektedir. Bu şekilde hazırlanan ölçme aracının geçerliği artırılmış olur. Ancak, bir ölçme aracında yer alan maddenin yanlı olduğunu söyleyebilmemiz için yalnızca yapılan DMF analizleri yeterli olmayacaktır. DMF içerdiği işaret edilen madde için uzman görüşü alınmakta ve alınan görüşler doğrultusunda maddenin yanlı olup olmadığına karar verilmektedir. O halde burada karşımıza çıkan problem, maddelerin yanlı olmaması halinde de analizler sonucunda, madde DMF içeriyor şeklinde sonuç elde edebilecek olmamızdır. Oysa ki araştırmalarda istenen durum, gerçekte DMF içeren ya da içermeyen maddenin, analiz sonrasında da gerçek durumu yansıtır şekilde çıkmasıdır. Literatürde bu durum I. Tip hata ve istatistiksel güç kavramları ile açıklanmaktadır. Roussos ve Stout(1996)' ında belirttiği gibi I.Tip hatanın pek çok olumsuz özelliğinin yanı sıra DMF analizlerinde ayırt ediciliği yüksek maddeler içinde tehdit oluşturmaktadır. Çünkü bu maddeler DMF içeriyor olarak tespit edilmeye daha yatkındırlar, bu şekilde testte bulunması gereken bir maddenin testten çıkarılması söz konusu

olacaktır. Bu nedenlerle bir DMF belirleme tekniğinin I. Tip hata ve istatistiksel güç oranlarının belirlenmesinin pratik ve önemli sonuçları vardır. Bu amaç ile bu araştırmada bazı koşullar altında farklı tekniklerin nasıl sonuçlar verdiği ortaya konulmak istenmektedir. Literatürde DMF belirleme teknikleri KTK ve MTK' ye dayanmaktadır. Bu kapsamda hem KTK' ye hem de MTK'ye dayalı DMF belirleme tekniklerinin, farklı koşullar altındaki I.Tip hata ve istatistiksel güç oranları ele alınmıştır. Bu araştırmanın sonuçlarının; örneklem büyüklüğü, örneklem büyüklüğü oranları, test uzunluğu ve DMF içeren madde miktarı söz konusu olduğunda kullanılacak uygun DMF belirleme tekniklerinin tespit edilmesi ile geniş ölçekli testlerle çalışan kurumlara, başlangıçta da değinildiği gibi yapılan testlerin geliştiricilerine ve ilgili testlere ait puanlara göre karar verenlere faydalı olacağı düşünülmektedir.

1.4. Araştırmanın Varsayımları

Araştırma kapsamında R programında üretilmiş olan verilerin, simülasyon koşullarını yansıttığı varsayılmıştır.

1.5. Araştırmanın Sınırlılıkları

Araştırma KTK' ye dayalı LR ve MH, MTK' ye dayalı Raju' nun Alan Ölçümleri ve Lord' un χ^2 si teknikleri ile sınırlıdır. Başlangıçta dahil edilmesi amaçlanan MTK –OO tekniğinin, R programında analizinin tamamlanması çok uzun sürdüğü için ilgili teknik araştırmadan çıkartılmıştır.

Veriler 2PL modele dayalı olarak üretilmiş olup simülasyon deseninde belirlenmiş olan koşullar ile sınırlıdır.

II. KURAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

Günümüzde eğitim programları, eğitim politikaları, bireylerin yeterlilikleri, öğretim yöntem ve teknolojileri gibi unsurların oluşturulması amacıyla veriler elde edilip, değerlendirmeler yapılmaktadır. Yapılan ölçme ve değerlendirme işlemleri, ülkemizin eğitim sisteminin her aşamasında karşımıza çıkmaktadır ve sistemin oluşturulmasında oldukça önemli bir yere sahiptir. Eğitimdeki ölçme işlemleri uygulanan grubun çok fazla bireyden oluşması sebebiyle, büyük çoğunlukla doğrudan değil; ölçme aracı kullanılarak dolaylı şekilde yapılmaktadır. Bu şekilde elde edilen bilimsel verilerle bireylerin gelecekteki yaşantılarını belirleyecek olması, yapılan ölçümlerin sonuçlarına göre alınan kararların önemini arttırmaktadır. Baykul (2000)' unda belirttiği gibi sınavlar, amacına bağlı olarak, alınan puanlar ile bireyler hakkında kararlar almamıza dayanak olmaktadır. Güvenirliği ve geçerliği düşük bir ölçme aracından elde edilen puanlara dayanarak seçme ve yerleştirme yapılması, gerçek durumu yansıtmayacaktır. Tamamen güvenilir ve geçerli puanlara ulaşabilmek için ise, yalnızca ölçülmek istenen yapının bireylerin puanını etkilemesi gerekir. Ancak test puanlarının, ölçülmesi amaçlanan yapı dışında, farklı değişkenlik kaynaklarından etkilenmemesi de imkânsızdır. Bahsedilen değişkenlik kaynaklarının her birini ortadan kaldırmak mümkün olmamasına rağmen; maddelerin bir alt gruba yönelik avantaj ya da dezavantaj sağlayacak şekilde olmadığından emin olunmalıdır (Crocker ve Algina, 1986). Bir ölçme aracındaki performans ile yansıtılan bilgi ve ölçme aracındaki maddeleri yanıtlamak için gereken bilgi ve becerilerin farklı olması söz konusu olabilir. Bu durum testteki maddelerin gruplardan birine üstünlük kazandırdığının göstergesi olabilir. Bu durum madde yanlılığı olarak adlandırılmaktadır. Testte yer alan maddelerin yanlı olması, test puanlarına göre alınacak kararların da hata içermesi anlamına gelir (Crocker ve Algina, 1986; Dorans ve Holland, 1993).

Ölçme aracının şartlarından veya yer alan maddelerin bazı özellikleri sebebiyle, maddeyi belirli bir grubun doğru yanıtlama olasılığının diğer gruba göre daha az veya çok olmasına madde yanlılığı denilmektedir (Camilli ve Shepard, 1994; Mantel ve Haenszel, 1959; Zumbo, 1999). Testler içerisindeki maddelerin yanlılığının belirlenmesi, ölçme aracının güvenilirliğini ve geçerliğini artırmak amacıyla yapılan önemli çalışmalar arasındadır.

2.1. Madde Yanlılığı ve Değişen Madde Fonksiyonu

Madde yanlılığı üzerine yapılan ilk çalışmalara 1960' lı yıllarda başlanmıştır. Ölçme aracı geliştirenler, yer alan maddelerin azınlık ve çoğunluk gruba göre farklılaşıp farklılaşmadığını araştırmışlardır. 1960' lı yıllardan günümüze kadar süren araştırmalar ile madde yanlılığının belirlenmesi amacıyla çeşitli teknikler ortaya çıkmıştır. Madde yanlılığı çalışmaları, aynı yetenek düzeyinde olmasına rağmen farklı gruplarda yer alan bireylerin puanlarının farklılaşıp farklılaşmadığının belirlenmesi amacıyla yapılan istatistiksel bir süreç ile başlar (Demir ve Köse, 2014). Aynı yetenek düzeyindeki farklı alt gruplarda farklılaşan fonksiyonlar veren maddeler, değişen madde fonksiyonlu (DMF) madde olarak adlandırılırlar (Dorans ve Holland, 1993). Öncelikle istatistiksel tekniklerin kullanıldığı DMF analizleri ile grupların fonksiyonlarının farklılaşıp farklılaşmadığı ortaya konulur. Eğer manidar bir farklılaşma söz konusu ise, bu maddelerin saptanmasının ardından farklılığın kaynağı tespit edilmeye çalışılır. Maddenin içeriği ele alınarak üzerinde çalışılan bu süreçte, farklılık kaynağı testin amacının dışında ise maddenin yanlı olduğu kararı alınır (Camili ve Shepard, 1994; Dorans ve Holland, 1993). Bu noktada, DMF içeren bir maddenin belirlenmesi yanlılık için bir ön koşul iken tek başına yeterli değildir (Zumbo, 1999).

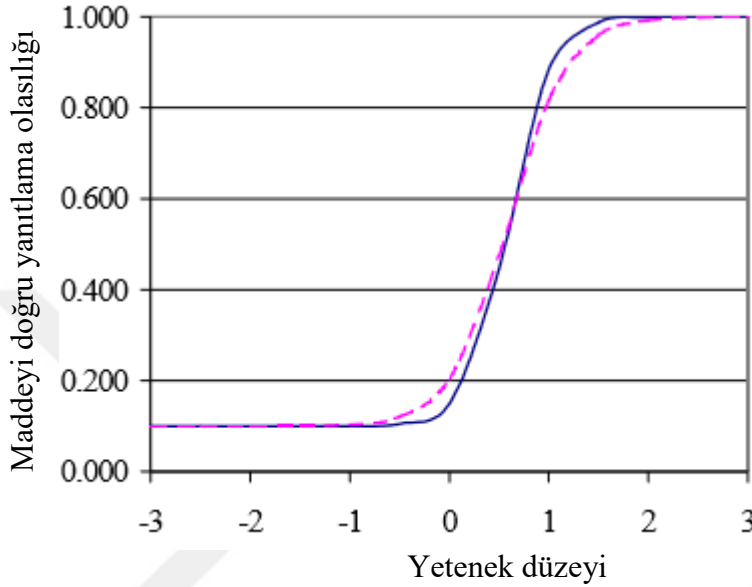
2.2. Değişen Madde Fonksiyonu

Hambleton, Swaminathan ve Rogers (1991) göre DMF, aynı yetenek düzeyinde olmasına rağmen farklı gruplarda olan bireylerin, bir maddeyi doğru yanıtlama olasılıklarının farklı olmasıdır.

DMF nin ortaya çıkmasının iki olası nedeni olabileceği Camilli ve Shepard (1994) tarafından belirtilmiştir. Bunlar, madde etkisi ve madde yanlılığı olarak adlandırılmaktadır. Madde etkisi, farklı gruplarda yer alan yanıtlayıcıların, bir maddeyi doğru yanıtlama olasılıkları arasındaki gerçek farklılıktır (Zumbo,1999). Bu farklılık gruplardan birinin önceden sahip olduğu tecrübe veya bilgiyle açıklanabilmektedir. Eğer yanıtlayıcılar sahip oldukları bilgiler açısından farklılaşıyorsa, maddelere verilen yanıtlarında farklılaşması beklenir. Bu durumda ölçme aracı hatalı değil ve madde yanlı değildir. Madde yanlılığında ise ölçme aracına ölçme esnasında, ölçülmek istenilen yeteneğin dışındaki etmenlerin dahil olması söz konusudur. DMF' nin

belirlenmesindeki amaç da yanlılıktan kaynaklanan farklılıkların açıklanmasıdır. DMF, daha öncede belirtildiği şekilde, madde yanlılığının belirlenmesinin ilk adımındır.

DMF çalışmalarında, testlerde eşleştirme kriteri olarak bireylerin testten aldıkları toplam puanlar kullanılmaktadır. Şekil 2.1.' de, iki gruba dair DMF göstermeyen madde grafiği gösterilmiştir.

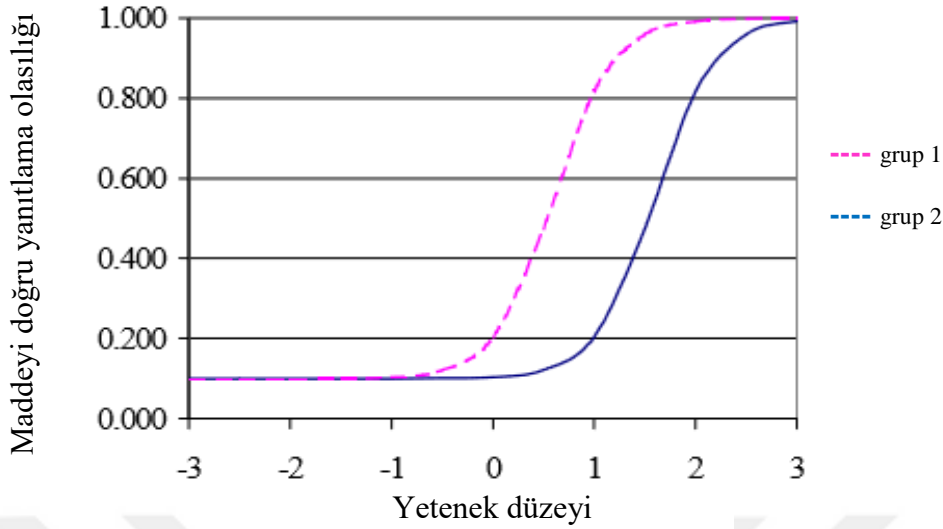


Şekil 2.1. DMF Göstermeyen Madde Grafiği (Zumbo, 1999)

Yukarıda görüldüğü gibi madde karakteristik eğrileri her grup için farksız ya da farksız olmaya çok yakın ise madde DMF göstermiyor demektir. Eğer madde karakteristik eğrileri, iki grup arasında önemli derecede farklılık gösteriyorsa, maddenin DMF gösterdiği söylenir (Zumbo, 1999). DMF, madde veya maddelerin grupların hepsinde aynı biçimde çalışmadığını, madde düzeyindeki değişmezliğin bozulduğunu ve gruplardan biri için maddenin yapı geçerliğinin düşük olduğunu gösterir (Steinberg ve Thissen, 2006).

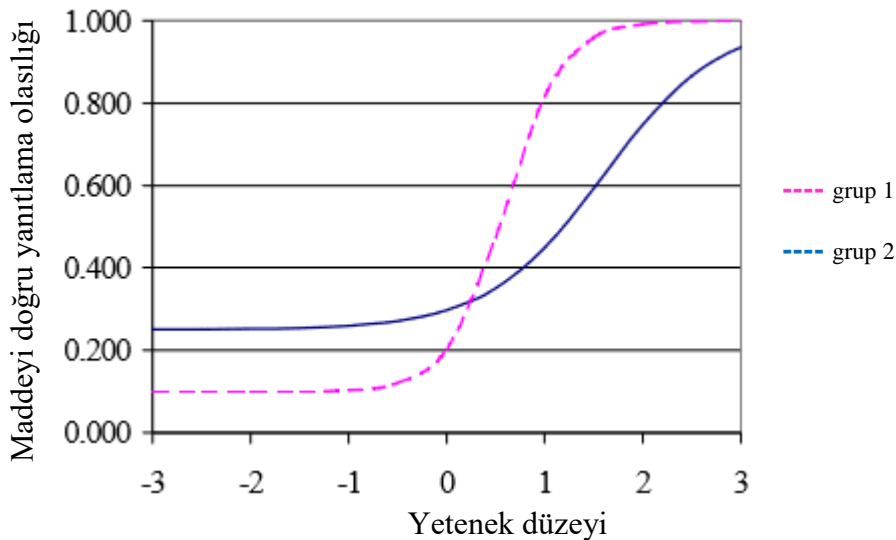
DMF, tek biçimli ve tek biçimli olmayan fonksiyonlar şeklinde ortaya çıkabilir. Tek biçimli DMF, yetenek düzeyleri ve grup arasında etkileşimin olmadığı durumlarda ortaya çıkar; tek biçimli olmayan DMF' de ise bir etkileşim söz konusudur (Swaminathan ve Rogers, 1990).

Şekil 2.2.'de tek biçimli DMF gösteren bir maddeye ait madde karakteristik eğrisi verilmiştir.



Şekil 2.2. Tek Biçimli DMF Gösteren Madde Grafiği (Zumbo, 1999).

Şekil 2.2. incelendiğinde, grup 1 ve grup 2 için madde karakteristik eğrilerinde kesişme olmadığı görülmektedir. Grup 1 için maddenin doğru cevaplanma olasılığı daima daha yüksektir. Tüm yetenek düzeylerinde madde, grup 1'nin lehine işlemiştir. Diğer bir deyişle madde, bütün yetenek düzeyleri boyunca aynı gruba avantaj sağlıyorsa tek biçimli DMF olarak adlandırılır. Tek biçimli DMF gösteren maddelerde iki gruba ait madde karakteristik eğrileri kesişmez. Şekil 2.3.'de tek biçimli olmayan DMF gösteren bir maddeye ait madde karakteristik eğrisi verilmiştir.



Şekil 2.3. Tek Biçimli Olmayan DMF Gösteren Madde Grafiği (Zumbo, 1999).

Şekil 2.3. incelendiğinde, alt yetenek düzeylerinde grup 2' nin maddeyi doğru cevaplama olasılığı grup 1'den daha yüksek iken, üst yetenek düzeylerinde grup 1'in maddeyi doğru yanıtlama olasılığı grup 2'den daha yüksektir. Madde, yetenek düzeyleri boyunca farklı gruplara avantaj sağlamaktadır. Maddenin doğru cevaplanma olasılığının, farklı yetenek düzeylerinde farklı gruplar lehine işlemesi söz konusudur (Zumbo, 1999). Ayrıca, tek biçimli DMF' de madde karakteristik eğrileri kesişmezken; tek biçimli olmayan DMF' de madde karakteristik eğrilerinin kesiştiği görülmektedir.

Genel olarak bakıldığında DMF belirleme teknikleri, Klasik Test Kuramı (KTK)' na ve Madde Tepki Kuramı (MTK)' na dayalı teknikler olarak sınıflandırılmaktadır. KTK' ye dayalı DMF belirleme teknikleri grupların puan dağılımlarını karşılaştırırken; MTK' ye dayalı DMF belirleme tekniklerinde grupların ilgili maddeye doğru yanıt verme olasılıkları karşılaştırmaktadır. Tablo 2.1.' de kuramlara, DMF türüne ve grup sayısı değişkenlerine göre hangi tekniklerin kullanılabileceği verilmiştir.

Tablo 2.1. DMF belirleme teknikleri (Magis, Beland, Tuerlinckx ve Boeck, 2010)

Kuram Türü	DMF Türü	Grup Sayısı	
		2	>2
KTK	Tek Biçimli	Mantel- Haenszel	Genelleştirilmiş Mantel-Haenszel
		Standartlaştırma	
	Lojistik Regresyon		
	SIBTEST		
MTK	Tek Biçimli	Mantel- Haenszel	Genelleştirilmiş Lord'un χ^2 si
		Lojistik Regresyon	
	Olmayan	SIBTEST	
		Breslow-Day	
MTK	Tek Biçimli/ Tek Biçimli Olmayan	Olabilirlik Oran	Genelleştirilmiş Lord'un χ^2 si
		Testi	
		Lord χ^2	
		Raju Alan Ölçümleri	

KTK' ye dayalı DMF belirleme teknikleri, Mantel- Haenszel (MH), Lojistik Regresyon (LR), Standartlaştırma ve SIBTEST; MTK' ye dayalı DMF belirleme teknikleri ise Raju' nun Alan Ölçüleri, Olabilirlik Oran Testi, Lord' un χ^2 si şeklinde sıralanmakla birlikte daha pek çok sayıda istatistiksel teknik geliştirilmiştir (Zumbo, 1999; Camilli ve Shepard, 1994; Hambleton, Swaminathan ve Rogers, 1991; Osterlind,1983).

MTK' ye dayalı teknikler, KTK' ye dayalı tekniklerden farklı olarak model dayanaklıdır. Bu sebeple, doğru yanıt verme olasılığının modellenmesine göre karşılaştırılan parametreler değişiklik gösterebilir. Örneğin, 1 parametrelili modelde (1PLM) gruplar, madde güçlüğüne karşılık gelen b parametresine göre karşılaştırılırken; 2 parametrelili modelde (2PLM) hem b parametresine hem de madde ayırt ediciliği şeklinde yorumlanan a parametresine göre karşılaştırılabilir. b parametresinde ortaya çıkan farklılık, tek biçimli DMF' ye işaret ederken; a parametresindeki farklılık, tek biçimli olmayan DMF' ye işaret eder (Kabasakal, Arsan, Gök, Kelecioğlu, 2014). Bu araştırmada hem madde güçlüğü (b parametresi) hem de madde ayırt ediciliği (a parametresi) kullanıldığı için 2PL model tercih edilmiştir. Kullanılan modeldeki b parametresinin değerlerinde değişiklik yapıldığı için tek biçimli DMF' nin belirlenmesi amacıyla kullanılan teknikler tercih edilmiştir. DMF belirlemek amacıyla geliştirilen tekniklerin her birinin performansı aynı değildir. MTK' ye dayalı teknikler, DIF belirlemek için geniş bir çerçeve oluşturur ve teorik olarak güçlü tekniklerdir. Ancak pratikte sağlanması güç olan büyük örneklem gerektirir (Narayanan ve Swaminathan, 1994). Bu nedenle bu araştırmada yalnızca MTK' ye dayalı teknikler kullanılmamış; yanı sıra KTK' ye dayalı tekniklerde kullanılmıştır. DMF belirleme tekniklerinden, Mantel- Haenszel, Lojistik Regresyon, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si tekniklerinin I. Tip hata ve istatistiksel güç oranları incelenmiştir.

2.3. Klasik Test Kuramına Dayalı Teknikler

Klasik Test Kuramı (KTK), ölçme tarihinde başlangıçtan beri bulunan, testlerin geliştirilmesi, analizler ve ölçüklerin puanlanması gibi pek çok bölümde yaygın bir şekilde kullanılan bir kuramdır. Bu kuram gözlenen puanlar üzerine kuruludur. Gözlenen puanlardan yola çıkılarak, gerçek puan kestirimlerine dayanır. Hata puanının ve gerçek puanın toplamının gözlenen puanı verdiği varsayılır. Eşitlik 1' de klasik test kuramına ait temel denklem verilmiştir (Crocker ve Algina, 1986).

$$X= T+ E \quad (1)$$

X: Gözlenen puan

T:Gerçek puan

E: Hata puanı

KTK' nin temel denklemi olan bu eşitlik incelendiğinde, hatanın düşmesi ile gerçek puan ve gözlenen puanın birbirine yaklaşacağı görülmektedir. Ancak hatasız bir ölçmenin mümkün olmadığı bilinmektedir. Dolayısıyla asla gerçek puan ile gözlenen puan eşit olmayacaktır. KTK' nin pek çok dezavantajı bulunmasına rağmen araştırmacılar tarafından sıklıkla kullanılan bir kuramdır. Ki bunun da sebebi olarak varsayımlarının karşılanmasının kolay olması ve küçük örneklem büyüklüklerinde de kullanılabilir olması gösterilebilir. KTK' ye ait belirtilen avantajlar, KTK' ye dayalı DMF belirleme tekniklerinde görülmektedir. Bu nedenle KTK' ye dayalı DMF belirleme teknikleri de literatürdeki çalışmalarda araştırmacılar tarafından sıkça tercih edilmektedir.

Bu araştırmaya dahil edilen, KTK' ye dayalı DMF belirleme tekniklerinden Mantel- Haenszel ve Lojistik Regresyon teknikleri üzerine ayrıntılı bilgiler aşağıda sunulmuştur.

2.3.1. Mantel – Haenszel (MH)

MH tekniği, ilk olarak Nathan Mantel ve William Haenszel tarafından 1950' li yıllarda, ki- kare uygulaması olarak geliştirilmiştir. Ardından, 1985 yılında Holland tarafından güncellenerek DMF belirlemek amacıyla kullanılmaya başlanmıştır (Osterlind ve Everson, 2009). MH tekniği, büyük örneklem gerektirmemesi, etki büyüklüğü değerini verebilmesi ve hesaplamalarının karmaşık olmaması gibi nedenlerle sıkça tercih edilen bir tekniktir (Samuelsen, 2005).

MH, DMF belirleme tekniği ile ilk olarak H_0 dediğimiz yokluk hipotezi “Yetenek düzeyleri eşleştirildiğinde herhangi bir maddenin yanıtlanması ile ait olduğu grup arasında ilişki yoktur” şeklinde oluşturulur. Yokluk hipotezinde de ifade edildiği gibi odak ve referans grubun bulunduğu teknikte, gözlenen puanlar çeşitli kategorilere ayrılır ve maddeyi doğru yanıtlama oranları açısından karşılaştırma yapılır. Bu karşılaştırmayı yapmak amacıyla bir madde için 2×2 ' lik, içerisinde hem grup üyelerini hem de doğru ve yanlış yanıtların sayısını sunan bir

kontincensi tablosu oluşturulur. Söz edilen kontincensi tablosuna ait bir örnek 2.2.' de verilmiştir (Magis, Beland, Tuerlinckx ve Boeck, 2010).

Tablo 2.2. Eşleştirilen Değişkenin j Düzeyinde Bir Maddeye Ait Kontincensi Tablosu

	1	0	Toplam
Referans	A_j	B_j	n_{Rj}
Odak	C_j	D_j	n_{Oj}
Toplam	m_{1j}	m_{0j}	T_j

A_j : j düzeyinde referans grupta maddeyi doğru yanıtlayanların sayısı

B_j : j düzeyinde referans grupta maddeyi yanlış yanıtlayanların sayısı

C_j : j düzeyinde odak grupta maddeyi doğru yanıtlayanların sayısı

D_j : j düzeyinde odak grupta maddeyi yanlış yanıtlayanların sayısı

m_{1j} ve m_{0j} : j düzeyinde referans ve odak gruptaki toplam birey sayısı

n_{Rj} ve n_{Oj} : j düzeyinde maddeyi doğru ve yanlış yanıtlayanların sayısı

Yukarıda verilmiş olan tablodaki değerleri de içinde barındıran MH tekniği ile DMF belirlerken kullanılan formül Eşitlik 2' de verilmiştir.

$$MH = \frac{\left(\left| \sum_j A_j - \sum_j E(A_j) \right| - 0.5 \right)^2}{\sum_j \text{Var}(A_j)}, \quad E(A_j) = \frac{n_{Rj} m_{1j}}{T_j}$$

$$\text{Var}(A_j) = \frac{n_{Rj} n_{Fj} m_{1j} m_{0j}}{T_j^2 (T_j - 1)} \quad (2)$$

Yapılan işlemlerin ardından elde edilen ki- kare değeri, serbestlik derecesi ile birlikte ki-kare dağılımındaki tablo ile karşılaştırılır. Ortaya çıkan değer, kritik değerden büyük gelirse H_0 red edilir ve maddenin DMF içerdiğini sonucuna ulaşılır. Ancak hesaplamalar sonucunda ortaya çıkan oranların daha basit bir şekilde yorumlanabilmesi için değer, “delta ölçüğü (Δ)” adı verilen ölçek değerine dönüştürülür. Elde edilen değerler 0' dan büyük yani pozitif ise madde odak grup lehine, 0' dan küçükse negatif değer ile referans grubun lehine olarak yorumlanmaktadır. MH tekniği ile DMF' li maddenin düzeyini belirlemek amacıyla, Tablo 2.3.' de verilmiş olan,

ETS (The Educational Test Service) önerilen kategori düzeyleri kullanılmaktadır (Wiberg, 2007).

Tablo 2.3. *MH DMF İstatistiğinin Yorumlanması*

Düzyey	Değer	Açıklaması
A	$ \Delta_{MH} < 1$	İhmal edilebilir düzeyde DIF var.
B	$1 \leq \Delta_{MH} < 1,5$	Orta düzeyde DIF var.
C	$ \Delta_{MH} \geq 1,5$	Yüksek düzeyde DIF var.

Tabloda ifade edildiği gibi A düzeyindeki DIF içeren bir madde ihmal edilebilirdir ve madde testte kalmaya devam edebilir. B düzeyindeki madde ise, üzerinde çalışılarak düzeltme yapılmalı ya da ilgili madde yerine başka bir madde eklenebilir. C düzeyindeki madde ise ölçme aracından mutlaka çıkarılması gereken bir maddedir.

2.3.2. Lojistik Regresyon(LR)

KTK' ye dayalı DMF belirleme tekniklerinden biri olan LR, MH tekniğinin geliştirilmiş halidir ve hem tek biçimli hem de tek biçimli olmayan DMF belirlemede duyarlı bir tekniktir. Swaminathan ve Roger (1990) tarafından önerilen bu teknikte, maksimum olabilirlik teknikleriyle tahmin edilen olasılık fonksiyonu kullanılır (Osterlind, 1983). Teknik, iki bağımlı değişkeni tahmin etmek için bağımsız değişkenleri kullanan standart lojistik regresyon modeli üzerine kuruludur.

LR' de ilgili maddeye verilen tepkiler ve toplam puan üzerinden DMF' nin varlığı sınanır. Grup aidiyeti değişkeni ve test puanı ile aidiyet değişkeni arasında olan etkileşimin modele dahil edilmesi halinde model uyumu test edilerek belirlenir (Erdem Keklik, 2012). Teknik, bağımsız değişkenlerin dahil edilmesi ve etkileşimleri bakımından farklılık gösteren modelleri karşılaştırır. Eşitlik 3' de temel model ifade edilmiştir.

$$P(u_{ij} = 1) = \frac{e^z}{[1+e^z]} \quad (3)$$

Eşitlik 2' de yer alan u_{ij} , i bireyinin bir maddeye vermiş olduğu tepki; j ise i bireyinin hangi gruba ait olduğunu gösterir. LR tekniğinde üç model değerlendirilmektedir (Dainis, 2008). Model 1, Eşitlik 4' de verilmiştir.

$$z = T_0 + T_1\theta \quad (4)$$

Model 1' de bağımsız değişken olarak yalnızca yetenek düzeyi dahil edilir ve yokluk hipotezi modelidir. Maddenin DMF içermediğini gösterir. Model 1' e grup değişkeni dahil edilmesi halinde Eşitlik 5' te verilmiş olan Model 2 elde edilir.

$$z = T_0 + T_1\theta + T_2g \quad (5)$$

Model 2' deki “g” grup değişkenini temsil etmektedir. Bu değer 1' e eşit olması bireyin birinci grubun üyesi; 0' a eşit olması ise ikinci grubun üyesi olduğunu ifade eder. Eşitlikte yer alan T_2 , grubun madde üzerindeki performans farkıdır. Bu model yalnızca tek biçimli DMF' nin test edilmesine imkan verir. Son olarak teknikte kullanılan en karmaşık model olan Model 3, Eşitlik 6' te verilmiştir.

$$z = T_0 + T_1\theta + T_2g + T_3(\theta g) \quad (6)$$

Model 3' te yer alan T_3 , grup değişkeni ile yetenek düzeyi arasındaki etkileşimi göstermektedir. Burada belirtilmesi gereken önemli bir nokta, LR' de yetenek düzeyinin gözlenen ham puan olarak tanımlanmasıdır. Yetenek düzeyi (θ) ile grup üyeliği (g), T_3 ile ilişkili olan “ θg ” terimi ile verilir. Model 3 ile hem tek biçimli DMF hem de tek biçimli olmayan DMF test edilebilmektedir.

LR tekniğinde ilk olarak, etkileşim teriminin (T_3) anlamlı olup olmadığının belirlenmesi amacıyla Model 2 ve 3 arasında karşılaştırma yapılır, bu şekilde tek biçimli olmayan DMF belirlenir. T_3 anlamlı değilse, ana etkinin veya grup üyeliğinin önemini test etmek amacıyla Model 2, Model 1 ile karşılaştırılabilmektedir. Bu ana etki (T_2), tek biçimli DMF' yi temsil eder ve T_2 ' nin kendisi olan yorumlanabilir bir etki büyüklüğüne sahiptir. Bu terim, maddeyi doğru yanıtlama ve yeteneğin kontrol edilmesi olasılığı arasındaki logit fark olarak yorumlanmaktadır. Ek olarak, MH etki büyüklüğü ile karşılaştırılabilmesi için ETS delta ölçeğine dönüştürülebilir.

LR tekniği programlama kolaylığı ve tek biçimli olmayan DMF' yi ele alırken güçlü olması gibi nedenler ile yaygın olarak kullanılan bir tekniktir (Swaminathan ve Rogers, 1990; Clauser ve Mazor, 1998).

2.4. Madde Tepki Kuramına Dayalı Teknikler

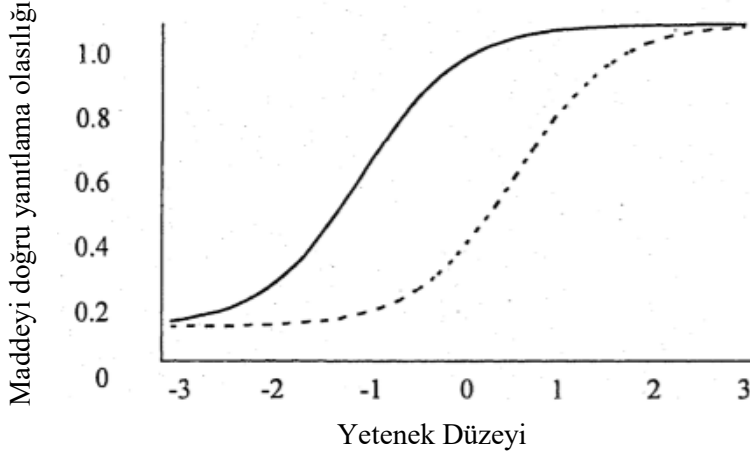
1930'lu yılların başından itibaren Klasik Test Kuramı'nın yetersiz ve sınırlı taraflarına çözüm olmak amacıyla, Madde Tepki Kuramı ortaya çıkmıştır. MTK, bireylerin yetenekleri doğrultusunda, bir maddeye verecekleri tepkiyi olasılık hesaplamalarıyla ortaya koyan bir kuramdır. Kuram ile birlikte, bireylerin yetenekleri maddelerden bağımsız şekilde; maddelerin psikometrik özellikleri ise bireylerden bağımsız bir şekilde kestirilebilmektedir (Embretson ve Reise, 2000; Lord, 2012).

MTK'ye dayalı DMF tanımı şu şekildedir; farklı gruplarda yer alan ancak aynı yetenek düzeyine sahip bireylerin, aynı maddeyi doğru yanıtlama olasılıklarının aynı olmamasıdır (Camilli ve Shepard, 1994). DMF' de bu gruplar referans ve odak grup olarak adlandırılmaktadır. Referans grubu maddeye karşı avantajlı olan bireylerden oluşurken, odak gruptaki bireyler ise dezavantajlıdır. Oluşturulan madde karakteristik eğrisi ile de bu gruplarda yer alan bireylerin yetenek düzeylerine göre maddeye doğru yanıt verme olasılıkları ortaya çıkmaktadır (Clauser ve Mazor, 1998).

MTK'ye dayalı DMF belirleme tekniklerinden bu araştırmada kullanılan Raju'nun Alan Ölçümleri ve Lord'un χ^2 'si teknikleri aşağıda ayrıntılı bir şekilde verilmiştir.

2.4.1. Raju' nun Alan Ölçümleri

Raju' nun alan ölçümleri, MTK altında gelişen ve örtük özellikler üzerine çalışan bir tekniktir. Raju (1988) madde yanlılığının ortaya konulmasına dönük çalışmalarda iki gruba ait madde karakteristik eğrileri (MKE)'nin kullanılabilmesini belirtmiştir. Bu teknikte, DMF' nin belirlenmesinde MKE' leri dikkate alınmaktadır. Raju' nun alan ölçümleri ifadesinde söz edilen alan, iki grubun kestirilen MKE arasında ortaya çıkan boşluktur (Camilli ve Shepard, 1994). Şekil 2.4.' de bir maddenin, iki farklı gruba ait olan MKE'sine yer verilmiştir.



Şekil 2.4. İki MKE Arasındaki Alan (Dainis, 2008)

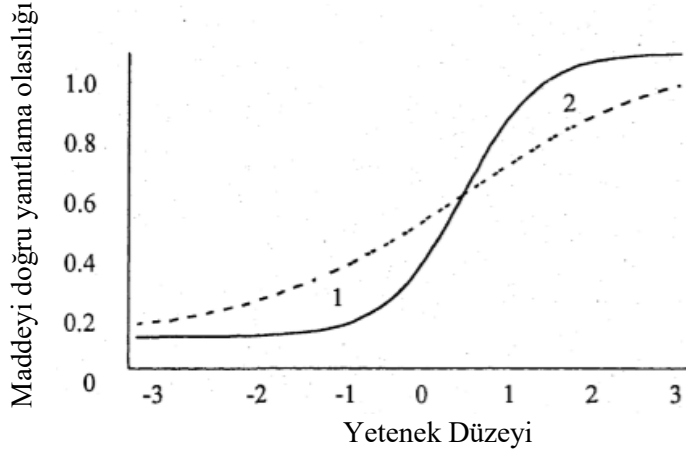
Şekil 2.4.' de yer alan MKE grafiği incelendiğinde, iki eğri arasında bir boşluk/alan olduğu görülmektedir. Eğer madde, iki gruba göre farklı işliyorsa, iki eğri arasında bu şekilde bir alan gözlemlenir. Bu da DMF' nin varlığına işarettir (Raju, 1988). Rudner (1977), bu alanın matematiksel bir şekilde hesaplanabilmesi için işaretli alan (signed area (SA)) olarak adlandırılan bir integral formülü önermiştir (Akt: Camilli ve Shepard, 1994). SA' nın formülü eşitlik 7'de verilmiştir.

$$SA = \int [P_R(\Theta) - P_O(\Theta)] d\Theta \quad (7)$$

P_R : Referans grup için olasılıkları,

P_O : Odak grup için olasılıkları temsil etmektedir.

Eşitlik 7' de de gösterildiği gibi, SA formülünde, odak grubun maddeyi doğru yanıtlama olasılığı ($P_O(\Theta)$), referans grubun maddeyi doğru yanıtlama olasılığından ($P_R(\Theta)$) çıkartılır. Bu yüzden, beklenmeyen bir şekilde referans grubun $P_R(\Theta)$ değeri daha yüksek çıkarsa, denklemin sonucu pozitif olur. Tersi bir durumda ise sonuç negatif olacaktır. Bu da performans farklılıklarının odak grubun lehine olması anlamına gelir. Ancak bilindiği üzere MKE' leri her zaman Şekil 2.5' deki gibi olmamaktadır. Şekil 2.5.' de daha farklı biçimde olan iki MKE gösterilmiştir.



Şekil 2.5. İki Farklı Grup İçin MKE'nin Grafiği (Dainis, 2008)

İki grup için kestirilen MKE Şekil 2.5.' deki gibi de olabilmektedir. Bu şekilde olduğunda Θ aralığının belirli bir noktasına kadar avantaj sağlayan grup sonrasında dezavantajlı olacaktır. Bu gerçeğe uygun olarak, geliştirilen SA formülünün uygulanması halinde MKE'lerin birbirini kesmesi durumunda, kesişen yerler birbirlerini sıfırlayacaktır. Bu da, DMF' nin olmadığı sonucunu üretebilmektedir. Bu durumun ortadan kaldırılması amacıyla farklı bir hesaplama tekniği olan, UA değerinin karesi alınarak hesaplanan işaretli alan ölçümleri (signed-area) (UA) ortaya çıkmıştır. UA'nın formülü eşitlik 8' de verilmiştir.

$$UA = \int (P_R(\Theta) - P_O(\Theta))^2 d\Theta \quad (8)$$

Her zaman pozitif değerler veren işaretli alan ölçümlerinde, Şekil 2.5.' deki gibi durumlarda ortaya çıkan iki diferansiyel performansın toplanması mümkün hale gelir. UA değerinin, SA değerinden büyük ölçüde farklı olması ise MKE'nin kesiştiğinin göstergesidir.

2.4.2. Lord' un χ^2 'si (Ki-Kare) Tekniği

Lord (1982) tarafından önerilmiş olan bu teknikte, iki farklı grubun bir maddeye ilişkin madde parametrelerinin farklılıkları karşılaştırılmaktadır (Akt: Price, 2014). Yapılan karşılaştırma sonucunda aralarında fark bulunması halinde, ilgili maddenin farklı biçimde bir işleyiş gösterdiği söylenebilir. Daha genel bir ifadeyle, aynı yetenek düzeyinde farklı gruplarda yer alan bireylerin, maddeyi doğru yanıtlama olasılıkları farklılaşıyorsa, maddenin DMF gösterdiği söylenmektedir (Kim ve Cohen, 1994).

Uygun MTK modellerinin kullanılması halinde, hem tek biçimli DMF hem de tek biçimli olmayan DMF' nin belirlenmesinin mümkün olması tekniğin avantajlarından (Price, 2014).

Lord' un χ^2 ' sinde her iki grup için kestirilmiş olan a ve b parametreleri arasındaki farklar hesaplanmakta ve bu farklar arasındaki varyans –kovaryans matrisleri kullanılmaktadır. Odak grup ve referans grup arasında yapılan bu karşılaştırma farkın istatistiksel olarak manidar olup olmadığını görmek amacıyla yapılır. Tekniğe ilişkin formül, eşitlik 9' da verilmiştir (Camilli ve Shepard, 1994).

$$V' = (a_O - a_R, b_O - b_R) \quad (9)$$

$$X^2(2) = V' S^{-1} V$$

S: Madde parametreleri arasında olan farkın Varyans –Kovaryans matrisi.

Analiz sonucunda elde edilen değerlerin, kritik değerle karşılaştırılmasıyla DMF olup olmadığı tespit edilir (Osterlind, 1983).

Bu araştırmada daha öncede belirtildiği şekilde yukarıda açıklanan tekniklerin çeşitli koşullar altındaki I.Tip hata ve istatistiksel güç oranlarının incelenmesi planlanmıştır. I. Tip hataya, DMF bağlamında bakıldığında, DMF' nin yanlış tanımlanması söz konusudur. Madde, gerçekte DMF içermiyor olmasına rağmen, istatistiksel olarak anlamlı bir şekilde DMF içeriyor şeklinde karar verilir. Daha farklı bir ifade ile, ilgili madde odak ve referans gruptaki bireyler arasında benzer performans göstermesine rağmen, kullanılan teknik, maddenin değişen madde fonksiyonu gösterdiğine işaret eder (Dainis, 2008). Bu problem araştırmalara bazı nedenler ile konu olmaktadır (Jodoin ve Gierl, 2001). İlk neden olarak maddenin DMF içeriyor şeklinde tespit edilmesinin zaman ve kaynak kaybına sebep olması söylenebilir. Eğer bu DMF tespiti yanlış ise test geliştirme süreci ile kaynaklar israf edilmiş olur. İkinci olarak, yine yanlış tanımlanması halinde, araştırmacılar gerek olmadığı halde kaynaklarını tüketerek yeniden DMF analizleri yapmak için vakit harcarlar. Son olarak ise araştırma açısından bakıldığında, çalışmanın amacı DMF belirleme tekniklerinin gücünü belirlemek ya da gerçek veri üzerinden DMF' li maddelerin tespiti ise aslında I. Tip hatalar nedeniyle karşılaştırmalardan yapılan çıkarımlar (teknik kalitesi ve etkinliği açısından) geçerli olmayacaktır.

I. Tip hatanın düzeylerine ilişkin literatürde çeşitli öneriler mevcuttur. Bu araştırmada, Bradley (1978)' in ortaya koymuş olduğu tutucu (conservative), korunan (maintained) ve şişirilmiş (inflated) şeklindeki sınıflandırması dikkate alınacaktır. Bradley' in sınıflandırmasında sınıflandırılmış nominal α değerleri Tablo 2.4.'de verilmiştir (Akt: Price, 2014).

Tablo 2.4. *Bradley' in I. Tip Hata Oranlarına İlişkin Sınıflandırması*

Düzyey	Değer Aralığı
<i>Tutucu (conservative)</i>	$\alpha < 0.045$
<i>Korunan (maintained)</i>	$0.045 < \alpha < 0.055$
<i>Şişirilmiş (inflated)</i>	$\alpha > 0.055$

Tablo 2.4. incelendiğinde, α değerinin 0.045 ile 0.055 arasında olduğu korunan (maintained) düzeydeki I. Tip hata, doğru istatistiksel sonuçlar çıkarabilmemiz için gerekli olan düzeydir. Korunan düzeyde olmayan bir α değeri yanlış sonuçlar elde etmemize sebep olmaktadır. Örnek olarak, tutucu I. Tip hata gösteren bir tekniğin gücü azalmakta ve şişirilmiş I. Tip hata oranı artmaktadır. Diğer bir deyişle H_0 hipotezi daha sık reddedilmektedir. Ki bu da istenmeyen I. Tip hata kavramının bir başka tanımı olan, doğru olan H_0 hipotezinin red edilerek, yanlış kararlar alınması demektir. I. Tip hata oranının etkilediği bir oran olan istatistiksel güç ise DMF içeren maddenin, analiz sonucunda verilen doğru karar ile DMF içeriyor olarak belirlenmesidir. Belirli koşullar altında gücü incelemek, deneysel yanlılık araştırmasının ilk adımıdır ve bu oranın yüksek olması beklenir (Dainis, 2008). Bu sebeple araştırmada farklı DMF belirleme tekniklerin I. Tip hata oranları ve gücü çeşitli koşullar altında incelenmiştir.

2.5. İlgili Araştırmalar

Araştırmanın bu bölümünde, literatürdeki DMF belirleme tekniklerinin performanslarının ortaya koyulması amacıyla yapılan çalışmalara yer verilmiştir.

Sünbül ve Ömür Sünbül (2016) çalışmalarında, DMF belirlemek amacıyla kullanılan tekniklerin belirli koşullarda, I. Tip hata ve güç oranlarını ortaya koymayı amaçlamıştır. Simülasyon veriler üzerinden yürütülen bu çalışmada, MH, LR, Lord' un χ^2 si ve Raju' nun alan ölçümleri teknikleri kullanılmıştır. Çalışmanın sonucunda, KTK' ye dayalı tekniklerde I. Tip hata oranları ortalamalarını; madde sayısının, örneklem büyüklüğünün ve referans gruptaki örneklem büyüklüğünün toplam örneklem büyüklüğüne oranının değişimlenmesinin çok değiştirmediğine ulaşılmıştır. MTK' ye dayalı tekniklerde ise örneklem büyüklüğünün artması ile I. Tip hata oranları ortalamasının azaldığı; madde sayısının artması ile herhangi bir değişim olmadığı ve referans ile odak gruplarında örneklem büyüklüğü oranlarının eşit olması halinde I. Tip hata oranlarının ortalamasının daha da düşük olduğu bulgularına ulaşılmıştır.

Çalışmanın güç oranlarına ilişkin sonuçlarında ise, her iki kurama dayalı tekniklerinde örneklem büyüklüğünün ve DMF miktarının artırılması ile güç oranları ortalamasının da arttığına ulaşılmıştır. Yine tüm teknikler için değişimlenen madde sayısının, güç oranları ortalamasını çok fazla değiştirmedigine, ek olarak da DMF' li madde sayısı oranının artırılmasının ve referans ile odak gruplardaki örneklem büyüklüğü oranlarının farklı olması halinde güç oranlarının ortalamasını azalttığı bilgisine ulaşılmıştır.

Kabasakal, Arsan, Gök ve Kelecioğlu (2014) belirli koşullar altında MH, SIBTEST ve MTK – OO tekniklerinin performanslarını, daha açık bir ifadeyle I. Tip hata ve güçlerini karşılaştırmıştır. Simülasyon veri ile yürütülen çalışmada, yetenek dağılımı farklılıkları, DMF içeren madde oranı, örneklem büyüklüğü oranı, test uzunluğu ve model türü değişkenlerinin DMF üzerine etkisine bakılmıştır. Çalışmadan elde edilen sonuçlar ile tek biçimli DMF' nin belirlenmesinde, ele alınan tüm koşullarda MH tekniğinin en yüksek güce, SIBTEST tekniğinin ise en yüksek hataya sahip olduğu ortaya çıkmıştır.

E. Keklik (2012), MH, LR ve MTK –OO tekniklerinin iki kategorili veriler üzerindeki I. Tip hata oranları ve istatistiksel güç değerlerinin çeşitli koşullar altında nasıl değiştiğini ortaya koymayı amaçlamıştır. Bu koşullar referans ve odak grubun yetenek dağılımı, örneklem büyüklüğü ile örneklem büyüklüğü oranlarının değiştiği çeşitli koşullardır. Çalışma, simülasyon veriler üzerinden yürütülmüştür. Sonucunda ise, grupların yetenek dağılımları normallik gösterdiğinde LR ve MH tekniklerinde benzer; MTK –OO tekniğine kıyasla da daha düşük I. Tip hata oranlarının olduğu ortaya çıkmıştır. Ancak grupların yetenek dağılımlarında farklılaşma olduğunda MTK –OO tekniğinin diğer tekniklere kıyas ile daha iyi I. Tip hata kontrolü gerçekleştirdiği ortaya çıkmıştır.

Magis ve De Boeck (2012) tarafından yapılan simülasyon çalışmasında farklı koşullar altında MH tekniğinin performansının incelenmesi amaçlanmıştır. Bu doğrultuda yapılan analizlerin ardından MH tekniğinin gruplar arası yetenek farkı olduğu durumlarda ve büyük örneklem büyüklüklerinde şişirilmiş I. Tip hata oranları verdiğine ulaşılmıştır.

Vaughn ve Wang, 2010 yılında yaptıkları çalışmada ise sınıflama ağacı, MH ve LR tekniklerinin, örneklem büyüklüğü, DMF içeren madde oranı ve yetenek dağılımı koşullarında I. Tip hata oranlarını tespit etmeyi amaçlamışlardır. Hem simülasyon hem de gerçek veri

üzerinden yürütülen bu çalışmanın sonucunda, kullanılan üç farklı örneklem büyüklüğü ve oranları için MH ve LR tekniklerinin düşük I. Tip hata oranlarına sahip olduğuna ulaşılmıştır.

Demars (2009) çalışmasında, DMF belirlemede kullanılan MH, LR ve SIBTEST tekniklerinin çeşitli koşullar altındaki I. Tip hata oranlarının karşılaştırmasını amaçlamıştır. Çalışmasını, test uzunluğu, örneklem büyüklüğü ve yetenek dağılımları koşullarını kullanarak simülasyon veri ile yürütmüştür. Yapılan bu araştırma sonucunda, test uzunluğunun kısılması ve örneklem büyüklüğünün artması ile MH ve LR tekniklerinin şişirilmiş I. Tip hata oranları verdiğine ulaşılmıştır.

Dainis (2008) çalışmasını hem gerçek hem de simülasyon verileri kullanarak dört farklı DMF belirleme tekniğinin performanslarını karşılaştırmıştır. MTK- OO, DFIT, MH ve LR bu araştırmaya dahil edilen tekniklerdir. Çalışmada manipüle edilen koşullar örneklem büyüklüğü, etkinin varlığı ve yokluğu, TB ve TBO Değişen Madde Fonksiyonu durumları ile DMF' nin oranıdır. Çalışmanın sonucunda, tüm simülasyon koşullarında DFIT tekniğinin düşük I. Tip hata oranları vermiştir, MH tekniği ise tüm koşullar altında kabul edilebilir düzeyde I. Tip hata oranları vermiştir. Ancak MH tekniğinin istatistiksel gücü TBO değişen madde fonksiyonunda düşük çıkmıştır. DFIT ve MTK- OO teknikleri yüksek istatistiksel güç oranları verirken LR tekniğinin madde etkisi koşulunda düşük güç oranları verdiği rapor edilmiştir.

Atar (2007) çalışmasında, diğer çalışmalardan farklı olarak iki kategorili ve çok kategorili puanlanan maddelerin olduğu bir testin yanı sıra karışık puanlanan maddelerden oluşan bir testteki maddelerinde DMF içerip içermediğinin incelenmesini ve kullanılan tekniklerin performansını ortaya koymayı amaçlamıştır. Bu amaç ile çalışmasını LR, MTK- OO ve GLLAMM tekniklerini kullanmıştır. Simülasyon verileri ile yürütülen bu çalışmada örneklem büyüklüğü, örneklem büyüklüğü oranları ve DMF içeren madde oranları koşulları değişimlenmiştir. Çalışma ile örneklem büyüklüğünden LR ve MTK- OO tekniklerinin etkilendiğini ve yine bu tekniklerin istatistiksel güç oranlarının daha yüksek olduğuna ulaşılmıştır. Çalışmanın daha genel bir sonucu olarak ise örneklem büyüklüklerinin büyümesinin hem istatistiksel güç oranlarını hem de I. Tip hata oranlarını arttırdığını ifade etmiştir.

Jodoin ve Gierl (2001) simülasyon çalışması ile LR tekniği için geliştirmiş oldukları sınıflandırmanın I. Tip hata ve istatistiksel güç oranlarını etkileyip etkilenmediğinin belirlenmesini amaçlamışlardır. Çalışmalarında etki büyüklüğü, örneklem büyüklüğü, DMF

içeren madde oranı ve grupların yetenek dağılımı değişkenlerini kullanarak araştırmalarını yürütmüşlerdir. Çalışma sonucunda, DMF içeren madde oranının 0.10 olduğu koşul altında örneklem büyüklüğü arttıkça, tekniğin istatistiksel güç oranlarının da arttığına ulaşılmıştır. DMF içeren madde oranı 0.20' ye çıktığında ise istatistiksel güç oranları düşmüştür. Ek olarak, referans grubun 1000; odak grubun 250 örneklem büyüklüğüne sahip olduğu koşulda tekniğin istatistiksel güç oranının beklenmeyen bir biçimde düşük çıktığı sonucuna ulaşılmıştır.

Gierl, Jodoin ve Ackerman (2000) tarafından yürütülen, MH, LR ile SIBTEST tekniklerinin, DMF içeren madde oranı, örneklem büyüklüğü ve yetenek dağılımı olmak üzere farklı koşullar altındaki I. Tip hata ve istatistiksel güç oranlarının incelendiği bir simülasyon çalışmasıdır. Çalışmanın sonucunda, örneklem büyüklüklerinin küçük olduğu durumlarda dahi üç tekniğinde I. Tip hata oranlarının nominal alfa düzeyi çevresinde olduğu belirtilirken; SIBTEST tekniğinin de en yüksek istatistiksel güç oranlarına sahip olduğu belirtilmiştir.

Ankenmann, Witt ve Dunbar (1999) çalışmasında DMF belirleme tekniklerinden MH ve LR tekniklerin I. Tip hata oranları ve istatistiksel güç oranlarının karşılaştırılmasını amaçlamışlardır. Çalışmada, örneklem büyüklüğü ile yetenek dağılımının farklı ve aynı olduğu koşullar altında tekniklerin performansları incelenmiştir. Çalışmanın sonucunda, MH tekniğinin, LR tekniğine göre istatistiksel güç oranlarının daha iyi olduğuna ancak her iki tekniğinde örneklem büyüklüklerinden etkilendiğine ulaşılmıştır. I. Tip hata oranları açısından bakıldığında ise iki tekniğinde genel olarak nominal alfa düzeyinin çevresinde olmasına karşın hemen hemen tüm örneklem büyüklükleri için MH tekniğinin nominal alfa düzeyini aştığını belirtmişlerdir.

Cohen, Kim ve Wollack tarafından 1996 yılında yürütülen araştırmada, Monte- Carlo simülasyonu kullanılarak DMF belirleme tekniklerinden MTK- OO tekniğinin I. Tip hata oranları incelenmiştir. İlgili araştırmada 2PL ve 3PL model kullanılarak, test uzunluğunun 50 ve örneklem büyüklüklerinin 250 ile 1000 olduğu koşullar altında analizler gerçekleştirilmiştir. Araştırmanın sonucunda, 2PL model için DMF belirleme tekniklerinin I. Tip hata oranlarının nominal alfa düzeyinde olduğunu ve 3PL modelde bu oranların nominal düzeyi aştığı bulgusuna ulaşılmıştır.

III. YÖNTEM

Araştırmanın bu bölümünde sırası ile araştırmanın modeli, simülasyon deseni, verilerin üretimi ve verilerin analizi başlıkları yer almaktadır.

3.1. Araştırmanın Modeli

Araştırmada, belirlenen koşullar altında 2PL modele uygun veriler üretilmiştir. Üretilen veriler ile DMF belirleme tekniklerinin I. Tip hata ve istatistiksel güç oranları incelenmiştir. DMF belirleme tekniklerinden MH, LR, Lord' un χ^2 ' si ve Raju' nun Alan Ölçümleri tekniklerinin performansları hakkında bilgi vererek, literatürde varolan önceki bilgilere katkı sağlayacak bir araştırma olması nedeniyle, araştırmanın modeli temel araştırmadır.

3.2. Simülasyon Deseni

Araştırmada, odak ve referans grup için cevaplayıcı tepkileri, yetenek düzeyleri ve madde parametreleri 2PL lojistik modele uygun olarak üretilmiştir.

Bu başlık altında I. Tip hata ve istatistik güç oranlarının farklılaşıp farklılaşmadığının ortaya konulması amacıyla verilerin hangi koşullar altında üretildiği ve bu koşulların sabit tutulan ile manipüle edilenleri açıklanmıştır.

3.2.1. Sabit Koşullar

3.2.1.1. İki Parametrelili Madde Tepki Modeli

Bu araştırmada, odak ve referans grup için cevaplayıcı tepkileri, yetenek düzeyleri ve madde parametreleri 2PL lojistik modele uygun olarak üretilmiştir. 2PL modelin veri uyumunun 1PL modele kıyas ile daha iyi olması sebebiyle, 1PL model araştırmaya dahil edilmemiştir. Yine 3PL modelde, gerçek test koşulları altında incelenmesi halinde c parametresi sebebiyle kestirimlerde standart hatanın büyüdüğü ve DMF belirleme çalışmalarında güçlü bir istatistik olmadığı belirtilmiştir (Hambleton, Swaminathan ve Rogers, 1991). Bu nedenler göz önünde bulundurularak araştırmaya yalnızca 2PL model dahil edilmiştir.

3.2.1.2. Odak ve Referans Grubun Yetenek Dağılımları

Araştırmada, odak ve referans grubun yetenek düzeyleri(Θ) sabit ve aynı tutulmuştur. Karşılaştırılan iki grubunda yetenek dağılımı aynıdır ve ortalaması 0, standart sapması 1 olan birim normal dağılım göstermektedir (Dedeen, 2004; Hauck Filho, Machado ve Damásio, 2014; Roussos ve Stout, 1996).

3.2.1.3. DMF Türü

Araştırmada verilerin üretilmesi aşamasında her iki grup içinde a parametreleri eşit tutulurken; referans grubun b parametresi değiştirilmiştir. İki gruba ait parametreler bu şekilde olduğunda ortaya tek biçimli DMF çıkmaktadır (Erdem Keklik, 2014). Bu araştırmada DMF türü olarak tek biçimli DMF kullanılmıştır.

3.2.2. Manipüle Edilen Koşullar

3.2.2.1. Örneklem Büyüklüğü ve Örneklem Büyüklüğü Oranları

Swaminathan ve Rogers(1990), yapılan çalışmalarda istatistiksel kestirimi etkileyebilecek faktörlerden birinin de odak ve referans grubun örneklem büyüklüğü ve örneklem büyüklüğü oranları olduğunu belirtmiştir. Parametrik olmayan tekniklerin, grupların örneklem sayısı eşit olmadığı durumlarda, DMF' li maddeyi belirleme gücü artış göstermektedir (Kristjansson, Aylesworth, McDowell ve Zumbo, 2005). Literatür incelendiğinde, KTK' ye dayalı tekniklerde grup başına en az 200- 250 bireyin olması halinde tekniklerin yeterli güce ulaştıkları bulgusu elde edilmiştir (Narayanan ve Swaminathan, 1996; Rogers ve Swaminathan, 1993). MTK' ye dayalı tekniklerde ise en az 1000 kişilik örneklemelere ihtiyaç duyulmaktadır (Shepard, Camilli ve Averill, 1981). Bu araştırmada iki farklı kurama ait teknikler incelendiği için ancak bu kuramlardan biri büyük örneklem gerektiren MTK olduğundan, araştırmada küçük örneklem tercih edilmemiştir. Örneklem büyüklükleri $N_1= 1800$ ve $N_2=3000$, referans grubun örneklem büyüklüğünün; toplam örneklem büyüklüğüne oranı ise $R/T_1= 0.50$ ve $R/T_2= 0.75$ olarak manipüle edilmiştir.

3.2.2.2. Test Uzunluğu

DMF kapsamında incelenen bir diğer faktör ise test uzunluğu, diğer bir deyişle madde sayısıdır. Literatürde yapılan diğer çalışmalar incelendiğinde, madde sayıları genel olarak 20, 40 ve 80

olarak alınmıştır (Kabasakal ve diğeri, 2014, Price, 2014, Wang, Tay ve Drasgow, 2013, Narayanan ve Swaminathan, 1994). Test uzunlukları için 20 küçük, 40 orta ve 80 büyük test uzunluğu olarak ifade edilebilmektedir. Türkiye’ de yapılan ulusal sınavlar göz önünde bulundurulduğunda ise LGS’ de 10- 20 madde; ALES, YDS ve YÖKDİL gibi sınavların ise 50-80 maddelik testlerden oluştuğu bilinmektedir. Türkiye’ de yapılan geniş ölçekli sınavlar göz önünde bulundurularak, araştırmada 20 ve 80 olmak üzere iki farklı test uzunluğu belirlenmiştir.

3.2.2.3. DMF İçeren Madde Oranı

DMF içeren maddelerin testte bulunmasının, ilgili testte, madde yanlılığının varlığına işaret olduğu bilinmektedir ve elbette ki bu istenmeyen bir durumdur. DMF’ nin varlığı, testlerin geçerliğini düşürür ve DMF belirlemek amacıyla kullanılan tekniklerin gücü üzerinde etkileri vardır (Jodoin ve Gierl, 2001). Ancak gerçek veriler kullanılarak yapılan bir çalışmada analizler yapılmadan önce DMF gösteren maddenin olup olmadığı bilinemez ve hangi tekniğin DMF içeren madde oranlarından etkilendiği sınınamaz. Bu amaçla araştırmada, DMF içeren madde oranının 0, 0.05 ve 0.10 olduğu koşullarda, kullanılan tekniklerde nasıl sonuçlar verdiği incelenmiştir.

3.3. Verilerin Üretimi

Araştırmada, farklı kuramlara dayanan DMF belirleme tekniklerinin belirli koşullar altında I. Tip hata ve istatistiksel güç oranlarının belirlenmesi simülasyon veriler ile çalışılarak yapılmıştır. Kullanılan simülasyon veriler, iki kategorili ve 2PL madde tepki modeline uygun olarak R programında I. Tip hata ve istatistiksel güç oranları için ayrı ayrı üretilmiştir.

Belirli koşullar altında farklı DMF belirleme tekniklerinin I. Tip hata oranlarının belirlenmesi amacıyla, örneklem büyüklüğü $N_1= 1800$ ve $N_2=3000$; örneklem büyüklüğü oranları $R/T_1= 0.50$ ve $R/T_2= 0.75$; test uzunluğu $TU_1= 20$ ve $TU_2= 80$ olacak şekilde değişimlenerek veriler üretilmiştir. Üretilen maddeler her iki grup için de, “a” parametresi ortalaması (\bar{x}) 0.8 ve standart sapması (SS) 0.04 olan bir normal dağılım üzerinden elde edilirken (Sünbül ve Ömür Sünbül, 2016); “b” parametresi minimum ve maksimum değerleri sırası ile -2 ve +2 olan bir dağılımdan seçkisiz biçimde çekilerek elde edilmiştir (Desa, 2012; Koğar, 2018). Bireylerin yetenek dağılımına dair değerler ise $\bar{x}=0$ ve $SS=1$ olan bir normal dağılım üzerinden elde edilmiştir.

Belirli koşullar altında farklı DMF belirleme tekniklerinin istatistiksel güç oranlarının belirlenmesi amacıyla ise örneklem büyüklüğü $N_1= 1800$ ve $N_2=3000$; örneklem büyüklüğü oranları $R/T_1= 0.50$ ve $R/T_2= 0.75$; test uzunluğu $TU_1= 20$ ve $TU_2= 80$ olan koşullara ek olarak DMF içeren madde oranı için de odak grubun belirli sayıdaki maddelerinin b parametrelerine 0.50 eklenmiştir. İstatistiksel güç oranları için üretilen verilerde, odak ve referans gruba ait a parametreleri yine $\bar{x}=0.8$ ve $SS= 0.04$ olan bir normal dağılım üzerinden; “b” parametresi minimum ve maksimum değerleri sırası ile -2 ve +2 olan bir dağılımdan seçkisiz biçimde çekilerek elde edilmiştir. Ancak DMF içeren maddelerin üretimi, odak grubun veri üretimindeki b parametrelerinin, referans grubun b parametresi üzerinde yapılan manipülasyonlar ile elde edilmiştir. Daha açık bir ifade ile DMF içeren madde oranı koşulundaki değerler göz önünde bulundurularak, DMF’li olması gereken madde sayısınca b parametresine 0.05 ve 0.10 değerleri eklenmiştir. Yapılan bu işlem dışında odak ve referans grubun yetenek düzeyleri ve a parametreleri denk iken b parametresi değiştirildiği için tek biçimli (uniform) DMF gösteren veriler elde edilmiştir.

Tablo 3.1.’ de araştırma kapsamında oluşturulan manipüle edilen koşullar verilmiştir.

Tablo 3.1. *Araştırmanın manipüle edilen koşulları*

Örneklem Büyüklüğü	Test Uzunluğu	Örneklem Büyüklüğü Oranı (R/T)	DMF İçeren Madde Oranı
1800	20	900/ 900 (0.50)	0.05
			0.10
		1350/ 450 (0.75)	0.05
		0.10	
	80	900/ 900 (0.50)	0.05
			0.10
1350/ 450 (0.75)		0.05	
		0.10	
3000	20	1500/ 1500 (0.50)	0.05
			0.10
		2250/ 750 (0.75)	0.05
		0.10	
	80	1500/ 1500 (0.50)	0.05
			0.10
2250/ 750 (0.75)		0.05	
		0.10	

DMF analizleri için iki farklı örneklem büyüklüğü, iki farklı örneklem büyüklüğü oranı, iki farklı test uzunluğu ve üç farklı DMF içeren madde oranına göre üretilen (2x2x2x3) toplam 24 veri seti için 25 tekrarlı yapılmıştır.

3.4. Verilerin Analizi

DMF belirlemek amacıyla geliştirilmiş olan pek çok teknik mevcuttur. Bu araştırmada iki kategorili veriler için KTK'ye dayalı Lojistik Regresyon(LR) ve Mantel- Haenszel(MH) tekniği ile MTK' ye dayalı Lord' un χ^2 'si ve Raju' nun alan ölçümleri kullanılmıştır. Verilerin analizi, R programında yer alan ve DMF belirleme tekniklerini içerisinde barındıran; bu tekniklerin hem tek başlarına hem de karşılaştırmalı olarak analiz edilmesine olanak sağlayan “difR” paketi kullanılarak yapılmıştır (Magis, Beland ve Raiche, 2018). MTK' ye dayalı teknikler için yapılan analizlerde, madde tepki modellerine uygun model kestirimlerine olanak sağlayan bir paket olan “ltm” paketi de kullanılarak; 2PL madde tepki modeline uygun kestirimler yapılmıştır (Rizopoulos,2018).

Araştırmada DMF belirleme teknikleri, I. Tip hata ve istatistiksel güç oranları olmak üzere iki ölçüt ile değerlendirilmiştir. Analizlerin ardından I. Tip hata için DMF içermezken, DMF içeriyor olarak analiz çıktısı veren maddelerin oranı hesaplanarak raporlaştırılmıştır. İstatistiksel güç oranları için ise madde, DMF içeriyorken, DMF içerdiğine dair analiz çıktısı elde edilmiş olan maddelerin oranı raporlaştırılmıştır. Raporlaştırma esnasında dikkate alınan ölçütler, MH, LR, Lord' un χ^2 ' si ve Raju' nun Alan Ölçümleri tekniklerinin I. Tip hata oranları için, Bradley(1978)'in sınıflandırması dikkate alınarak nominal alta düzeyi olan 0.05 düzeyinde ya da daha altında olduğunda, I. Tip hata oranlarının istenilen düzeyde olduğu; 0.05' den yüksek olduğu durumlarda ise şişirilmiş I. Tip hata olarak kabul edilmiştir (Akt: Price, 2014). İstatistiksel güç oranları için ise 0.80' in üstündeki tekniklerin gücünün yeterli ve yüksek olduğu; altında olanların ise yetersiz olduğu ölçütü dikkate alınmıştır (Kabasakal ve diğerleri, 2014).

IV. BULGULAR

Bu bölümde araştırmanın amaçları doğrultusunda elde edilen bulgular “I. Tip Hata Oranlarına İlişkin Bulgular” ve “İstatistiksel Güç Oranlarına İlişkin Bulgular” olmak üzere iki alt başlık altında verilmiştir.

4.1. I. Tip Hata Oranlarına İlişkin Bulgular

Araştırmanın bu bölümünde, örneklem büyüklüğü $N_1=1800$ ve $N_2=3000$; örneklem büyüklüğü oranları $R/T_1=0.50$ ve $R/T_2=0.75$; test uzunluğu $TU_1=20$ ve $TU_2=80$ olacak şekilde değişimlenerek oluşturulan koşullara dair bulgular verilmiştir.

4.1.1. 1. Alt Probleme İlişkin Bulgular

“Örneklem büyüklüğünün 1800; örneklem büyüklüğü oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF içeren madde oranının 0 olduğu koşullar altında MH, LR, Raju’ nun Alan Ölçümleri ve Lord’ un χ^2 si tekniklerinin 1. Tip hata oranları nasıl değişmektedir?” alt problemine ilişkin bulgular Tablo 4.1.’ de verilmiştir.

Tablo 4.1. MH, LR, Raju’ nun Alan Ölçümleri ve Lord’ un χ^2 si Tekniklerinin 1800 Örneklem Büyüklüğü için I. Tip Hata Oranları

Örneklem Büyüklüğü	Madde Sayısı	R/T	MH	LR	Lord χ^2	Raju
1800	20	0.50	0.924	0.918	0.864	0.740
		0.75	0.902	0.888	0.788	0.678
	80	0.50	0.887	0.880	0.880	0.793
		0.75	0.861	0.860	0.862	0.791

Tablo 4.1. incelendiğinde 1800 örneklem büyüklüğü ve 20 madde için farklı koşullar altında I. Tip hata oranı değerlerinin, 0.678 ile 0.924 arasında değiştiği görülmektedir. Örneklem büyüklüğü oranları 0.50 olduğunda α değerleri MH tekniği ile 0.924 ile en yüksek; Raju’ nun alan ölçümleri tekniği ile ise 0.74 ile en düşük hatalı olarak bulunmuştur. Örneklem büyüklüğü oranı 0.75 iken yine aynı şekilde 0.902 ile en yüksek α değeri MH tekniğinde; en düşük α değeri ise 0.678 ile Raju’ nun alan ölçümleri tekniğinde çıkmıştır.

Örneklem büyüklüğünü 1800 ve madde sayısının 80 olarak alındığı analizlerin sonucunda ise α değeri 0.791 ile 0.887 arasında değişmektedir. Analizde, örneklem büyüklüğü oranları değiştikçe α değerleri yine değişiklik göstermektedir. Örneklem büyüklüğü oranının 0.50 olduğu koşulda, 0.887 α değeri ile MH tekniği en yüksek hataya; 0.793 α değeri ile de Raju' nun alan ölçümleri tekniği en düşük hataya sahip olarak çıkmıştır. Örneklem büyüklüğü oranı 0.75 olarak alındığında ise Lord χ^2 ' si tekniği 0.862 ile en yüksek; Raju' nun alan ölçümleri de 0.791 α değeri ile en düşük hataya sahip olarak bulunmuştur.

4.1.2. 2. Alt Probleme İlişkin Bulgular

“Örneklem büyüklüğünün 3000; örneklem büyüklüğü oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF içeren madde oranının 0 olduğu koşullar altında MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si tekniklerinin I. Tip hata oranları nasıl değişmektedir?” alt probleminde ilişkin bulgular Tablo 4.2.' de verilmiştir.

Tablo 4.2. MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 si Tekniklerinin 3000 Örneklem Büyüklüğü için I. Tip Hata Oranları

Örneklem Büyüklüğü	Madde Sayısı	R/T	MH	LR	Lord χ^2	Raju
3000	20	0.50	0.924	0.928	0.888	0.778
		0.75	0.886	0.888	0.856	0.768
	80	0.50	0.916	0.913	0.901	0.838
		0.75	0.908	0.906	0.905	0.851

Tablo 4.2. incelendiğinde 3000 örneklem büyüklüğü ve 20 madde için örneklem büyüklüğü oranının değiştiği koşullar altında α değerinin 0.768 ile 0.928 arasında değiştiği görülmektedir. Örneklem büyüklüğü oranları 0.50 olduğunda α değerleri LR tekniğinde 0.928 ile en yüksek; Raju' nun alan ölçümleri tekniğinde ise 0.778 ile en düşük hata içeriyor şeklinde bulunmuştur. Örneklem büyüklüğü oranı 0.75 iken yine aynı şekilde 0.888 ile en yüksek α değeri LR tekniğinde; en düşük α değeri ise 0.768 ile Raju' nun alan ölçümleri tekniğinde bulunmuştur.

Örneklem büyüklüğünü 3000 ve madde sayısının 80 olarak alındığı analizlerin sonucunda ise α değeri 0.838 ile 0.916 arasında değişmektedir. Analizde, örneklem büyüklüğü oranları değiştikçe α değerleri değişiklik göstermektedir. Örneklem büyüklüğü oranının 0.50 olduğu

koşulda, 0.916 α değeri ile MH tekniği en yüksek hataya; 0.838 α değeri ile de Raju' nun alan ölçümleri tekniği en düşük hataya sahip olarak çıkmıştır. Örneklem büyüklüğü oranı 0.75 olarak alındığında ise MH tekniği 0.908 ile en yüksek; Raju' nun alan ölçümleri tekniği ise 0.851 α değeri ile en düşük hataya sahip olarak bulunmuştur.

4.2. İstatistiksel Güç Oranlarına İlişkin Bulgular

Araştırmanın bu bölümünde, örneklem büyüklüğü $N_1= 1800$ ve $N_2=3000$; örneklem büyüklüğü oranları $R/T_1= 0.50$ ve $R/T_2= 0,75$; test uzunluğu $TU_1= 20$ ve $TU_2= 80$; DMF içeren madde oranının 0.05 ve 0.10 olacak şekilde değiştirilerek oluşturulan koşullara dair bulgular verilmiştir.

4.2.1. 3. Alt Probleme İlişkin Bulgular

“Örneklem büyüklüğünün 1800; örneklem büyüklüğü oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF içeren madde oranının 0.05 ve 0.1 olduğu koşullar altında MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 ' si tekniklerinin istatistiksel güç oranları nasıl değişmektedir?” alt problemine ilişkin bulgular Tablo 4.3.' de verilmiştir.

Tablo 4.3. MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 ' si Tekniklerinin 1800 Örneklem Büyüklüğü için İstatistiksel Güç Oranları

Örneklem Büyüklüğü	Madde Sayısı	DMF İçeren Madde Oranı	R/T	MH	LR	Lord χ^2	Raju
1800	20	0.05	0.50	0.80	0.68	0.64	0.52
			0.75	0.64	0.44	0.56	0.44
		0.1	0.50	0.60	0.54	0.50	0.38
			0.75	0.20	0.14	0.34	0.32
	80	0.05	0.50	0.11	0.12	0.41	0.39
			0.75	0.73	0.71	0.68	0.70
		0.1	0.50	0.09	0.10	0.31	0.24
			0.75	0.59	0.51	0.53	0.51

Tablo 4.3. incelendiğinde 1800 örneklem büyüklüğü ve 20 madde için örneklem büyüklüğü ve DMF içeren madde oranlarının değiştiği koşullar altında istatistiksel güç değerlerinin 0.14 ile

0.80 arasında deđiřtiđi grlmektedir. 20 madde iin DMF ieren madde oranı 0.05 ve rneklem byklđ oranları 0.50 ve 0.75 olduđunda sırasıyla 0.80 ve 0.64 ile MH tekniđi en yksek gce sahip olarak bulunurken; 0.50 rneklem byklđ oranı iin 0.52 ile Raju' nun alan lmleri tekniđi; 0.75 rneklem byklđ oranı iin ise 0.44 ile LR ve Raju'nun alan lmleri teknikleri en dřk gce sahip olarak ıkmıřtır. 20 madde iin DMF ieren madde oranı 0.1 ve rneklem byklđ oranları sırasıyla 0.50 ve 0.75 olduđunda sırasıyla 0.60 ile MH ve 0.34 ile Lord χ^2 si teknikleri en yksek gce sahip olarak bulunurken; 0.50 rneklem byklđ oranı iin 0.38 ile Raju'nun alan lmleri ve 0.75 rneklem byklđ oranı iin 0.14 ile LR teknikleri en dřk gce sahip olarak bulunmuřtur.

1800 rneklem byklđ ve 80 madde iin rneklem byklđ ve DMF ieren madde oranlarının deđiřtiđi kořullar altında istatistiksel g deđerlerinin 0.11 ile 0.73 arasında deđiřtiđi grlmektedir. 80 madde iin DMF ieren madde oranı 0.05 ve rneklem byklđ oranları 0.50 ile 0.75 olduđunda sırasıyla 0.41 ile Lord χ^2 si ve 0.73 ile MH teknikleri en yksek gce sahip olarak bulunurken; rneklem byklđ oranı 0.50 iin 0.11 ile MH tekniđi ve rneklem byklđ oranı 0.75 iin ise ve 0.68 ile Lord χ^2 si teknikleri en dřk gce sahip olarak bulunmuřtur. 80 madde iin DMF ieren madde oranı 0.1 ve rneklem byklđ oranları 0.50 ve 0.75 olduđunda sırasıyla 0.31 ile Lord χ^2 si ve 0.59 ile MH teknikleri en yksek gce sahip olarak bulunmuřtur. rneklem byklđ oranları 0.50 iken 0.09 ile MH tekniđi ve rneklem byklđ oranı 0.75 iken 0.51 ile LR ve Raju' nun alan lmleri teknikleri en dřk gce sahip olarak bulunmuřtur.

4.2.2. 4. Alt Probleme İliřkin Bulgular

“rneklem byklđnn 3000; rneklem byklđ oranının 0.50 ve 0.75, madde sayısının 20 ve 80, DMF ieren madde oranının 0.05 ve 0.1 olduđu kořullar altında MH, LR, Raju' nun Alan lmleri ve Lord' un χ^2 si tekniklerinin istatistiksel g oranları nasıl deđiřmektedir?” alt problemine iliřkin bulgular Tablo 4.4.' de verilmiřtir.

Tablo 4.4. *MH, LR, Raju' nun Alan Ölçümleri ve Lord' un χ^2 ' si Tekniklerinin 3000 Örneklem Büyüklüğü için İstatistiksel Güç Oranları*

Örneklem Büyüklüğü	Madde Sayısı	DMF İçeren Madde Oranı	R/T	MH	LR	Lord χ^2	Raju	
3000	20	0.05	0.50	0.12	0.08	0.28	0.20	
			0.75	0.20	0.12	0.12	0.04	
	80	0.1	0.50	0.14	0.08	0.18	0.20	
			0.75	0.08	0.04	0.42	0.38	
	80	0.05	0.50	0.50	0.08	0.12	0.32	0.28
			0.75	0.85	0.78	0.60	0.58	
80	0.1	0.50	0.50	0.77	0.76	0.60	0.53	
		0.75	0.13	0.14	0.48	0.44		

Tablo 4.4. incelendiğinde 3000 örneklem büyüklüğü ve 20 madde için örneklem büyüklüğü ve DMF içeren madde oranlarının değiştiği koşullar altında istatistiksel güç değerlerinin 0.04 ile 0.28 arasında değiştiği görülmektedir. 20 madde için DMF içeren madde oranı 0.05 ve örneklem büyüklüğü oranları 0.50 ve 0.75 olduğunda sırasıyla 0.28 ile Raju' nun alan ölçümleri tekniği ve 0.20 le MH tekniği en yüksek güce sahip olarak bulunurken; 0.50 örneklem büyüklüğü oranı için 0.08 ile LR tekniği; 0.75 örneklem büyüklüğü oranı için ise 0.04 ile Raju' nun alan ölçümleri teknikleri en düşük güce sahip olarak çıkmıştır. 20 madde için DMF içeren madde oranı 0.1 ve örneklem büyüklüğü oranları sırasıyla 0.50 ve 0.75 olduğunda sırasıyla 0.20 ile Raju' nun alan ölçümleri ve 0.42 ile Lord χ^2 ' si teknikleri en yüksek güce sahip olarak bulunurken; 0.50 örneklem büyüklüğü oranı için 0.08 ile LR tekniği ve 0.04 örneklem büyüklüğü oranı için 0.14 ile LR teknikleri en düşük güce sahip olarak bulunmuştur.

3000 örneklem büyüklüğü ve 80 madde için örneklem büyüklüğü ve DMF içeren madde oranlarının değiştiği koşullar altında istatistiksel güç değerlerinin 0.08 ile 0.85 arasında değiştiği görülmektedir. 80 madde için DMF içeren madde oranı 0.05 ve örneklem büyüklüğü oranları 0.50 ile 0.75 olduğunda sırasıyla 0.32 ile Lord χ^2 ' si ve 0.85 ile MH teknikleri en yüksek güce sahip olarak bulunurken; örneklem büyüklüğü oranı 0.50 için 0.08 ile MH tekniği ve ve örneklem büyüklüğü oranı 0.75 için ise 0.58 ile Raju' nun alan ölçümleri tekniği en düşük güce sahip olarak bulunmuştur. 80 madde için DMF içeren madde oranı 0.1 ve örneklem büyüklüğü oranları 0.50 ve 0.75 olduğunda sırasıyla 0.77 ile MH tekniği ve 0.48 ile Lord χ^2 ' si en yüksek

güce sahip olarak bulunmuştur. Örneklem büyüklüğü oranları 0.50 iken 0.53 ile Raju' nun alan ölçümleri ve 0.13 ile MH teknikleri en düşük güce sahip olarak bulunmuştur.



V. SONUÇ, TARTIŞMA VE ÖNERİLER

5.1. Sonuç ve Tartışma

Araştırmanın bu bölümünde sırasıyla araştırmanın alt problemlerine ilişkin sonuç ve tartışmalara yer verilmiştir.

Araştırmanın birinci ve ikinci alt problemi olan farklı koşullar altındaki DMF belirleme tekniklerinin I. Tip hata oranlarına genel olarak bakıldığında, karşılaştırmaya dahil edilen dört tekniğinde belirlenen tüm koşullar altında ortaya çıkan I. Tip hata oranlarının nominal hata düzeyinden (0.05) yüksek olduğu bulunmuştur.

Birinci alt problem doğrultusunda elde edilen bulgular sonucunda, 1800 örneklem büyüklüğü için belirlenen tüm koşullar incelendiğinde en yüksek I. Tip hata oranlarına sahip teknik olarak MH bulunmuştur. Bununla birlikte elde edilen tüm değerler, şişirilmiş I. Tip hata göstermektedir. Kullanılan tekniklerin koşullar içerisinde incelenmesi yapıldığında, MH ve LR tekniklerinin çok yakın değerler aldığı ve 1800 örneklem büyüklüğü için madde sayısının 20 olduğu koşulda, 80 olduğundaki koşula göre daha yüksek hatalar verdiği bulunmuştur. Her iki teknikte de en yüksek hata oranını, madde sayısının 20, R/ T oranının 0.50 olduğu koşul altında verirken; en düşük hata oranını ise madde sayısının 80 ve R/ T oranının 0.75 olduğu koşul altında vermiştir.

Aynı alt problem altında Lord' un χ^2 ' si ve Raju' nun alan ölçümleri teknikleri incelendiğinde ise her iki tekniğinde madde sayısının 80 ve R/ T oranının 0.50 olduğu koşul altında en yüksek I. Tip hata oranları verdiğine ulaşılmıştır. En düşük hata oranları ise madde sayısının 20; R/ T oranının 0.75 olduğu koşul altında ulaşılmıştır.

Literatürdeki diğer araştırmalar incelendiğinde, bu araştırmanın bulguları ile hem tutarlılık gösteren hem de farklı sonuçların elde edildiği görülmüştür. Kim (2010) tarafından yürütülen araştırmada, farklı koşullar altında LR, MH, DFIT ve Lord' un χ^2 ' si tekniklerinin I. Tip hata oranları üzerinde çalışılmıştır. Araştırması sonucunda, bu araştırmaya ait sonuçlar ile tutarlılık göstererek test uzunluğunun daha az olduğu, geniş örneklem büyüklüğü ile odak ve referans grubun eşit olduğu koşullar altında tüm tekniklerin I. Tip hatayı şişirme eğiliminde olduğuna ulaşmıştır. Bir başka çalışmada da benzer şekilde test uzunluğu kıaldıkça ve örneklem büyüklüğü arttıkça, MH ve LR tekniklerinin şişirilmiş I. Tip hata oranları verdiğine belirtilmiştir.

(Demars, 2009). Diğer taraftan, Dainis (2008) tarafından hem gerçek veri hem de simülasyon veri üzerinden yürütülen bir çalışmada ise MTK-OO, DFIT, MH ve LR tekniklerinin I. Tip hatalarının kıyaslanması sonucunda, MH tekniğinin hata oranları kabul edilebilir bir şekilde nominal düzeyde olduğu bulunmuştur. Gierl, Jodoin ve Ackerman (2000) ise yetenek dağılımlarının eşit olan koşullarda örneklem büyüklüğünün küçük olması halinde bile MH ve LR tekniklerinin I. Tip hata oranlarının nominal hata düzeylerinin çevresinde hatta altında bile olabileceğini vurgularken; benzer şekilde Ankenmann, Witt ve Dunbar(1999) özellikle LR tekniğinin genel olarak koşullar kapsamında nominal hata düzeyinde I. Tip hata oranlarını verdiğini ve MH tekniğinin bu nominal hata düzeyinin üzerinde hata oranları verdiğini belirtmişlerdir . Ancak bu çalışmada bu durumun aksine örneklem büyüklüğünün geniş ya da küçük olarak kabul edildiği her durum için her iki teknikte, kabul edilebilen hata oranlarının çok daha üstünde I. Tip hata oranları vermiştir.

Birinci alt probleme ait sonuçları genel olarak değerlendirecek olursak, 1800 örneklem büyüklüğü, 20 ve 80 madde sayısı ile R/ T oranlarının 0.50 ve 0.75 olduğu değişimlenen koşullar altında tekniklerden elde edilen I. Tip hata oranlarının, tekniklerin ait oldukları kuramlar bazında, KTK ve MTK' ye dayalı tekniklerin kendi içerisinde etkilendikleri koşullar açısından benzerlik gösterdiği, Kan, Sünbül ve Ömür (2013)' ün çalışmasında da olduğu gibi yorumlanabilmektedir. Her iki kuramdaki tekniklerde, odak ve referans grubun örneklem büyüklüğünün eşit olmadığı 0.75 olduğu koşulda daha düşük I. Tip hata oranları vermektedir. Keklik (2012)' de araştırmasında benzer şekilde örneklem büyüklükleri oranlarının 1:1 olduğu durumlardan ziyade 1:2 olduğu durumlarda daha düşük I. Tip hata oranları elde edildiğini belirtmiştir.

Yukarıdaki ifadelere ek olarak MTK' ye dayalı Lord' un χ^2 ' si ve Raju' nun alan ölçümleri teknikleri birinci alt problemdeki koşullar için daha düşük I. Tip hata oranları verdiği de söylenebilir.

İkinci alt problem doğrultusunda elde edilen bulgular sonucunda, 3000 örneklem büyüklüğü için belirlenen tüm koşullar incelendiğinde en yüksek I. Tip hata oranlarına sahip teknik olarak, 0.928 değeri ile LR bulunmuştur. Ancak, bu değere oldukça yakın bir şekilde 0.924 hata oranı ile MH tekniğini yine en yüksek I. Tip hata oranına sahip olarak kabul edebiliriz. Bununla birlikte elde edilen tüm değerler, tıpkı 1800 örneklem büyüklüğü ile oluşturulan koşullarda olduğu gibi şişirilmiş I. Tip hata göstermektedir.

I. tip hata oranlarının belirlenmesi amacıyla oluşturulan koşullarda örneklem büyüklüğünün 1800' den 3000' e çıkartılarak genişletilmesi, üç koşul dışında tamamında I. Tip hata değerlerini yükseltmiştir. Bunun aksine bir durum gözlemlenen 3 koşul ve teknikler ise şu şekildedir: R/ T oranının 0.75 olduğu 20 madde için MH tekniği daha düşük I. Tip hata oranı vermiştir. Yine 20 madde için R/ T oranının 0.50 olduğu koşulda hem 1800 hem de 3000 örneklem büyüklüğü için MH tekniği aynı hata oranı değerini verirken R/T oranının 0.75 olduğunda da LR tekniği aynı hata oranı değerini vermiştir. Araştırma çerçevesinde kullanılan tekniklerin koşullar içerisinde incelemesi yapıldığında, MH ve LR tekniği 1800 örneklem büyüklüğünde olduğu gibi burada da tüm koşullarda birbirlerine çok yakın değerler almıştır. Her iki teknikte en düşük I. Tip hata oranlarını, madde sayısının 20 ve R/ T oranının 0.75 olduğu koşul altında vermiştir.

İkinci alt problem olan 3000 örneklem büyüklüğü ile değişimlenen koşullar altında elde edilen Lord' un χ^2 ' si ve Raju' nun alan ölçümleri tekniklerine ait bulgulara bakıldığında ise her iki tekniğinde madde sayısının 80 ve R/ T oranının 0.75 olduğu koşul altında en yüksek I. Tip hata oranları verdiğine ulaşılmıştır. En düşük I. Tip hata oranlarını ise tıpkı KTK' ye dayalı tekniklerde olduğu gibi madde sayısının 20; R/ T oranının 0.75 olduğu koşul altında ulaşılmıştır.

Daha önceden belirtilmiş olan örneklem büyüklüğünün genişlemesi ile I. Tip hata oranlarının yükselmesi sonucu literatürdeki diğer çalışmalar ile tutarlıdır. Sünbül ve Ömür Sünbül (2016) tarafından yapılan araştırmada da Raju' nun Alan Ölçümleri, Lord' un χ^2 ' si, MH ve LR tekniklerinin farklı koşullar altında karşılaştırılması yapılmıştır. Sonucunda ise özellikle MTK' ye dayalı tekniklerde örneklem büyüklüğünün artması ile I. Tip hata oranlarının azaldığı sonucuna ulaşımlardır. Yine simülasyon veri üzerinden Magis ve De Boeck (2012)' tarafından yürütülen bir başka araştırmada ise bu durumun I. Tip hata oranlarını şişirdiğini belirtmişlerdir. Bunun yanı sıra daha farklı bir sonuç ile Vaughn ve Wang (2010) MH ve LR tekniğini; Price (2014) ise MH tekniğini farklı örneklem büyüklüğü altındaki koşullarda sınamasına rağmen düşük I. Tip hata oranlarına sahip olarak bulmuştur. Araştırmanın bulguları ile tutarlılık gösteren farklı simülasyon çalışmaları da vardır (Erdem Keklik, 2014; Li, Brooks ve Johanson, 2012; Rousses ve Stout, 1996).

I. Tip hata oranlarının belirlenmesi amacıyla oluşturulan koşullar, hep birlikte değerlendirildiğinde ise her iki alt problem için de Raju' nun Alan Ölçümleri tekniği tüm koşullar altında en düşük I. Tip hataya sahip olarak çıkmıştır.

Yapılan tüm bu yorumlara ek olarak, çeşitli koşullar altında araştırmaya dahil edilen teknikler ile elde edilen I. Tip hata oranlarının nominal hata düzeyinin üstünde olmasının nedeni olarak oluşturulan koşullar gösterilebilir. Kullanılan örneklem büyüklüğü ve test uzunluğu için bu durum söz konusu olabilir. Araştırmada kullanılan örneklem büyüklükleri, küçük değildir; büyük örneklem kullanılmıştır ki büyük örneklem büyüklüklerinin I. Tip hata oranlarını yükselttiği bulgusu literatürdeki diğer çalışmalar ile desteklenmektedir. Test uzunluğu için ise, 20 madde sayısı ile yapılan analizlerde 80 madde sayısı ile oluşturulan koşullara nispeten daha yüksek I. Tip hata oranları elde edilmiştir. Yine KTK' ye dayalı tekniklerin daha yüksek oranların vermesi ise ilgili tekniklerin kuramsal yapısı ile ilişkili olarak açıklanabilir. İlk olarak değinebileceğimiz nokta, MH ve LR tekniklerinin gözlenen puana dayalı teknikler olmasıdır. Bu teknikler, DMF' nin belirlenmesi sürecinde eşleştirme değişkeni olarak gerçek puanları kullanmaktadır. Kullanılan tekniklerin bu yönden değerlendirilmesinde Zwick, Thayer ve Mazzeo (1997)' nin de belirttiği gibi, gözlenen puana dayalı olan tekniklerin ölçme hataları, testin güvenilirliği düşürebilmektedir. Dolayısı ile bu durum ortalamadan sapmış olan gerçek puanlara neden olur. Değinebilecek bir diğer problem ise toplam test puanlarının kullanılıyor olmasıdır. Yapılan araştırmalar göz önünde bulundurulduğunda MTK modellerine uygun olarak üretilen veriler ile MH ve LR gibi gözlenen puana dayalı teknikler ile çalışmak şişirilmiş I. Tip hataya sebep olduğu görülmektedir.

Üçüncü alt probleme ait bulgular doğrultusunda 1800 örneklem büyüklüğü için oluşturulmuş farklı koşullar altında yapılan analizler sonucunda, en yüksek istatistiksel güç oranı değerini MH tekniği 0.80 ile vermiştir. MH tekniği, madde sayısının 20; R/T oranının 0.75 ve DMF içeren madde oranının 0.10 olduğu koşul ile madde sayısının 80, R/ T oranının 0.50 ve DMF miktarının 0.05 ile 0.10 olduğu üç koşul dışında tüm koşullarda en yüksek istatistiksel güç oranlarını vermiştir. En yüksek değerine ise beklenilen aksine madde sayısının 80 değil, 20 olduğu durumda ulaşmıştır. Araştırmanın sonuçlarına benzer şekilde Kabasakal, Arsan, Gök ve Kelecioğlu (2014)' nun çalışmasında da tüm koşullar altında MH tekniği en yüksek istatistiksel güç oranlarına sahip olarak bulunmuştur. Bir başka araştırmada da yine MH tekniğinin en güçlü teknik olduğu bulgusu elde edilmiştir (Kristjansson, Aylesworth, McDowell ve Zumbo, 2005).

Araştırmaya dahil edilen teknikler bazında sonuçlar ele alındığında ise LR, Lord' un χ^2 ' si ve Raju' nun Alan Ölçümleri teknikleri en yüksek istatistiksel güç oranlarına madde sayısının 80, DMF içeren madde oranının 0.05 ve R/ T oranının 0.75 olduğu koşulda ulaşmışlardır. Ancak

elde edilen oranlar, 0.70 çevresinde dağılım göstermekte ve istenilen değer olan 0.80' e ulaşamamışlardır. Tekniklerin dayalı oldukları kurama ait genel bir değerlendirmesi yapılırsa, üçüncü alt problem kapsamında yer alan sekiz koşulun, beşinde KTK' ye dayalı olan MH ve LR teknikleri daha yüksek değerler alırken; üçünde MTK' ye dayalı olan Lord' un χ^2 si ve Raju' nun Alan Ölçümleri teknikleri daha yüksek değerler almıştır. Buradan hareketle koşuldan koşula tekniklerin düzenli bir farklılık göstermemesi sebebiyle yukarıdaki sonuca dayanarak KTK' ye dayalı tekniklerin 1800 örneklem büyüklüğünde daha yüksek istatistiksel güç oranları verdiğini yorumu yapılabilir.

Yukarıda üçüncü alt probleme dayalı olarak verilen sonuçlara ek olarak tüm teknikler için ortak olan diğer sonuçlar ise şu şekildedir; madde sayısının hem 20 hem de 80 olduğu koşullar için DMF içeren madde oranının 0.05' den 0.10' a çıkması istatistiksel güç oranlarını olumsuz yönde etkilemekte ve düşürmektedir. Ancak literatürdeki diğer araştırmalar incelendiğinde, Atar ve Kamata (2011) özellikle LR tekniğinin örneklem büyüklüğü küçük ise DMF içeren madde oranı düştükçe istatistiksel güç oranlarında düştüğünü vurgulamıştır. Elde edilen sonuçlar, bu araştırmanın aksine bulgular vermiştir. Bu çalışmada DMF içeren madde oranının daha düşük olduğu 0.05 olduğu koşullar altında daha yüksek istatistiksel güç oranlarına ulaşılmıştır. Yine başka bir çalışmada LR ve MH tekniklerinin performanslarının karşılaştırması yapılmıştır ve istatistiksel güç oranlarının DMF içeren madde oranı arttıkça yükseldiğine ulaşılmıştır (Hidalgo ve Lopez- Pina, 2004). Bu doğrultuda DMF içeren madde oranına dair bir genelleme yapılamamaktadır. R/ T oranının 0.75 olduğu koşullar altında, madde sayısı 20 olduğunda istatistiksel güç oranları düşüş gösterirken; madde sayısının 80 olduğu durumlarda yükselmektedir. Diğer bir deyiş ile, madde sayısının 20 olduğu, R/ T oranı 0.50 olduğu tüm koşullarda istatistiksel güç oranları daha yüksekken; madde sayısı 80 iken R/ T oranı 0.75 olduğunda daha yüksektir. referans ve odak grupların örneklem büyüklüğü oranlarına ait literatürdeki diğer çalışmalar incelendiğinde ise bu orandan en çok etkilenen tekniğin MH olduğu ve iki grubun örneklem büyüklüğünün eşit olmadığı durumlarda istatistiksel güç oranlarının düştüğü belirtilmiştir (Kabasakal, Arsan, Gök ve Kelecioğlu, 2014).

Dördüncü alt probleme ait bulgular doğrultusunda 3000 örneklem büyüklüğü için oluşturulan farklı koşullar altında yapılan analizler sonucunda, yine 1800 örneklem büyüklüğünde olduğu gibi en yüksek istatistiksel güç oranları burada da MH tekniği ile 0.85 olarak elde edilmiştir. Bununla birlikte teknik, 1800 örneklem büyüklüğüne kıyasla koşulların çoğunda diğer

tekniklerden daha düşük performans göstermiştir. Bu noktada MH tekniğinin örneklem büyüklüğünden etkilendiği söylenebilir. Literatürde MH ve LR tekniklerinin örneklem büyüklüğünün artması ile istatistiksel güç oranlarını düşürdüğü belirtilen araştırma sonuçları ile tutarlı çalışmalar bulunmaktadır (Erdem Keklik, 2012; Vaughn ve Wang, 2010). Ankenmann, Witt ve Dunbar (1999) ise ilgili tekniklerin, örneklem büyüklüklerinden etkilendiğini vurgulamıştır.

Araştırma kapsamında değerlendirilen teknikler göz önünde bulundurulduğundan hemen hemen tüm koşullar altında madde sayısının 80 olduğu koşullar daha yüksek istatistiksel güç oranları vermiştir. Tekniklerin en yüksek istatistiksel güç oranları elde ettiği koşul ise madde sayısının 80, DMF içeren madde oranının 0.05 ve R/ T oranının 0.75 olduğudur. Bu doğrultuda madde sayısının artmasının istatistiksel güç oranları üzerinde olumlu yönde bir etkisi olduğu yorumu yapılabilir. Burada değinilebilecek bir diğer nokta ise R/ T oranının güç oranların yansımadır. Bu çalışmada belirtildiği gibi oran 0.75 olduğunda daha yüksek istatistiksel güç oranları elde edilmiş olmasına karşın farklı araştırmalarda, grupların örneklem büyüklüklerinin aynı olmadığı koşullarda tekniklerin dengesizliklerin olabildiği de ifade edilmiştir (Jodoin ve Gierl 2001; Narayanan ve Swaminathan, 1994). Dolayısı ile R/ T oranına ait kesin bir yorum yapılamamaktadır.

Dördüncü alt problemde elde edilen sonuçları koşullar bazında incelersek, teknikler büyük çoğunlukla madde sayısının 80 olduğu koşullar, 20 olduğundan daha iyi performans göstermiştir. DMF oranının 0.10 olduğu koşullar göz önünde bulundurulduğunda, madde sayısı farketmeksizin R/ T oranı hem 0.50 hem de 0.75 olduğunda, birkaç koşul dışında DMF içeren madde oranı 0.05' e göre daha yüksek istatistiksel güç oranlarına ulaşılmıştır. Bu bulgu literatürdeki diğer çalışmalar ile tutarlılık göstermektedir (Atar ve Kamata, 2011). Jodoin ve Gierl (2001) tarafından gerçekleştirilen bir çalışmada ise DMF içeren madde oranı 0.10 olduğunda, tekniklerin daha iyi performans gösterdiği ifade edilirken; Erdem Keklik (2012) b parametresine eklenen DMF miktarının artışının, tekniğin gücünü arttırdığını ifade etmiştir. Chang, Mazzeo ve Roussos(1996) ise a parametresi arttırıldıkça DMF belirleme tekniklerinin de buna bağlı olarak güç oranlarının artacağını eklemiştir.

Araştırma kapsamında daha öncede belirtildiği gibi istatistiksel güç oranları için üçüncü ve dördüncü alt problemler oluşturulmuştu. Bu doğrultuda iki farklı örneklem büyüklüğüne sahip iki alt problemin karşılaştırılması yapıldığında, tekniklerden elde edilen istatistiksel güç

oranları, 1800 örneklem büyüklüğü ile oluşturulmuş koşulların yarısından çoğunda; 3000 örneklem büyüklüğündeki koşullara kıyasla daha iyi performans göstermiştir. Ortaya çıkan oranlar incelendiğinde yalnızca dokuz tanesinin 3000 örneklem büyüklüğünde daha yüksek olduğu söylenebilir. Ancak bu değerlerin koşullar anlamında genellenebilirliğini yapmak güçtür. Bununla birlikte yalnızca madde sayısının 80, DMF içeren madde oranının 0.10 ve R/ T oranının 0.50 olduğu koşulda tüm tekniklerin daha yüksek istatistiksel güç oranları verdiği söylenebilir.

Kuramlar açısından yapılan kıyaslama doğrultusunda ise genel olarak 1800 örneklem büyüklüğünde daha yüksek istatistiksel güç oranları elde edilmiştir. Ki bu da Sünbül ve Ömür Sünbül (2016)'ın bulgularının aksine bir sonuçtur. Bahsedilen araştırmada örneklem büyüklüğünün artmasının hem KTK hem de MTK'ye dayalı tekniklerin istatistiksel güç oranlarını yükselttiği belirtilmiştir. Benzer şekilde büyük örneklem büyüklüklerinin güç oranlarını arttırdığı sonucuna ulaştığını belirten farklı araştırmalarda mevcuttur (Atar, 2007; Narayanan ve Swaminathan, 1994). Ancak bu çalışmada, aşağıda belirtilen birkaç koşul dışında tüm teknikler 1800 örneklem büyüklüğünde daha iyi performans göstermiştir. Bir koşulda KTK'ye dayalı teknikler 3000 örneklemde daha iyi bir performans gösterirken; bir diğer koşulda MTK'ye dayalı teknikler göstermiştir. Her iki kuramında ortak olarak büyük örneklem büyüklüğünde daha iyi performans gösterdiği koşul, madde sayısının 80, DMF içeren madde oranının 0.10 ve R/ T oranının 0.50 olduğudur.

5.2. Öneriler

5.2.1. Uygulamaya Dönük Öneriler

Yapılan bu araştırmadan elde edilen bulgular doğrultusunda yapılabilecek olan uygulamaya dönük öneriler aşağıda sunulmuştur.

Bu araştırma kapsamında I. Tip hata oranlarına ilişkin sonuçlar göz önünde bulundurulduğunda, en yüksek hata oranlarına MH ve LR teknikleri ile ulaşılmıştır. Ancak araştırmada büyük örneklem büyüklüklerinin kullanılmış olması bunun bir nedeni olarak gösterilebilir. Bu nedenle daha küçük örneklem ile çalışacak olan uygulamacıların KTK'ye dayalı teknikleri tercih etmesi önerilebilir.

MTK' ye dayalı teknikler olan Lord' un χ^2 ' si ve Raju' nun Alan Ölçümleri tekniklerinde ise örneklem büyüklüğünün artması ile I. Tip hata oranları düşmektedir. Bu durumun MTK' nin kuramsal yapısı ile ilgili olan bir durum olmasının yanı sıra, büyük örneklem büyüklükleri üzerinde çalışacak uygulamacılar MTK' ye dayalı teknikleri tercih edebilir.

Ek olarak bu araştırmada kullanılan R/ T oranları göz önünde bulundurulduğunda odak ve referans grubun örneklem büyüklüklerinin eşit olmadığı koşullar altında I. Tip hata oranları daha düşük değerlerdir. Bu nedenle uygulamacıların grupların oluşturulması noktasında imkanları olursa, eşit örneklem büyüklüğü olmayacak şekilde oluşturması önerilebilir.

Bu araştırma kapsamında tekniklere ilişkin istatistiksel güç oranları üzerinde durulduğunda, en güçlü teknik olarak MH tekniği bulunmuştur. Ancak daha önce de belirtildiği gibi I. Tip hata oranı en büyük olan teknikte MH tekniği olduğu için, uygulamacıların bu tekniğin örneklem büyüklüğünden etkileniyor olduğu bulgusunu göz önünde bulundurarak tercih etmesi önerilmektedir. Küçük örneklem büyüklükleri ve kısa test uzunluğunda MH tekniği tercih edilebilir.

Genel olarak KTK' ye dayalı bir yorum yapıldığında ise uygulamacılara yine benzer şekilde küçük örneklem büyüklüklerinde bu kurama dayalı tekniklerin tercih edilmesi gerektiği tavsiye edilebilir. Çünkü LR tekniği de tıpkı MH gibi örneklem büyüklüğünden etkilenmektedir.

Bu araştırmada oluşturulmuş olan simülasyon koşulları dikkate alındığında ise I. Tip hata oranlarında olduğu gibi istatistiksel güç oranları içinde R/ T oranının referans ve odak grup örneklem büyüklüklerinin eşit olmadığı koşullar daha iyi performans göstermiştir. Ancak uygulamacıların üzerinde çalıştıkları örneklemde gruplar eşit büyüklüğe sahip ise madde sayısının az olması önerilebilir. Bu şekilde tekniklerin daha iyi performans göstermeleri sağlanacaktır. Bunlara ek olarak uygulamacılar büyük örneklem büyüklüğüne sahipse madde sayısının fazla olduğu ölçme araçlarını kullanmaları tavsiye edilebilir.

5.2.2. İleriye Dönük Araştırmalara İlişkin Öneriler

Değişen madde fonksiyonu alanında yapılacak olan ileriye dönük araştırmalara ilişkin öneriler aşağıda sunulmuştur.

Gelecek çalışmalarda, araştırmacılar bu araştırma kapsamında kullanılan sabit tutulan ve manipüle edilen koşullar üzerinden değişiklik yaparak yeniden çalışabilir. Grupların yetenek

dağılımları manipüle edilerek farklı dağılımlara sahip gruplar üzerinde çalışma yürütülebilir. Bu araştırmada grupların yetenek düzeyleri sabit tutulan bir koşul olarak ele alınmıştır ve grupların yetenek dağılımları, belirlenen değer aralıklarında tesadüfi bir şekilde seçilmiştir. Yine hem aynı yetenek dağılımına sahip hem farklı yetenek dağılımına sahip gruplar ile çalışılabilir.

Bu araştırmada DMF içeren madde oranına göre b parametresi üzerinde değişiklikler yapılarak tek biçimli DMF oluşturulmuştu. Gelecek çalışmalarda a parametresindeki değerler değiştirilerek hem madde ayırt ediciliğinin DMF belirleme tekniklerinin performansına etkisi hem de tek biçimli olmayan DMF' nin belirlenmesi noktasında çalışmalar yürütülebilir.

Bu araştırma yalnızca iki kategorili puanlanan maddeler üzerinden yürütülmüştür. Gelecek çalışmalarda bu araştırmada kullanılan teknik ve koşulların maddelerin kategori sayıları artırıldığında tekniklerin performansları üzerine nasıl etkiler gösterebileceği incelenebilir.

Bu araştırmada MH, LR, Lord' un χ^2 si ve Raju' nun Alan Ölçümleri teknikleri kullanılmıştır. Ancak, araştırmanın bulguları doğrultusunda yukarıda belirtilen tekniklerin performansları noktasında belirsizlikler ve eksikliklerin var olduğu ortaya çıkmıştır. Bu nedenle, MTK- OO ya da Genelleştirilmiş MH gibi farklı teknikler tercih edilerek çalışma yürütülebilir.

Bu araştırma yalnızca simülasyon veriler üzerinde çalışılarak yürütülmüştür. Yapılan çalışmanın uygulamadaki yansımalarının daha açık bir şekilde ortaya konulabilmesi ve karşılaştırılabilmesi amacıyla gelecek çalışmalarda hem simülasyon hem de gerçek veri üzerinden araştırmaların yapılması uygun olacaktır.

Bu araştırmada tekniklerin performansları incelenirken I. Tip hata ve istatistiksel güç oranları kullanılmıştır. Araştırmacılar kullanılan tekniklerin II. Tip hata oranlarını yönünden incelenmesini de çalışmalarına dahil edebilirler.

Bu araştırma kapsamında üretilen verilerin tamamı 2PL madde tepki modele uygun bir şekilde üretilerek bu koşul sabit tutulmuştur. Ancak bilindiği üzere 1PL ve 3PL modellerde mevcuttur. Farklı parametrelili modeller araştırmaya dahil edilerek hangi tekniğin hangi modelde daha iyi performans göstereceği tespit edilebilir.

Son olarak yapılan bu araştırmada verilerin analizi aşamasında R programı kullanılmıştır. Ancak literatürde kullanılan programlara göre DMF belirleme tekniklerinin performansının da

değişebileceği belirtilmektedir. Bu amaçla gelecek çalışmalarda R programının yanında farklı programlar kullanılarak teknikler karşılaştırılabilir.



KAYNAKÇA

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Atalay Kabasakal, K., Arsan, N., Gök, B., Kelecioğlu, H. (2014). Değişen Madde Fonksiyonunun Belirlenmesinde MTK Olabilirlik Oranı SIBTEST ve Mantel- Haenszel Yöntemlerinin Performanslarının (I. Tip Hata ve Güç) Karşılaştırılması. *Kuram ve Uygulamada Eğitim Bilimleri*, 14 (6), 2175- 2193.
- Atar, B. (2007). *Differential Item Functioning Analyses for Mixed Response Data Using IRT Likelihood-Ratio Test, Logistic Regression, and GLLAMM Procedures*. (Yayımlanmamış doktora tezi). Florida State University, Tallahassee.
- Atar, B., Kamata, A. (2011). Comparison of IRT Likelihood Ratio Test and Logistic Regression DIF Detection Procedures. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 41, 36–47.
- Ankenmann, R. D., Witt, E. A., Dunbar, S. B. (1999). An Investigation of the Power of the Likelihood Ratio Goodness-of-Fit Statistics in Detecting Differential Item Functioning. *Journal of Educational Measurement*, 36 (4), 277–300.
- Baykul, Y. (2000). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulanması*. Ankara: ÖSYM Yayınları.
- Camilli, G., Shepard, L.A. (1994). *Methods for Identifying Biased Test Items*. California: Sage Publications.
- Chang, H., Mazzeo, J., Roussos, L. (1996). Detecting DIF for Polytomously Scored Items: An Adaptation of SIBTEST Procedure. *Journal of Educational Measurement*, 33(3), 333–353.
- Clauser, B. E., Mazor, K. M. (1998). Using Statistical Procedure to Identify Differential Item Functioning Test Items. *Educational Measurement: Issues and Practice*, 17, 31-44.
- Cohen, A. S., Kim, S. H., and Wollack, J. A. (1996). An Investigation of The Likelihood Ratio Test for Detection of Differential Item Functioning. *Applied Psychological Measurement*, 20(1), 15–26.

- Crocker, L., Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: CBS College Publishing.
- Dainis, A. M. (2008). *Method for Identifying Differential Item and Test Functioning: an Investigation of Type 1 Error Rates and Power*. (Yayımlanmamış doktora tezi). James Madison University, Harrisonburg.
- DeMars, C. E. (2009). Modification of the Mantel-Haenszel and Logistic Regression DIF Procedures to Incorporate the SIBTEST Regression Correction. *Journal of Educational and Behavioral Statistics*, 34 (2), 149- 170.
- Demir, S., Köse, İ. A. (2014). Mantel- Haenszel, SIBTEST ve Lojistik Regresyon Yöntemleri ile Değişen Madde Fonksiyonunu Analizi. *International Journal of Human Sciences*, 11 (1), 700- 714.
- Desa, Z. N. (2012). *Bi-factor Multidimensional Item Response Theory Modeling for Subscores Estimation, Reliability, and Classification*. (Yayımlanmamış doktora tezi). University of Kansas.
- Dodeen, H. (2004). The Relationship Between Item Parameters and Item Fit. *Journal of Educational Measurement*, 41 (3), 261- 270.
- Dorans, N. J., Holland, P. W. (1993). DIF Detection and Description: Mantel-Haenszel and Standardization. P. W. Holland, H. Wainer (ed.), *Differential Item Functioning* içinde (s. 35-66). Hillsdale: Lawrence Erlbaum.
- Ellis, B. B., Raju, N. S. (2003). Test and Item Bias: What They Are, What They Aren't, and How to Detect Them. *ERIC Institute Education Sciences*. <https://files.eric.ed.gov/fulltext/ED480042.pdf>.
- Embretson, S.E., Reise, S.P. (2000). *Item Response Theory for Psychologists*. London: Lawrence Erlbaum Associates, Publishers.
- Erdem Keklik, D. (2012). *İki Kategorili Maddelerde Tek Biçimli Değişen Madde Fonksiyonu Belirleme Tekniklerinin Karşılaştırılması: Bir Simülasyon Çalışması*. (Yayımlanmamış doktora tezi). Ankara Üniversitesi/ Eğitim Bilimleri Enstitüsü, Ankara.

- Erdem Keklik, D. (2014). Değişen Madde Fonksiyonunu Belirlemede Mantel- Haenszel ve Lojistik Regresyon Tekniklerinin Karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5 (2), 12- 25.
- Gierl, M. J., Jodoin, M. G., and Ackerman, T. A. (2000). *Performance of Mantel-Haenszel, Simultaneous Item Bias Test, and Logistic Regression When the Proportion of DIF Items is Large*. The Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, USA.
- Hambleton, R. K., Swaminathan, H., Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. London: Sage Publications.
- Hauck Filho, N., Machado, W. D. L., ve Damásio, B. F. (2014). Effects of Statistical Models and Items Difficulties on Making Trait-Level Inferences: A Simulation Study. *Psicologia: Reflexão e Crítica*, 27(4), 670- 678.
- Hidalgo, M. D., Lopez-Pina, J. A. (2004). Differential Item Functioning Detection and Effect Size: A Comparison Between Logistic Regression and Mantel-Haenszel Procedures. *Educational and Psychological Measurement*, 64(6), 903-915.
- Jodoin, M. G., Gierl, M. J. (2001). Evaluating Type I Error and Power Rates Using an Effect Size Measure With the Logistic Regression Procedure for DIF Detection. *Applied Psychological Measurement*, 14 (4), 329- 349.
- Kan, A., Sünbül, Ö., Ömür, S. (2013). 6. - 8. Sınıf Seviye Belirleme Sınavları Alt Testlerinin Çeşitli Yöntemlere Göre Değişen Madde Fonksiyonlarının İncelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 9 (2), 207- 222.
- Kaya, M. (2017). Psikolojide Eğitimde Ölçme ve Değerlendirme. Mustafa Otrar (Ed.), *Ölçme Sürecinde Aranılan Özellikler: Geçerlik içinde* (s. 154- 198). Ankara: Nobel Yayın.
- Kim, J. (2010). *Controlling Type I Error Rate in Evaluating Differential Item Functioning for Four DIF Methods: Use of Three Procedures for Adjustment of Multiple Item Testing*. (Yayımlanmamış doktora tezi). Georgia State University/ Department of Educational Policy Studies, Atlanta.

- Kim, S., Cohen A. S. (1994). An Investigation of Lord' s Procedure for the Detection of Differential Item Functioning. *Applied Psychological Measurement*, 18 (3), 217- 228.
- Kođar, H. (2018). An Examination of Parametric and Nonparametric Dimensionality Assessment Methods with Exploratory and Confirmatory Mode. *Journal of Education and Learning*, 7 (3), 148- 158.
- Kristjansson , E., Aylesworth, R., Mcdowell, I., Zumbo, B. D. (2005). Comparison of Four Methods for Detecting Differential Item Functioning in Ordered Response Items. *Educational and Psychological Measurement*, 65 (6), 935-953.
- Li, Y., Brooks, G. P., Johanson, G. A. (2012). Item Discrimination and Type I Error in the Detection of Differential Item Functioning. *Educational and Psychological Measurement*, 72(5), 847-861.
- Lord, F. M. (2012). *Applications of Item Response Theory to Practical Testing Problems*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Magis, D., Beland, B., Raiche, G. (2018). difR: Collection of methods to detect dichotomous differential item functioning (DIF). R Package Version 5.0. <https://cran.r-project.org/web/packages/difR/difR.pdf>.
- Magis, D., Beland, B. Tuerlinckx, F., De Boeck P. (2010). A General Framework and An R Package for The Detection of Dichotomous Differential Item Functioning. *Behavior Research Methods*, 42 (3), 847- 862.
- Magis, D., De Boeck, P. (2012). A Robust Outlier Approach to Prevent Type I Error Inflation in Differential Item Functioning. *Educational and Psychological Measurement*, 72(2), 291-311.
- Mantel, N., Haenszel, W. (1959). Statistical Aspects of The Analysis of Data From Retrospective Studies of Disease. *J. Natl. Cancer Inst.*, 22 (4), 719-748.
- Narayanan, P., Swaminathan, H. (1994). Performance of The Mantel-Haenszel and Simultaneous Item Bias Procedures for Detecting Differential. *Applied Psychological Measurement*, 18(4), 315-328.

- Narayanan, P., Swaminathan, H. (1996); Identification of Items that Show Nonuniform DIF. *Applied Psychological Measurement*, 20 (3), 257- 274.
- Osterlind, S. J. (1983). *Test Item Bias*. California: Sage Publications.
- Osterlind, S. J., Everson, H. T. (2009). *Differential Item Functioning*. CA: Sage Publications.
- Price, E. A. (2014). *Item Discrimination, Model-Data Fit, and Type I Error Rates in DIF Detection Using Lord's χ^2 , the Likelihood Ratio Test, and the Mantel-Haenszel Procedure*. (Yayımlanmamış doktora tezi). Ohio University, Columbus.
- Raju, N. S. (1988). The Area Between Two Item Characteristic Curves. *Psychometrika*, 23 (4), 495- 502.
- Rizopoulos, D. (2018). Latent trait models under IRT. R Package Version 1.1-1. <https://cran.r-project.org/web/packages/ltm/ltm.pdf>.
- Rogers, H. J., Swaminathan, H. (1993). A Comparison of Logistic Regression and Mantel-Haenszel Procedures for Detecting Differential Item Functioning. *Applied Psychological Measurement*, 17 (2), 105- 116.
- Roussos, L. A., and Stout, W. F. (1996). Simulation Studies of the Effects of Small Sample Size and Studied Item Parameters on SIBTEST and Mantel-Haenszel Type I Error Performance. *Journal of Educational Measurement*, 33 (2), 215-230.
- Samuelsen, K. M. (2005). *Examining Differential Item Functioning From a Latent Class Perspective*. (Yayımlanmamış doktora tezi). University of Maryland/ Department of Measurement, Statistics and Evaluation, ABD.
- Shepard, L., Camilli, G., Averill, M. (1981). Comparison of Procedures for Detecting Test-Item Bias with Both Internal and External Ability Criteria. *Journal of Educational Statistics*, 6 (4), 317- 375.
- Steinberg, L., Thissen, D. (2006). Using Effect Sizes for Research Reporting: Examples Using Item Response Theory to Analyze Differential Item Functioning. *Psychol Methods*, 11 (4), 402–415.

- Sünbül, Ö., Ömür Sünbül, S. (2016). Değişen Madde Fonksiyonunun Belirlenmesinde Kullanılan Yöntemlerde I. Tip Hata ve Güç Çalışması. *Elementary Education Online*, 15 (3), 882- 897.
- Swaminathan, H. ve Rogers, H. J. (1990). Detecting Differential Item Functioning Using Logistic Regression Procedures. *Journal of Educational Measurement*, 27 (4), 361-370.
- Vaughn, B. K., Wang, Q. (2010). DIF Trees: Using Classification Trees to Detect Differential Item Functioning. *Educational and Psychological Measurement*, 70(6) 941–952.
- Wang, W., Tay, L., Drasgow, F. (2013). Detecting Differential Item Functioning of Polytomous Items for an Ideal Point Response Process. *Applied Psychological Measurement*, 37 (4), 316- 335.
- Wiberg, M. (2007). *Measuring and Detecting Differential Item Functioning in Criterionreferenced Licensing Test: A Theoretic Comparison of Methods*. (Yayımlanmamış doktora tezi). Umea University, Sweden.
- Zumbo, B. D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. Ottawa, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- Zwick, R., Thayer, D. T., Mazzeo, J. (1997). *Describing and Categorizing DIF in Polytomous Items*. Princeton, New Jersey: Educational Testing Service.

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Ayşe BİLİCİOĞLU

Doğum Tarihi : 02.10.1994

Doğum Yeri : Karabük

Eğitim Durumu

Lisans Öğrenimi : Kastamonu Üniversitesi 3,47/ 4

İlköğretim Bölümü, Sınıf Öğretmenliği Anabilim Dalı

Yüksek Lisans : Akdeniz Üniversitesi

Eğitim Bilimleri Enstitüsü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı (2016- 2019)

Bildiği Yabancı Diller: İngilizce

Bilimsel Faaliyetler

Katılan Bilimsel Kongre/ Sempozyum ve Bilimsel Toplantılar:

5. International Eurasian Educational Research Congress(EJER), Antalya, Mayıs 2018

2. Uluslararası Çağdaş Eğitim Araştırmaları Kongresi, Muğla, Eylül 2017

26. Uluslararası Eğitim Bilimleri Kongresi(UEBK), Antalya, Nisan 2017

Dergilerde Yayımlanan Makaleler:

Bilicioğlu A., Yılmaz K. , "Öğrencilerin Sınav Kaygısı, Fen'e Yönelik İlgi ve Ebeveyn Desteği Değişkenleri Üzerine Uluslararası Bir Karşılaştırma: Türkiye- Singapur", Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, cilt.17, ss.1-12, 2017

Hakemli Kongre/Sempozyum Bildiri Kitaplarında Yer Alan Yayınlar:

Bıçak B., Biliciođlu, A., Ayvalli M., "Öđretmen adaylarının bilimsel araştırma okuryazarlık düzeyleri", 5. International Eurasian Educational Research Congress(EJER), ANTALYA, TÜRKİYE, 2-5 Mayıs 2018.

Biliciođlu A., Yılmaz K. , "PISA 2015 Sonuçlarına göre Öğrencilerin Sınav Kaygısı, Fen'e Yönelik İlgi ve Ebeveyn Desteđi Deđişkenleri Üzerine Uluslararası Bir Karşılaştırma: Türkiye-Singapur", 26. Uluslararası Eğitim Bilimleri Kongresi(UEBK), ANTALYA, TÜRKİYE, 20-23 Nisan 2017.

Biliciođlu A., Dođan E., "6.Sınıf Matematik Dersi Tam Sayılar Öğrenme Alanında Başarı Testinin Geliştirilmesi ve Madde Tepki Kuramı ile Klasik Test Kuramına göre Madde Parametrelerinin Karşılaştırılması", 2. Uluslararası Çađdaş Eğitim Araştırmaları Kongresi, MUĞLA, TÜRKİYE, 28 Eylül - 1 Ekim 2017.

İş Deneyimi

Projeler : Erasmus+ Öğrenci Deđişim programı kapsamında 2014/15 yılı güz dönemi Portekiz' de öğrenim görüldü.

Çalıştığı Kurumlar : 2017- 2019 Akdeniz Üniversitesi, Eğitim Fakültesi, Eğitimde Ölçme ve Deđerlendirme Anabilim Dalı (Araştırma Görevlisi)

İletişim

Cep Telefonu : 506 7968058

E-posta : biliciođluayse1@gmail.com

İNTİHAL RAPORU

DEĞİŞEN MADDE FONKSİYONU BELİRLEMEDE KULLANILAN TEKNİKLERİN İ. TIP HATA VE İSTATİSTİKSEL GÜÇ ORANLARININ KARŞILAŞTIRILMASI

ORIGINALITY REPORT

9%

SIMILARITY INDEX

7%

INTERNET SOURCES

5%

PUBLICATIONS

%

STUDENT PAPERS

PRIMARY SOURCES

1	ilkogretim-online.org.tr Internet Source	2%
2	acikarsiv.ankara.edu.tr Internet Source	2%
3	ERDEM KEKLİK, Devrim. "Değişen Madde Fonksiyonunu Belirlemede Mantel-Haenszel ve Lojistik Regresyon Tekniklerinin Karşılaştırılması", Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 2014. Publication	1%
4	egitimvebilim.ted.org.tr Internet Source	1%
5	ÖMÜR SÜNBÜL, Seçil and KAN, Adnan. "Bilişsel Tanı Modellerinde Parametre Kestirimini ve Sınıflama", Hacettepe Üniversitesi, 2016. Publication	<1%

adumilas.adu.edu.tr

6

Internet Source

<1%

7

TUNÇ, Emine Burcu and KUTLU, Ömer. "İki ve Çok Kategorili Puanlanan Maddelerde Değişen Madde Fonksiyonlarının Karşılaştırılması", Başkent Üniversitesi Eğitim Fakültesi, 2018.

Publication

<1%

8

www.dergipark.ulakbim.gov.tr

Internet Source

<1%

9

DOĞAN, Nuri and ÖĞRETMEN, Tuncay. "Değişen Madde Fonksiyonunu Belirlemede Mantel-Haenszel, Ki-Kare ve Lojistik Regresyon Tekniklerinin Karşılaştırılması", Kaligrafi Yayıncılık, 2008.

Publication

<1%

10

KAN, Adnan, SÜNBÜL, Önder and ÖMER, Seçil. "6. - 8. sınıf seviye belirleme sınavları alt testlerinin çeşitli yöntemlere göre değişen madde fonksiyonlarının incelenmesi", Mersin Üniversitesi, 2013.

Publication

<1%

11

hacettepe.academia.edu

Internet Source

<1%

12

d.ogren-sen.com

Internet Source

<1%

UZUN, Nezaket Bilge and GELBAL, Selahattin.

13

"PISA Fen Başarı Testinin Madde Yanlılığının
Kültür ve Dil Açısından İncelenmesi",
Kastamonu Üniversitesi, 2017.

Publication

<1%

14

AKIN ARIKAN, Çiğdem, UĞURLU, Seçil and
ATAR, Burcu. "MIMIC, SIBTEST, Lojistik
Regresyon ve Mantel-Haenszel Yöntemleriyle
Gerçekleştirilen DMF ve Yanlılık Çalışması",
Hacettepe Üniversitesi, 2016.

Publication

<1%

15

www.researchgate.net

Internet Source

<1%

16

KARTAL, Seval KULA and DİRLİK, Ezgi MOR.
"GEÇERLİK KAVRAMININ TARİHSEL GELİŞİMİ
VE ", Abant İzzet Baysal Üniversitesi, 2016.

Publication

<1%

17

dergipark.gov.tr

Internet Source

<1%

18

TARIM, Kamuran and BULUT, M.Sencer.
"Okulöncesi öğretmenlerinin matematik ve
matematik öğretimine ilişkin algı ve tutumları",
Çukurova Üniversitesi, 2006.

Publication

<1%

19

ÖZKAN, Yeşim ÖZER. "Üniversite Mezuniyet
Notları ve Dönüşüm Tablolarında ", Şükrü
Oğuz Özdamar, 2016.

Publication

<1%

20

ÖMÜR SÜNBL, Seçil and AŞİRET, Semih.
"Madde Tepki Kuramına Dayalı Madde-Uyum
İndekslerinin I.Tip Hata ve Güç Oranlarının
İncelenmesi", Eğitimde ve Psikolojide Ölçme ve
Değerlendirme Dergisi, 2017.

Publication

<1%

Exclude quotes Off
Exclude bibliography Off

Exclude matches Off

Causes
1