

T.C.

MERSİN ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANABİLİM DALI

KLASİK TEST VE MADDE TEPKİ KURAMLARINA  
DAYALI DEĞİŞEN MADDE FONKSİYONU  
BELİRLEME TEKNİKLERİNİN FARKLI PUANLAMA  
DURUMLARINDA İNCELENMESİ

DOKTORA TEZİ

Hüseyin SELVİ

Mersin, 2013

T.C.

MERSİN ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANABİLİM DALI

KLASİK TEST VE MADDE TEPKİ KURAMLARINA  
DAYALI DEĞİŞEN MADDE FONKSİYONU  
BELİRLEME TEKNİKLERİNİN FARKLI PUANLAMA  
DURUMLARINDA İNCELENMESİ

DOKTORA TEZİ

Hüseyin SELVİ

Danışman

Yrd. Doç. Dr. Devrim ALICI

Mersin, 2013

## KABUL VE ONAY

Mersin Üniversitesi Eğitim Bilimleri Enstitüsü Müdürlüğü'ne,

Hüseyin SELVİ tarafından hazırlanan 'Klasik Test ve Madde Tepki Kuramlarına Dayalı Değişen Madde Fonksiyonu Belirleme Tekniklerinin Farklı Puanlama Durumlarında İncelenmesi' başlıklı bu çalışma, jürimiz tarafından Eğitimde Ölçme ve Değerlendirme Bilim Dalı'nda DOKTORA TEZİ olarak kabul edilmiştir.

Başkan..........(İmza)

Prof. Dr. Adnan ERKUŞ

Üye..........(İmza)

Doç. Dr. Duygu ANIL

Üye..........(İmza)

Doç. Dr. Tuncay ÖĞRETMEN

Üye..........(İmza)

Yrd. Doç. Dr. Devrim ALICI (Danışman)

Üye..........(İmza)

Yrd. Doç. Dr. Önder SÜNBÜL

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.



## TEŐEKKÜR

Tezimin bütn aŐamalarında desteęini esirgemeyen, yol gsteren ve benimle birlikte emek sarf eden deęerli bilim hocam Yrd. Doę. Dr. Devrim ALICI'ya sonsuz teŐekkrlerimi sunarım.

Doktora eęitimimin en baŐından itibaren bilgisini, nderlięini ve sevgisini esirgemeyen deęerli bilim hocam Prof. Dr. Adnan ERKUŐ'a ayrıca teŐekkr ederim.

Tez alıŐmamı gerekleŐtirmek iin ihtiya duyduęum verilere, doktora tez alıŐması kapsamında elde ettięi verileri kullanmama izin vererek ulaŐmamı saęlayan Yrd. Doę. Dr. Gksu GZEN ITAK'a teŐekkrlerimi sunarım.

AraŐtırmam boyunca bana destek olan ve yol gsteren Doę. Dr. Adnan KAN, Doę. Dr. Tuncay OęRET MEN, Yrd. Doę. Dr. nder SNBL, ArŐ. Gr. Seil MR SNBL ve tm deęerli bilim hocalarıma teŐekkr ederim.

Ayrıca destekleri ile zor dnemlerimde hep yanımda olan aileme ok teŐekkr ederim.

## ÖZET

# KLASİK TEST VE MADDE TEPKİ KURAMLARINA DAYALI DEĞİŞEN MADDE FONKSİYONU BELİRLEME TEKNİKLERİNİN FARKLI PUANLAMA DURUMLARINDA İNCELENMESİ

HÜSEYİN SELVİ

Doktora Tezi

Eğitim Bilimleri Enstitüsü

Yrd. Doç. Dr. Devrim ALICI (Danışman)

Kasım 2013

97 Sayfa

Bu araştırmanın amacı, farklı kuramlara dayalı farklı DMF belirleme tekniklerinin, farklı puanlama durumlarındaki işleyişlerini incelemektir.

Bu amaç doğrultusunda, ikili ve ağırlıklı puanlamaya uygun bir Sözel Akıl Yürütme Yeteneği Testinin 1593 kişiye uygulanmasıyla elde edilen veriler, Mantel-Haenszel/Mantel, Standardizasyon, Lojistik Regresyon ve Olabilirlik

Oran Testi teknikleriyle incelenmiş, kullanılan her bir teknik ve puanlama durumu için DMF gösteren maddeler ve DMF'li madde sayıları belirlenmiştir.

Elde edilen bulgulardan, KTK'na dayalı DMF belirleme tekniklerinin genel olarak kendi içlerinde tutarlı sonuçlar verdiği ve bu sonuçların birbirinden anlamlı derecede farklılaşmadığı; ancak ikili puanlama durumundan uzman yargısına dayalı ağırlıklı puanlama ve deneysel ağırlıklı puanlama durumuna gidildikçe bu uyumun bozulduğu sonucuna ulaşılmıştır. Ayrıca farklı puanlama durumlarında MTK ve KTK'na dayalı teknikler arası uyumun düşük olduğu ve elde edilen sonuçların genel olarak birbirinden anlamlı derecede farklılaştığı gözlenmiştir.

**Anahtar Kelimeler:** Değişen Madde Fonksiyonu, Mantel-Haenszel, Standardizasyon, Lojistik Regresyon, Olabilirlik Oran Testi

## **ABSTRACT**

# **INVESTIGATION OF DIFFERENTIAL ITEM FUNCTIONING PROCEDURES BASED ON CLASICAL TEST AND ITEM RESPONSE THEORY AT DIFFERENT SCORING SITUATIONS**

**HÜSEYİN SELVİ**

PhD Thesis

Institute of Education Sciences

Asst. Prof. Dr. Devrim ALICI (Supervisor)

November 2013

97 Pages

The purpose of this study is examining different differential item functioning procedures based on different theories at different scoring situations.

For this purpose, the data obtained from 1593 people by a verbal reasoning test is analyzed with Mantel-Haenszel/Mantel, Standardization, Logistic Regression and Likelihood Ratio Test procedures at different scoring situations. Than items and number of items which has DIF are identified.

From the results of analyses, DIF procedures based on Clasical Test Theory are producing generally consistent results that are not differing significantly from each other. But this harmony is disrupted when the scoring

method goes from dichotomous case to weighted case. In addition the harmony between procedures based on Classical Test Theory and Item Response Theory at different scoring situations is observed very low. And results are differing significantly from each other.

**Keywords:** Differential Item Functioning, Mantel-Haenszel, Standardization, Logistic Regression, Likelihood Ratio Test.



# İÇİNDEKİLER

KABUL VE ONAY.....	i
TEŞEKKÜR.....	ii
ÖZET.....	iii
ABSTRACT.....	v
İÇİNDEKİLER.....	vii
TABLolar DİZİNİ.....	x
ŞEKİLLER DİZİNİ.....	vii
SİMGELER VE KISALTMALAR DİZİNİ.....	viii
BÖLÜM I.....	1
GİRİŞ.....	1
I.1. Değişen Madde Fonksiyonu Belirlemede Kullanılan Teknikler.....	4
I.1.1. Klasik Test Kuramı'na Dayalı Teknikler .....	4
I.1.2. Madde Tepki Kuramı'na Dayalı Teknikler .....	17
I.3. İlgili Araştırmalar.....	32
I.4. Araştırmanın Amacı ve Önemi.....	39
I.5. Problem Cümlesi.....	41
I.5.1) Alt Problemler.....	41
I.6. Sayılıtlar.....	42
I.7. Sınırlılıklar.....	42

BÖLÜM II.....	43
YÖNTEM.....	43
II.1. Araştırmanın Türü.....	43
II.2. Çalışma Grubu.....	43
II.3. Ölçme Aracı.....	45
II.4. Araştırmanın Değişkenleri.....	48
II.5. Verilerin Analizi.....	48
BÖLÜM III.....	55
BULGULAR.....	55
III.1. Alt Problemler Doğrultusunda Elde Edilen Bulgular.....	56
III.1.1. Birinci Alt Problem Doğrultusunda Elde Edilen Bulgular.....	56
III.1.2. İkinci Alt Problem Doğrultusunda Elde Edilen Bulgular.....	62
III.1.3. Üçüncü Alt Problem Doğrultusunda Elde Edilen Bulgular.....	68
III.1.4. Dördüncü Alt Problem Doğrultusunda Elde Edilen Bulgular.....	73
BÖLÜM IV.....	76
TARTIŞMA VE YORUM.....	76
IV.1. Öneriler.....	82
KAYNAKÇA.....	84
EKLER.....	92
ÖZGEÇMİŞ.....	97

## TABLolar DİZİNİ

<b>Tablo 1:</b> Lojistik Regresyon TekniĐinde DeĐişen Madde Fonksiyonu Düzeyi İin $R^2$ Ölütleri.....	8
<b>Tablo 2:</b> Kontinjensi Tablosu .....	9
<b>Tablo 3:</b> Verilerin eřitli DeĐişkenlere Göre Elde Edilen Bazı Betimsel DeĐerleri.....	44
<b>Tablo 4:</b> Teste İlişkin Elde Edilen Betimsel İstatistikler.....	45
<b>Tablo 5:</b> Maddelerin (Seeneklerin) AĐırlıklı Puan DaĐılımları .....	47
<b>Tablo 6:</b> McNemar Testi İin Kullanılan Ki-Kare Testi Tablosu.....	53
<b>Tablo 7:</b> Odak ve Referans Grupları Aısından Araştırma Verilerine Ait Bazı Betimsel İstatistikler.....	55
<b>Tablo 8:</b> İkili Puanlama Durumunda, M-H/Mantel, Standardizasyon ve Lojistik Regresyon Teknikleriyle Elde Edilen Bulgular.....	57
<b>Tablo 9:</b> UYDA Puanlama Durumunda M-H/Mantel, Standardizasyon ve Lojistik Regresyon Teknikleriyle Elde Edilen Bulgular.....	59
<b>Tablo 10:</b> DA Puanlama Durumunda M-H/Mantel, Standardizasyon ve Lojistik Regresyon Teknikleriyle Elde Edilen Bulgular.....	61
<b>Tablo 11:</b> İkili Puanlama Durumunda Olabilirlik Oran Testi TekniĐiyle Elde Edilen Bulgular.....	63
<b>Tablo 12:</b> UYDA Puanlama Durumunda Olabilirlik Oran Testi TekniĐiyle Elde Edilen Bulgular.....	65
<b>Tablo 13:</b> DA Puanlama Durumunda Olabilirlik Oran Testi TekniĐiyle Elde Edilen Bulgular.....	67

<b>Tablo 14:</b> İkili, UYDA ve DA Puanlama Durumlarında, Saptanan DMF’li Maddeler Açısından KTK ve MTK’na Dayalı DMF Belirleme Teknikleri Arası Farka İlişkin Cochran’s Q testi Sonuçları .....	68
<b>Tablo 15:</b> KTK ve MTK’na Dayalı DMF Belirleme Teknikleri Kullanıldığında, Saptanan DMF’li Maddeler Açısından İkili, UYDA ve DA Puanlama Durumları Arası Farka İlişkin McNemar ve Cochran’s Q Testi Sonuçları .....	69
<b>Tablo 16:</b> KTK ve MTK’na Dayalı DMF Belirleme Teknikleriyle, İkili, UYDA ve DA Puanlama Durumlarında Belirlenen DMF’li Maddeler Arasındaki Uyum Yüzdeleri .....	71
<b>Tablo 17:</b> KTK ve MTK’na Dayalı DMF Belirleme Teknikleriyle, İkili, UYDA ve DA Puanlama Durumlarında Belirlenen DMF’li Madde Sayıları .....	74
<b>Tablo 18:</b> KTK ve MTK’na Dayalı DMF Belirleme Teknikleriyle, İkili, UYDA ve DA Puanlama Durumlarında Belirlenen DMF’li Madde Sayılarına İlişkin Birleştirilmiş Küme .....	75

## ŒEKİLLER DİZİNİ

<b>Œekil 1:</b> Madde Karakteristik Eđrileri Arası Alan.....	18
--	----

## SİMGELER VE KISALTMALAR DİZİNİ

<b>DMF:</b>	Değişen Madde Fonksiyonu
<b>KTK:</b>	Klasik Test Kuramı
<b>MTK:</b>	Madde Tepki Kuramı
<b>MKF:</b>	Madde Karakteristik Fonksiyonu
<b>MKE:</b>	Madde Karakteristik Eğrisi
<b>ACM:</b>	Ağırlıklandırılmış Cevap Modeli
<b>KPM:</b>	Kısmi Puan Modeli
<b>DÖM:</b>	Dereceleme Ölçeği Modeli
<b>STM:</b>	Sınıflamalı Tepki Modeli
<b>UYDA:</b>	Uzman Yargısına Dayalı Ağırlıklı Puanlama
<b>DA:</b>	DeneySEL Ağırlıklı Puanlama
<b>M-H/Mantel:</b>	Mantel Haenszel/Mantel
<b>LR:</b>	Lojistik Regresyon
<b>MTK-OO:</b>	Olabilirlik Oran Testi Tekniği
<b>TB:</b>	Tek Biçimli
<b>TBO:</b>	Tek Biçimli Olmayan

## BÖLÜM I

### GİRİŞ

Bilimin olmazsa olmaz ölçütlerinden biri ölçülebilirliktir. Bu nedenle bilim dallarındaki gelişmeler ölçme bilimindeki gelişmeler ile paralellik göstermektedir. Nitekim bilim dalları kullandıkları ölçme yöntemleri, teknikleri, araçları vb. geliştirebildiği ölçüde ilerleyebilmiştir. Thorndike'in 'Doğada bir şey varsa, bir miktar vardır' ve Mc.Call'ın 'Bir şey bir miktar varsa, o şey ölçülebilir' ifadeleri de göz önünde bulundurulduğunda, ölçmenin bütün bilim dalları açısından önemi ortaya çıkmaktadır (Akt: Erkuş, 2003).

Bu önemin farkında olan uzmanlar, farklı bilim dallarındaki ölçme işlemlerinde kullanılmak üzere pek çok ölçme aracı geliştirmiştir ve geliştirmeye de devam etmektedir. Geliştirilen bu ölçme araçlarının nitelikli ölçümler yapabilmesi için ise başlıca iki temel gereksinimi karşılaması gerekmektedir. Bunlar; güvenirlik ve geçerliktir.

Güvenirlik, bir ölçme aracının hatasız ölçme yapabilme yeterliği olarak tanımlanmakta ve geçerlik için bir ön koşul oluşturmaktadır. Geçerlik ise genel olarak ilgili ölçme aracının, ölçülmek istenilen niteliği, başka niteliklere karıştırmadan ölçebilme yeterliği olarak tanımlanmaktadır (Crocker ve Algina, 1986; Magnusson, 1967; Murphy ve Davidshofer, 2005; Anastasi ve Urbina, 1997; Thorndike, 1982; Aiken, 2000).

Eğitim ve psikoloji bilimleri açısından bir ölçme aracının güvenirliği; test uzunluğu, grup homojenliği, maddelerin ortalama güclüğü, puanlayıcılar arası tutarlılık vb. değişkenlerden etkilenmektedir. Geçerlik ise güvenirliği de içinde

barındıran daha geniş bir kavram olup, ölçülmek istenilen değişkenin yapısı, ölçme aracının ilgili konu alanını kapsama düzeyi, kullanım amacı, yanlılık vb. pek çok değişkenden etkilenmektedir (Crocker ve Algina, 1986; Magnusson, 1967; Murphy ve Davidshofer, 2005; Anastasi ve Urbina, 1997; Thorndike, 1982; Aiken, 2000). Diğer ifadeyle, seçkisiz hata kaynakları ölçme sonuçlarının güvenilirliğini etkileyen bir sorun olarak ortaya çıkarken; sistematik hata kaynakları geçerliği etkileyen bir sorun olarak ortaya çıkmaktadır. Geçerliği etkileyen değişkenler incelendiğinde ise, yanlılığın bu değişkenler arasında önemli bir yer tuttuğu görülmektedir Bu anlamda yanlılık, ölçme aracının geçerliğini etkileyen sistematik bir hata kaynağı olarak tanımlanabilir (Osterling, 1983).

Ölçülmek istenilen niteliğe başka niteliklerin karıştığı bazı durumlarda, farklı alt gruplarda bulunan bireyler bu durumdan sistematik şekilde etkilenmektedir. Geçerliği olumsuz yönde etkileyen ve yanlılık olarak nitelenen bu değişimin ilk farkına varan Alfred Binet'dir. Binet 1910 yılında, düşük sosyoekonomik yapıdan gelen bireylerin, bazı test maddelerine verdikleri tepkilerin mental yeteneklere ek olarak kültürel yapıdan da etkilendiğini fark etmiş ve bu tür maddeleri dikkat, evdeki eğitim, dil bilgisi vb. farklı alt kategorilere ayırmıştır (Camilli ve Shepard, 1994).

1970'li yıllardan sonra ise yanlılığın geçerlik üzerinde ciddi sıkıntı oluşturduğu ve hemen her konuda verilen kararların isabetsiz olmasına neden olduğu görülmüştür (Holland ve Wainer, 1993). Bu nedenle farklı alt gruplarda farklı şekilde işleyen maddelerin (yanlılığın) belirlenmesine yönelik çalışmalar, geçerlik çalışmaları içerisinde yer almaya başlamış ve gelişen bilgisayar altyapısı-kuramlarla birlikte kullanılan teknikler de oldukça kapsamlı hale gelmiştir. Günümüzde Değişen Madde Fonksiyonu (DMF) adını alan bu çalışmaların amacının, işleyişinin ve mantığının daha iyi anlaşılabilmesi için yanlılık, madde



etkisi, test etkisi, madde yanlılığı, test yanlılığı ve Değişen Madde Fonksiyonu terimlerinin ayrıca açıklanması yararlı görülmektedir.

*Yanlılık*, Angoff (1993) tarafından; ‘aynı yetenek düzeyine sahip farklı alt gruplarda bulunan bireylerin ilgili test maddesine doğru cevap verme olasılıkları arasındaki fark’ olarak tanımlanmıştır.

*Madde-test etkisi ise*, madde ve test istatistikleri yardımıyla ilgili maddenin ya da testin farklı alt gruplarda farklı şekilde işlediğinin belirlenmesi durumu olarak açıklanmaktadır. Ancak bu durum, maddenin ya da testin yanlı olduğu anlamına *gelmemektedir*. Diğer ifadeyle, ilgili maddenin ya da testin farklı alt gruplarda farklı şekilde işlediğinin belirlenmesi durumunda, bu farklılıkların gerçekten ölçülmek istenen niteliğin kendisinden mi, yoksa sistematik bir değişimden mi kaynaklandığının ayrıca araştırılması gerekmektedir. Bu aşamada ise genel olarak uzman görüşlerine başvurulmaktadır. Eğer karşılaşılan grup farklılıkları ölçülmek istenilen nitelikten kaynaklanmıyorsa, bu durum ‘*madde-test yanlılığı*’ olarak adlandırılmaktadır. Örneğin; bireylerin matematik başarısını ölçmeye yönelik olarak geliştirilmiş bir teste katılan yabancı uyruklu öğrencilerin düşük not almalarına, matematik başarısı yanında kelime bilgilerindeki yetersizlik de neden olmuş olabilir. Böyle bir durumun varlığı, ölçülmek istenilen niteliğin başka niteliklerle karıştırılması nedeniyle madde-test yanlılığını doğurmakta ve geçerliği olumsuz yönde etkilemektedir.

Madde-test yanlılığının belirlenmesine yönelik yapılan çalışmalar ise *Değişen Madde Fonksiyonu* belirleme çalışmaları olarak adlandırılmaktadır (Osterling, 1983).

Değişen Madde Fonksiyonu belirlemede, Klasik Test Kuramı (KTK) ve Madde Tepki Kuramı’na (MTK) dayalı olarak pek çok teknik geliştirilmiştir.

## **I.1. Değişen Madde Fonksiyonu Belirlemede Kullanılan Teknikler**

### **I.1.1 Klasik Test Kuramına Dayalı Teknikler**

Klasik Test Kuramı, temelde bireylerin gözlenen puanlarından (ham puan) hareketle gerçek puanlarının kestirilmesine dayanmakta ve bir teste ilişkin gözlenen puanın (ham puan) gerçek ve hata puanları toplamından oluştuğunu varsaymaktadır. Buradaki *Gözlenen puan*, yapılan ölçme işlemi sonucunda elde edilen puanı temsil etmekte olup, içinde gerçek ve hata puanlarını barındırmaktadır. *Gerçek puan*, ölçülen özelliğe ait olan ve içinde hata puanı barındırmayan puandır. Bu puan hipotetik bir puan olup, operasyonel olarak tanımlanamamakta; bu nedenle de doğrudan ya da dolaylı olarak ölçülememektedir. *Hata puanı ise*, yapılan her ölçme işlemine bir miktar karışan ve kaynağı bilinemeyen puanı temsil etmektedir. Bu hatalar, ölçme sonuçlarına karışan sabit ve sistematik hataların dışında kalan seçkisiz (random) hataları ifade etmektedir (Baykul, 2000).

Klasik Test Kuramına dayalı olarak Değişen Madde Fonksiyonunu belirlemek amacıyla geliştirilen teknikler genel olarak dış ve iç ölçüte dayalı teknikler olmak üzere ikiye ayrılır. Dış ölçüte dayalı teknikler, yordama geçerliği (predictive validity) üzerine kuruludur ve ölçütün kestirilmesinde kullanılan regresyon denkleminin bazı alt gruplar için farklı şekilde işlemesi, test yanlılığı olarak kabul edilmektedir. Dış ölçüte dayalı teknikler incelendiğinde, bu tekniklerin genel olarak maddeler bazında yetersiz kaldığı görülmektedir. Buna ek olarak, her zaman bir dış ölçüt bulunamaması ve ölçütün kendisinin de yanlı olması gibi sıkıntılar bulunmaktadır. Bu nedenlerle dış ölçüte dayalı teknikler günümüzde pek kullanılmamaktadır. İç ölçüte dayalı tekniklerde ise, farklı alt gruplarda bulunan bireyler, testten aldıkları toplam puanlar açısından eşleştirilmekte ve maddelerdeki performansları izlenmektedir. İç ölçüte dayalı tekniklerden bazıları, Dönüştürülmüş Madde Güçlüğü İndeksi, Altın Kural Prosedürü, Varyans Analizine Bağlı Teknik, Korelasyonel Teknik, Kontinjensi

Tablosu Teknikleri ve Lojistik Regresyon Tekniđi řeklinde sıralanabilir. Klasik Test Kuramına dayalı olarak geliřtirilen bu teknikler ařađıda detaylı olarak açıklanmıřtır.

*Dönüřtürülmüř Madde Güçlüđü İndeksi (DMG) (Transformed Item Difficulty Index)*. Bu teknikte her bir madde için referans ve odak gruplarından ayrı ayrı elde edilen madde güçlük ( $p$ ) deđerleri üzerinden hareket edilmektedir. Burada kullanılan odak grup terimi genellikle azınlık (dezavantaj sahibi olan) grubu; referans grup terimi ise çođunluk (avantaj sahibi olan) grubu ifade etmektedir (Santelices ve Wilson, 2012). DMG tekniđinde aynı madde için grup performanslarındaki deđiřimi incelemek amacıyla her iki grup için ayrı ayrı hesaplanan ‘1-p’ deđerleri, öncelikle normalleřtirilmiř z deđerlerine; elde edilen bu z deđerleri ise ortalaması 13 ve standart sapması 4 olan delta ölçeđine dönüřtürülmektedir. Elde edilen bu deđerler arasındaki farkın büyümesi ise (0’dan sapma), DMF varlıđına yönelik yorum yapılmasını gerektirmektedir.

Bu teknikte DMF göstergesi olarak yalnızca ‘ $p$ ’ deđerlerinin kullanılması; ‘ $p$ ’ deđerlerinin ortalama grup farklılıklarından (geçerlik kanıtı olan gerçek fark) ve madde ayırt ediciliđinden ( $r_{jx}$ ) etkilenmesi nedeniyle sıkıntı oluřturmaktadır. Bu sıkıntıya çözüm bulmak amacıyla Angoff (1982), bireylerin bir dıř ölçüt yardımıyla eřleřtirilmesini ve DMG indekslerinin de bu řekilde hesaplanmasını önermiřtir. Ancak her zaman bir dıř ölçüt bulmak mümkün olmamaktadır. Ayrıca dıř ölçüt kullanımının da ölçütün kendisinin yanlı olması, güvenilirliđinin düşük olması vb. farklı sıkıntıları bulunmaktadır (akt: Camilli ve Shepard, 1994).

*Altın Kural Prosedürü (The Golden Rule Procedure)*. Bu teknik adını bir sigorta řirketinden almaktadır. řirket, hayat ve kaza sigortası yapmadan önce bireyleri bir teste tabi tutmakta ve güvence altına alınmak istenen riskin büyüklüđüne göre sigorta fiyatını belirlemektedir. Burada kullanılan mantık ‘ $p$ ’

değerlerini mümkün olduğu kadar yüksek tutarak (kolay madde) yanlılık varsa bile bunun etkisini en aza indirmek şeklindedir. Şirket, maddeleri 2 gruba ayırmaktadır. 1. gruptaki maddeler her iki grubun da %40'ın ( $p>0.40$ ) altında başarı göstermeyeceği ve her iki grup arasındaki doğru cevaplama oranları farkının %15'i geçmeyeceği şeklindedir. Diğer maddeler ise 2. gruba atılmaktadır. Yansız bir test için de maddeler 1. gruptan seçilmektedir. Ancak bu durum ranj daralması nedeniyle iç tutarlığı ve geçerliği düşürmektedir (Angoff, 1993).

*Varyans Analizine bağlı Değişen Madde Fonksiyonu belirleme tekniği.* İki faktörlü tekrarlı ANOVA modeli kullanılarak, bir faktör 'grup etkisi' ve grup içi faktör de 'madde etkisi' olarak alındığında; grup ortalamaları farkının grup için ana etkiyi, grup madde etkileşiminin de madde etkisini göstermesi beklenmektedir. (Cleary ve Hilton, 1968, akt: Camilli ve Shepard, 1994 ). Ancak Jensen (1984) ve Gordon'un (1987) yapmış oldukları çalışmalar, bu tekniğin bütün maddelerde aynı yönde bulunan ve belirli düzeye ulaşmayan DMF etkisini saptamakta yetersiz kaldığını göstermektedir. Bütün maddelerde madde etkisi olduğunda, grup madde etkileşimi madde etkisi olarak alındığından, madde etkisi olmasına rağmen bu ana etki olarak karşımıza çıkmakta ve grup-madde etkileşimi görülmemektedir. Grup etkisinin madde etkisinden büyük olduğu durumlarda ise madde etkisi grup etkisi tarafından bastırılmakta ve madde etkisi hatalı şekilde ortaya konmaktadır. Ayrıca DMG'de olduğu gibi Varyans Analizi de ' $p$ ' değerlerinden ve ' $r_{jx}$ ' değerlerinden etkilenmektedir (akt: Camilli ve Shepard, 1994).

*Korelasyonel teknikler.* DMF saptanmasında kullanılan bir diğer teknik; her iki grup için hesaplanan ' $p$ ' değerlerini büyükten küçüğe sıralayarak Spearman Sıra Farkları Korelasyonuna bakmaktır. Elde edilen korelasyon değerinin 1.00'e yakın olması DMF olmadığı, 1.00'den uzaklaşması ise DMF olduğu yönünde yorumlanır. Ancak 1'den ne kadarlık bir sapmanın DMF etkisi olarak kabul edileceği yargıci kararına bağlıdır. Korelasyonel tekniklerde gruplar

için 'p' değerleri farklı olsa bile 'r<sub>ix</sub>' değerleri benzer şekilde işleyebilmekte ve korelasyon yüksek çıkabilmektedir. Bu da var olan DMF etkisinin algılanmamasına neden olabilmektedir. Ayrıca 'p' değerleri ortalamasının 0.50'ye yaklaşması bazı maddeler için sapma puanını maksimum yaptığından r<sub>ix</sub> değerleri kolay veya zor maddeye oranla yüksek çıkmaktadır. Bu da kullanılan ölçütün hatalı olmasına neden olmaktadır (Camilli ve Shepard, 1994).

*Lojistik Regresyon tekniği.* Değişen Madde Fonksiyonu belirlemede kullanılan bir diğer teknik olan Lojistik Regresyon analizinde amaç, bireylerin hangi grubun üyesi olduğunu kestirmede kullanılacak bir regresyon denklemi oluşturmaktır. Bu noktadan hareketle bağımsız değişken olarak grup aidiyeti ve toplam test puanı, bağımlı değişken olarak da madde puanı (1,0) alındığında, Lojistik Regresyon analiziyle DMF saptanımında bulunulabilmektedir. Bu bağlamda grup, toplam puan ve kesişim (interaction) parametreleri kestirilmede ve model karşılaştırma stratejileriyle istatistiksel fark test edilmektedir. Maksimum Likelihood (ML) prosedürlerinin kullanıldığı formüller eşitlik 1'de verilmiştir (Swaminathan ve Rogers, 2005).

$$P(u_i = 1) = \frac{\exp \Psi_i}{1 + \exp \Psi_i} \quad (1)$$

$$\Psi_i = \delta + t_1 G_i + t_2 X_i + t_3 (G_i X_i)$$

P(u<sub>i</sub>=1): i bireyinin u maddesine doğru cevap verme olasılığı

$\Psi_i$ : Bir maddeden i bireyi için beklenen log odds değeri,

i: Birey,

$X_i$ : Bireyin toplam puanı,

$G_i$ : Bireyin grubu,

$t_1$ : Odds oranı logaritmalarının birleştirilmiş hali,

$t_2$ : Birey yeteneklerinin eşit olduğunu gösteren parametre,

$t_3$ : Tek biçimli Değişen Madde Fonksiyonu parametresi.

Eşitlik 1 doğrultusunda aşağıdaki 3 modelde bulunan katsayılar ML prosedürleriyle kestirilmekte ve modellerin bir birlerine göre anlamlı fark oluşturup oluşturmadığı olabilirlik (likelihood) oran testiyle incelenmektedir.

1. model:

$$\psi_i = \delta + t_1 G_i + t_2 X_i + t_3 (G_i X_i)$$

2. model:

$$\psi_i = \delta + t_1 G_i + t_2 X_i$$

3. model:

$$\psi_i = \delta + t_2 X_i$$

2. ve 3. model arasında fark çıkması durumunda DMF varlığından söz edilmektedir. 1 ve 2. model arasında fark çıkması durumunda ise Tek Biçimli Olmayan DMF varlığından söz edilmektedir.

Lojistik Regresyon tekniğiyle DMF düzeyini belirlemek üzere araştırmacılar tarafından  $R^2$  değerlerine dayalı çeşitli ölçütler geliştirilmiştir. Bu ölçütlerden biri de Gierl ve ark. (2000) tarafından önerilmiştir. İlgili ölçütler Tablo 1’de sunulmuştur.

**Tablo 1:** Lojistik Regresyon tekniğinde Değişen Madde Fonksiyonu düzeyi için  $R^2$  ölçütleri

Değişen Madde Fonksiyonu Düzeyi	Gierl ve Arkadaşlarının Ölçütleri	Yorumlar
A	$R^2 < 0.035$	DMF yoktur.
B	$0.035 < R^2 < 0.070$	Orta düzeyde DMF vardır.
C	$R^2 \geq 0.070$	Yüksek düzeyde DMF vardır.

*Kontinjensi Tablosu teknikleri.* Bu tekniklerde parametrik olmayan istatistikler kullanılmaktadır. Bu nedenle parametrik testlerdeki sayılıtların karşılanması gerekmemektedir. Burada DMF etkisi, hipotez testleri yardımıyla test edilmekte olup, ne kadarlık bir farka DMF denileceği de böylece istatistiksel olarak karara bağlanmaktadır. Bu tekniklerde madde etkisini test etmenin yanında, miktarı ve düzeyi hakkında da bilgi sahibi olunabilmektedir. Kullanılan tekniklerin hesaplanması oldukça kolay olup, Madde Tepki Kuramı'ndaki gibi büyük örneklemelere ulaşılması gerekmemektedir. Ancak bu tekniklerdeki bazı hesaplamalarda ' $r_{jx}$ ' değeri hesaba katılmamakta ve her iki grup için de ' $r_{jx}$ ' değerlerinin eşit olduğu kabul edilmektedir. Bu da tekniklerin ulaşılması zor bir sayılıtya dayanmasına neden olmaktadır (Osterling, 1983).

Bu teknikleri kullanabilmek için öncelikle gruplar toplam puan açısından eşleştirilmekte ve her madde için veriler ayrı ayrı Tablo 2'deki gibi bir kontinjensi tablosuna yerleştirilmektedir.

**Tablo 2:** Kontinjensi tablosu

		Maddeden Alınan Puan		Toplam
		1	0	
Grup	Referans Grup	$p_{rj}$	$q_{rj}$	1
	Odak Grup	$p_{oj}$	$q_{oj}$	1

Tablo 2 yardımıyla bu tekniklere ait hesaplamalar yapılmakta olup, bu bağlamda kullanılan tekniklerden bazılarında aşağıda değinilmiştir.

a) *Oranlar arası basit farklar tekniği.* Hesaplaması oldukça kolay olan bu teknikte, yetenek düzeyleri bakımından eşleştirilmiş odak ve referans grup bireyler üzerinden hesaplanan ' $p$ ' değerleri arasındaki farka bakılmaktadır. Beklenen değer '0' olup, bu değerden sapmalar DMF varlığı yönünde bilgi

vermektedir. İstatistiksel olarak test edilecek olan hipotez ise eşitlik 2'den de görüleceği üzere, evren oranları arasındaki farkın '0' olduğu şeklindedir.

$$\Delta p_j = p_{R_j} - p_{F_j} \quad (2)$$

$$H_0 : \Pi_{R_j} = \Pi_{F_j}$$

$\Delta p_j$  : İlgili maddenin her iki grup için eşleştirilmiş  $j$  puanı seviyesindeki madde güçlüğü farkı,

$p_{R_j}$ : Referans grup için madde güçlüğü,

$p_{F_j}$ : Odak grup için madde güçlüğü.

Bu tekniklerde ' $p$ ' değerleri yerine odds oranları da kullanılabilir.  $\Delta p_j$  değerlerinden alınan bilgilere ek olarak odds oranlarıyla madde etkisinin düzeyi hakkında da bilgi sahibi olunabilmektedir. Odds oranları her bir grup için ayrı ayrı 3 no'lu eşitlikte yer alan formül yardımıyla hesaplanmaktadır.

$$\alpha_j = \frac{\frac{p_{R_j}}{q_{R_j}}}{\frac{p_{F_j}}{q_{F_j}}} \quad (3)$$

Pay: Referans grup için odds oranı,

Payda: Odak grup için odds oranı,

$q$ :  $1-p$  (maddenin doğru cevaplanmama oranı)

$\alpha$ : Odds oranı

$j$ : İncelenilen madde

$j$  maddesinde DMF bulunmadığı durumda bu maddeye ilişkin  $\alpha$  değerinin 1.00 olması beklenmektedir. 1.00'den sapmalar  $j$  maddesinde DMF etkisi olduğu yönünde bilgi vermektedir. Örneğin, buradaki ' $\alpha_j$ ' değerinin 2.00



çıkması;  $j$  maddesini bir grubun doğru cevaplama olasılığının, diğer gruptan 2 kat fazla olduğunu göstermektedir (Camilli ve Shepard, 1994).

*b) Logaritmik odds oranı tekniği.* Bilindiği gibi odak ve referans gruplardan aynı madde için elde edilen ' $p$ ' değerleri farkı  $(-1, 1)$  aralığında yer almaktadır ve 0' a göre simetriktir. ' $\alpha$ ' değerleri ise  $(0, \infty)$  aralığında değerler almaktadır.  $(-1, 1)$  aralığı DMF etkisini yansıtmakta yetersiz kalabilirken, ' $\alpha$ ' değerinin ise simetriklik konusunda sıkıntısı bulunmaktadır. Bu sıkıntılardan ' $\alpha$ ' değerinin logaritması alınarak  $(\beta)$  kurtulunabilmektedir. Böylece negatif değerler bir grubun lehine, pozitif değerler ise diğer grubun lehine, 0'a simetrik olacak şekilde yerleşmektedir. '0' değeri ise DMF olmadığı yönünde bilgi vermektedir.

*c) Yanlılıktan etkilenen bireylerin miktarının Kontinjensi Tablosu yardımıyla hesaplamaya dâhil edilmesi.* Bu tekniklerde yanlılıktan etkilenen birey sayıları hesaplamalara dâhil edilmektedir.

*c.1) Standardizasyon tekniği (Standardized Proportion Difference Measure).* Aynı yetenek düzeyine sahip birey sayılarının odak ve referans gruplar için eşit olmaması, araştırmacıları grup büyüklükleriyle ' $\Delta p_j$ ' değerlerini ağırlıklandırmaya itmiştir. Böylece gruplardaki birey sayılarının farklı olmasının oluşturabileceği baskılama etkisi en aza indirilebilmektedir.

Karşılaşılabilecek bir başka durum ise; madde etkisinin alt ve üst yetenek düzeylerinde farklı şekilde işlemesidir. Alt yetenek düzeyinde bir grubun lehine işleyen madde, üst yetenek düzeyinde diğer grubun lehine işleyebilmektedir. Tek Biçimli Olmayan (TBO) Değişen Madde Fonksiyonu (Non-Uniform DIF) olarak adlandırılan bu durum, yapılan analizlerin hatalı şekilde işlemesine neden olmaktadır. Tek Biçimli (TB) Değişen Madde Fonksiyonu (Uniform DIF) durumunda ise; madde-test etkisi bütün yetenek düzeylerinde aynı grubun lehine işlemektedir.

Bu teknikte Tek Biçimli Değişen Madde Fonksiyonu durumu için; her iki gruptan ayrı ayrı hesaplanan ‘ $p$ ’ değerleri arasındaki farkın birey sayısı ile ağırlıklandırılmış ortalaması alınmaktadır. Bu işleme ‘İşaretli Madde Güçlüğü Farkı (Signed  $p$  Difference (SPD-X))’ denilmektedir. Dorans ve Holland (1993) tarafından önerilen bu teknige ilişkin formül 4 no’lu eşitlikte verilmiştir (akt: Gonzales ve diğ., 2010).

$$SPD - X = \frac{\sum_{j=1}^s n_{Fj} \Delta_{pj}}{\sum_{j=1}^s n_{Fj}} \quad (4)$$

$j$  : Her iki grup için eşleştirilmiş puan seviyesi,

$n_{Fj}$  : Her iki grup için eşleştirilmiş  $j$  puanı seviyesinde, odak grupta bulunan birey sayısı,

$S$ : Her iki grup için eşleştirilmiş puan sayısı,

$\Delta_{pj}$  : İlgili maddenin her iki grup için eşleştirilmiş  $j$  puanı seviyesindeki madde güçlüğü farkı.

Tek Biçimli Olmayan Değişen Madde Fonksiyonu durumunda ise; yine her iki grup için ayrı ayrı hesaplanan ‘ $p$ ’ değerleri arasındaki farkın kareleri toplamının, birey sayısı ile ağırlıklandırılmış ortalamasının karekökü alınmaktadır. Bu işleme de ‘İşaretsiz Madde Güçlüğü Farkı (Un-Signed  $p$  Difference (UPD-X))’ denilmektedir. UPD-X’in SPD-X’den tek farkı; ‘ $p$ ’ farkları karesinin, karekökünün alınması ve tek biçimliliğin değer üzerindeki etkisinin ortadan kaldırılmasıdır. İlgili formül eşitlik 5’de verilmiştir.

$$UPD - X = \sqrt{\frac{\sum_{j=1}^s n_{Fj} (\Delta_{pj})^2}{\sum_{j=1}^s n_{Fj}}} \quad (5)$$

Ağırlıklı puanlanan maddeler için ise Zwick, Donoghue ve Grima (1993), SPD-X tekniğinde kullanılan 'p' değerleri yerine madde ortalamalarının kullanılmasını önermişler ve bu doğrultuda SMD formülünü geliştirmişlerdir (akt: Gonzales ve diğ., 2010). İlgili formül eşitlik 6'de verilmiştir.

$$SMD = \sum_{j=1}^K N_{F.j} MD_j / N_F. \quad (6)$$

$$MD_j = \sum_{c=1}^c N_{F.cj} R_c / N_{F.j} - \sum_{c=1}^c N_{R.cj} R_c / N_{F.j}$$

$j$  : Her iki grup için eşleştirilmiş puan seviyesi,

$N_{Fj}$ : Her iki grup için eşleştirilmiş  $j$  puanı seviyesinde, odak grupta bulunan birey sayısı,

$N_{Rj}$ : Her iki grup için eşleştirilmiş  $j$  puanı seviyesinde, referans grupta bulunan birey sayısı,

$K$ : Her iki grup için eşleştirilmiş puan sayısı.

Bu teknikte elde edilen değerler (-1, 1) aralığında yer almaktadır. -0.05 ile 0.05 aralığında yer alan değerler ihmal edilebilir düzeyde; -1 ile -0.05 ve 0.05 ile 1 aralığında yer alan değerler ise ihmal edilemeyecek düzeyde DMF varlığını göstermektedir (Gonzales ve diğ., 2010).

*c.2) Mantel-Haenszel logaritmik odds oranı (Mantel-Haenszel Log Odds Ratio).* Bu teknikte grup ağırlıkları da göz önünde bulundurularak toplam puan açısından eşleşen bütün alt grup bireylerinden elde edilen odds oranları birleştirilmektedir. İlgili formüller eşitlik 7'de verilmiştir.

$$\alpha_{MH} = \frac{\sum_j \frac{p_{r_j} q_{f_j} n_{r_j} n_{f_j}}{n_j}}{\sum_j \frac{q_{r_j} p_{f_j} n_{r_j} n_{f_j}}{n_j}} \quad (7)$$

$$\beta_{MH} = \log_e(\alpha_{MH})$$

$p_{r_j}$  : İlgili maddenin  $j$  puan seviyesindeki madde güçlüğü,

$q_{f_j}$  : İlgili maddenin  $j$  puan seviyesindeki doğru cevaplanmama oranı.

Formül 7'den de anlaşılacağı üzere, ' $\alpha$ ' değerinin logaritması alınarak ' $\beta$ ' değerleri elde edilmektedir. Pozitif ' $\beta$ ' değerleri referans, negatif değerler ise odak grup lehine DMF olduğunu göstermektedir. ' $\beta$ ' değerinin '0' olması ise DMF olmadığı yönünde bilgi vermektedir (Angoff, 1993).

d) *Madde etkisini test etmek için kullanılan Kontinjensi Tablosu teknikleri (Contingency Table Methods For Testing DIF)*. Kontinjensi Tablosu kullanılarak hesaplanan ' $\Delta p_j$ ' ve ' $\beta$ ' değerlerinin 0'dan anlamlı şekilde farklılaşıp farklılaşmadığını istatistiksel olarak test etmek için yapılan bu analizlerde en az bir tane odak-referans grup eşleşmesine ihtiyaç duyulmaktadır. ' $H_0$ ' hipotezinin kabul edilmesi DMF olmadığı, reddedilmesi ise DMF olduğu yönünde bilgi vermektedir. ' $H_0$ ' hipotezi;

- 1) Evren değerlerinin hepsi için ' $\Delta p_j$ ' ya da ' $\beta$ ' değerlerinin sıfırdan farklı olduğu ya da,
- 2) Bazı evren değerleri için ' $\Delta p_j$ ' ya da ' $\beta$ ' değerlerinin sıfırdan farklı olduğu durumlarda reddedilmektedir (Camilli ve Shepard, 1994).

d.1) *Mantel-Haenszel Ki-Kare Testi (Mantel-Haenszel Chi Square Test)*. Madde etkisini test etmek için kullanılan teknikler içerisinde en yaygın olarak kullanılanlardan biri, Mantel-Haenszel Ki-Kare Testi'dir. Madde etkisi

ölçümlerini veren ‘  $\alpha_{MH}$  ’ ve ‘  $\beta_{MH}$  ’ değerlerinin istatistiksel olarak manidarlığını test edebilmek amacıyla Mantel-Haenszel (1959), ‘1’ serbestlik derecesinde  $\chi^2$  dağılımı gösteren eşitlik 8’deki formülü geliştirmişlerdir (akt: Dorans ve Holland, 1993):

$$MH\chi^2 = \frac{\left\{ \sum_{j=1}^s (A_j - E(A_j)) - \frac{1}{2} \right\}^2}{\sum_{j=1}^s VAR(A_j)} \quad (8)$$

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

$$E(A_j) = \frac{n_{Rj}m_{1j}}{T_j}$$

$A_j$  : İlgili madde için referans grubun  $j$  puan seviyesindeki gözlenen doğru cevap sayısı,

$E(A_j)$  : İlgili madde için referans grubun  $j$  puan seviyesindeki beklenen doğru cevap sayısı

$T_j$  :  $j$  puan seviyesindeki toplam birey sayısı,

$m_{1j}$  :  $j$  puan seviyesi için odak ve referans gruptaki toplam doğru cevap veren birey sayısı,

$m_{0j}$  :  $j$  puan seviyesi için odak ve referans gruptaki toplam yanlış cevap veren birey sayısı (maddeyi boş bırakanlar dâhil).

Yapılan alan yazın incelemesinde, madde etkisinin istatistiksel olarak test edilmesinin pek çok araştırmacı tarafından önerildiği görülmüştür (Holland ve Thayer, 1988; Bishop ve diğ., 1975; Marascuilo ve Slaughter 1981; Angoff 1982; Angoff, 1993; akt: Holland ve Wainer, 1993).

Ağırlıklı puanlanan maddelerde kullanılmak üzere ise Mantel (1963) tarafından yine Mantel-Haenszel tekniğine dayanan ancak kontinjensi tablosunun

kategori sayısı doğrultusunda genişletilmesiyle elde edilen daha geniş bir formül önerilmiştir. İlgili formül eşitlik 9'da verilmiştir (akt: Gonzales ve diğ., 2010).

$$MANTEL = \frac{\left[ \sum_{j=1}^k F_j - \sum_{j=1}^k E(F_j) \right]^2}{\sum_{j=1}^k Var(F_j)} \quad (9)$$

$$F_j = \sum_{c=1}^c R_c N_{F_{cj}}$$

$F_j$ : İlgili madde için referans grubun  $j$  puan seviyesindeki puanı,

$E(F_j)$ : İlgili madde için referans grubun  $j$  puan seviyesindeki beklenen puanı,

$j$ : Her iki grup için eşleştirilmiş puan seviyesi,

$K$ : Her iki grup için eşleştirilmiş puan sayısı.

Ayrıca Educational Testing Service (ETS) tarafından  $\beta_{MH}$  değerleri kullanılarak her bir madde için madde etkisi düzeyini nesnelleştirmek amacıyla çeşitli çalışmalar yürütülmüştür. Bu çalışmalar sonucunda;

$D = -2.35 * \beta_{MH}$  olmak üzere;

A: Madde etkisi bulunamayan maddeyi	$ DI  < 1$
B: Anlamli derecede madde etkisi bulunan maddeyi	$1 <  DI  < 1.5$
C: Yüksek derecede madde etkisi bulunan maddeyi	$ DI  > 1.5$

ifade etmektedir (Zwick, 2012).

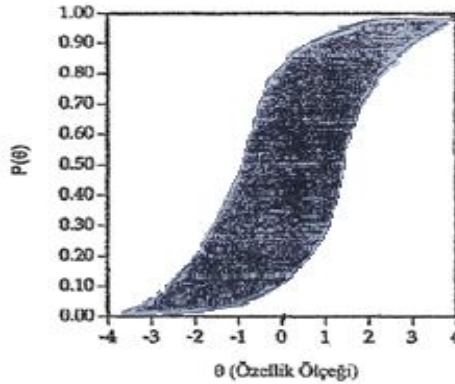
### **I.1. 2. Madde Tepki Kuramına Dayalı Teknikler**

Madde Tepki Kuramı, Klasik Test Kuramı'nın yetersiz ve sınırlı yanlarına çözüm bulmak amacıyla 1930'lu yılların başından itibaren geliştirilmeye başlanmıştır. Madde Tepki Kuramı kısaca; bireylerin kendi yetenekleri doğrultusunda bir maddeye nasıl tepki vereceklerini olasılık hesaplamalarıyla ortaya koyan bir kuramdır. Kuramın sağlamış olduğu en önemli avantaj, birey yeteneklerinin maddelerden bağımsız; maddelerin psikometrik özelliklerinin ise bireylerden bağımsız şekilde kestirilebilmesidir (Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1986; Lord, 1980).

Değişen Madde Fonksiyonunun Madde Tepki Kuramına dayalı tanımı; aynı yetenek düzeyine sahip farklı gruplardaki bireylerin aynı maddeye doğru cevap verme olasılıklarının eşit olmaması şeklindedir. Madde Tepki Kuramıyla Değişen Madde Fonksiyonunun belirlenmesinde öncelikli olarak model veri uyumunun sağlanması gerekmektedir. Daha sonra her iki grup için ilgili model doğrultusunda parametrelerin kestirilmesi, kestirilen bu parametrelerin ölçeklenmesi ve her iki grup için elde edilen Madde Karakteristik Eğri'leri arasındaki farkın incelenmesi gerekmektedir (Osterling, 1983). Madde Tepki Kuramına dayalı Değişen Madde Fonksiyonu belirleme tekniklerinden bazıları, Basit Alan İndeksi Tekniği, Olasılık Farkları İndeksi Tekniği, 'b' Parametresi Farkları Tekniği, Lord'un Ki-Kare Tekniği ve Olabilirlik Oran Testi Tekniği olarak sıralanabilir. Bu teknikler aşağıda kısaca açıklanmıştır.

*Madde Tepki Kuramıyla Değişen Madde Fonksiyonu düzeyinin belirlendiği teknikler.*

a) *Basit Alan İndeksleri.* Bu teknikte her iki grup için kestirilen Madde Karakteristik Eğrileri (MKE) arasında kalan alan hesaplanmaktadır. Bu alanın sıfır çıkması, madde etkisi olmadığı yönünde bilgi vermektedir. Belirtilen durum Şekil 1 üzerinde görülmektedir.



**Şekil 1:** Madde Karakteristik Eğrileri arası alan (Camilli ve Shepard, 1994).

Rudner (1977), bu alanın matematiksel olarak hesaplanmasını sağlayan bir integral formülü geliştirmiştir (akt: Bao & Dayton & Hendrickson, 2009). İlgili formül eşitlik 10'da verilmiştir.

$$\text{İşaretli Alan} = \int [P_R(\theta) - P_F(\theta)] d(\theta) \quad (10)$$

$P_r$ : Referans grup için olasılıkları,

$P_f$ : Odak grup için olasılıkları temsil etmektedir.

İşaretli Alan İndeksi, MKE'lerin birbirini kesmesi durumunda, kesişen kısımların integral hesaplamasında birbirini sıfırlaması nedeniyle 'DMF yok' sonucu üretebilmektedir. Bu yanılgıya düşmemek için hesaplanacak olan bir diğer indeks işaretsiz alan indeksidir. İşaretsiz alan indeksinde, işaretli alan indeksindeki farkın karesi alınarak alanların birbirini sıfırlaması engellenmektedir. İşaretsiz alan indeksinin işaretli alan indeksinden büyük çıkması MKE'lerin



birbirini kestiğinin (Tek Biçimli Olmayan Değişen Madde Fonksiyonu varlığı) bir göstergesidir. İşaretsiz alan indeksine ilişkin formül eşitlik 11’de verilmiştir.

$$\text{İşaretsiz Alan} = \left[ \int (P_R(\theta) - P_F(\theta))^2 d(\theta) \right]^{1/2} \quad (11)$$

Bu tekniklerin 2 önemli sıkıntısı bulunmaktadır.

- 1) Bireyler bazı durumlarda MKE üzerine eşit şekilde dağılmamaktadır. Bu da DMF miktarının (Değişen Madde Fonksiyonundan etkilenen birey sayısı) hatalı şekilde algılanmasına neden olabilmektedir.
- 2) İki grup için kestirilen ‘c’ parametrelerinin farklı olması durumunda integral eğrisinin bir ucu açık kalmakta bu da hesaplamanın sonsuza gitmesine ya da yanıltıcı olmasına neden olmaktadır (Bao & Dayton & Hendrickson, 2009).

*b) Olasılık Farkları İndeksleri.* Basit Alan İndekslerinde bireylerin yetenek boyutu üzerinde eşit dağılmaması yanıltıcı sonuçlara neden olabilmektedir. Bu sıkıntıya Linn ve Harnish (1981) tarafından ilgili alanların, bireylerin MKE üzerine dağılımları kullanılarak ağırlıklandırılmasının önerilmesiyle çözüm getirilmeye çalışılmıştır. Önerilen formül eşitlik 12’de verilmiştir (akt: Camilli ve Shepard, 1994).

$$\Delta P_j = P_R(\theta_j) - P_F(\theta_j)$$

$$SPD - \theta = \frac{\sum_{j=1}^{n_F} \Delta P_j}{n_F} \quad (12)$$

$$UPD - \theta = \frac{\sum_{j=1}^{n_F} (\Delta P_j)^2}{n_F}$$

$\Delta P_j$  : İlgili maddeye yönelik  $j$  yetenek seviyesindeki her iki grup için elde edilen doğru cevaplama olasılıkları farkı.

*c) b Parametresi Farkları.* Bu teknikte, aynı yetenek düzeyine sahip bireylerden, her iki grup için ayrı ayrı kestirilen  $b$  parametreleri arasındaki fark hesaplanmaktadır. Bu farkın '0' olması, maddede DMF olmadığı yönünde bilgi vermektedir.

*Madde Tepki Kuramıyla Değişen Madde Fonksiyonunun test edildiği teknikler.* Bu kısımda madde etkisinin manidarlığını test etmek üzere geliştirilmiş tekniklere değinilmiştir.

*a) b Parametresi farklarının test edilmesi.* Bu teknikte her iki grup için aynı maddeye yönelik olarak ayrı ayrı kestirilen  $b$  parametreleri arası farkların '0' dan anlamlı şekilde değişiklik gösterip göstermediği incelenmektedir.

Bu işlemi gerçekleştirebilmek için her iki gruptan kestirilen  $b$  parametrelerinin standart hatalarına ihtiyaç duyulmaktadır. İlgili formül eşitlik 13'de verilmiştir.

$$S_{\Delta b} = \sqrt{S_F^2 + S_R^2} \quad (13)$$

$$d = \frac{\Delta b}{S_{\Delta b}}$$

$$\Delta_b = b_F - b_R$$

Buradaki  $d$  deęerinin normal daęılım gsterdięi sayıtlısıyla kurulan hipotez ise eřitlik 14’de verilmiřtir.

$$H_0 = \Delta_b = 0 \quad (14)$$

*b) Lord’un Ki-Kare Teknięi.* Bu teknik Lord (1982) tarafından nerilmiř olup, ilgili forml eřitlik 15’de verilmiřtir (aktaran: Camilli ve Shepard, 1994).

$$V' = (a_F - a_R, b_F - b_R) \quad (15)$$

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

$S$ : Madde parametreleri arasındaki farkın Varyans-Kovaryans matrisi.

Ancak bu teknikte iki grup iin kestirilen MKE’ler benzer olsalar bile,  $H_0$  reddedilebilmektedir. ünkü farklı  $a$  ve  $b$  parametresi kombinasyonları benzer řekillerde MKE’ler oluřturabilmektedir. Bu da DMF etkisini, MKE’ler hemen hemen aynı olsa bile, parametreler arasında fark olması nedeniyle manidar dzeye ulařtırabilmektedir. Bu nedenle ilgili test nerilmemektedir (Camilli ve Shepard, 1994).

*c) Olabilirlik Oran Testi teknięi.* Bu teknikte odak ve referans gruplarına ynelik ayrı ayrı kestirilen madde parametreleri arasında anlamlı fark olup olmadıęı incelenmektedir. Bu amala madde parametrelerinin eřit olduęu varsayılan sınırlandırılmıř ve i.madde parametresinin eřit olmadıęı, dięer madde parametrelerinin ise eřit olduęu varsayılan geniřletilmiř bir model oluřturulmakta ve iki modelin olabilirlik deęerlerinin logaritmaları arasındaki fark hesaplanmaktadır. Bu deęer kullanılan modelin parametre sayısı kadar serbestlik derecesinde  $\chi^2$  daęılımı gstermektedir. Elde edilen deęerin tablo deęerini ařması durumunda yokluk hipotezi red edilmekte ve ilgili maddenin DMF varlıęı gsterdięi řeklinde yorum yapılmaktadır (Thissen, 2001).

Yukarıda Klasik Test ve Madde Tepki Kuramlarına dayalı olarak geliştirilen bazı DMF belirleme teknikleri açıklanmaya çalışılmıştır. Her iki kuram da günümüzde yaygın olarak kullanılmaktadır, ancak araştırmacılar Madde Tepki Kuramı'nın Klasik Test Kuramı'na göre pek çok avantajının bulunduğunu iddia etmektedir. Bunlardan DMF açısından önemli sayılabilecek bazılarına aşağıda kısaca değinilmiştir (Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1986; Lord, 1980).

- 1) Madde Tepki Kuramı'nda madde parametreleri gruptan bağımsızdır.
- 2) Gruplar ve bireyler yetenekleri doğrultusunda daha anlamlı şekilde eşleştirilebilmektedir.
- 3) Kestirilen parametreler standart hataları ile birlikte elde edilebilmektedir.
- 4) Madde özellikleri grafiksel olarak daha rahat şekilde gösterilebilmektedir. Böylece DMF'nin grafiksel olarak gösterimi kolaylaşmaktadır.

DMF saptanmasında Madde Tepki Kuramı modellerinin kullanılması yukarıda bahsedilen nedenlerden ötürü kolaylık ve kesinlik sağlamaktadır.

Yukarıda açıklanan KTK ve MTK'na dayalı bütün DMF belirleme teknikleri zaman içerisinde bilgisayar altyapısı ve kuramlardaki gelişmeler dâhilinde pek çok araştırmacının da katkısıyla geçerlik çalışmaları kapsamında geliştirilmiştir. Ancak teknikler hakkında yapılan açıklamalardan, bu tekniklerin hemen hepsinin güçlü ve zayıf yanlarının bulunduğu görülmektedir. Nitekim pek çok teknik birbirlerinin zayıf yanlarını gidermeye yönelik olarak geliştirilmiştir.

Farklı DMF belirleme tekniklerinin güçlü ve zayıf yanlarını belirlemek amacıyla bu tekniklerin farklı değişkenlere göre işleyişlerini inceleyen pek çok çalışma yapılmıştır. Yapılan bu çalışmalar farklı DMF belirleme teknikleriyle elde edilen bulguların, DMF'li madde sayısı-oranı, test uzunluğu, DMF düzeyi,

örneklem büyüklüğü, maddelerdeki DMF biçimi vb. değişkenlerden etkilendiğini ortaya koymaktadır (Gelin ve Zumbo, 2003; Osterling, 1983; Camili ve Shepard, 1994; Gierl, Jodoin ve Ackerman, 2000; Narayanan ve Swaminathan, 1994; Padilla ve diğ., 2012).

*Değişen Madde Fonksiyonu içeren madde sayısı ve test uzunluğu.* Test uzunluğunun artması Klasik Test Kuramı'na göre güvenilirliği de arttırmaktadır (Gulliksen, 1950). Daha güvenilir testlerden elde edilen toplam puanlar ise bireylerin yetenek düzeylerini daha güvenilir şekilde yansıtmaktadır. Test içerisindeki DMF'li madde sayısının-oranının artması ise toplam puanların bireylerin gerçek yetenek düzeylerini yansıtmalarını engellemekte ve odak-referans grup eşleştirmelerini olumsuz yönde etkilemektedir (Narayanan & Swaminathan, 1994). Ayrıca yanlış madde oranının yükselmesi geçerliği de olumsuz yönde etkilemekte ve DMF belirleme tekniklerinin etkililiğini azaltmaktadır (Osterling, 1983; Camili & Shepard, 1994; Gierl & Jodoin & Ackerman, 2000). Narayanan ve Swaminathan (1994); yapmış oldukları çalışmada standartlaştırılmış başarı testlerinde %10 ve %15 düzeyinde DMF içeren madde ile karşılaşılmasının olası, %20 ve üzerindeki bir düzeyin ise test için oldukça kötü bir senaryo olduğunu belirtmişlerdir.

*Maddelerdeki Değişen Madde Fonksiyonu düzeyi.* Maddelerdeki DMF düzeyi, bireylerin DMF'den etkilenme derecelerini ve DMF belirleme tekniklerinin sonuçlarını değiştirebilmektedir. Bu nedenle araştırmacılar DMF içeren maddelerdeki DMF düzeyini nesnelleştirmek amacıyla çeşitli ölçütler geliştirmişlerdir. Bu ölçütlerden birine göre M-H prosedüründeki alfa değerinin  $-2,35 \ln'$  alınmakta ve bu değer mutlak değer 1'den küçük olması durumunda 'A', 1 ve 1,5 arasında olması durumunda 'B' ve 1,5'den büyük olması durumunda ise 'C' düzeyinde DMF varlığı kabul edilmektedir. Bu ölçüt halen Educational Testing Service (ETS) tarafından kullanılmaktadır (Zwick., 2012).

*Örneklem büyüklüğü.* Alanyazında M-H prosedürü üzerinde yapılan çalışmalar artan örneklem büyüklüğünün DMF içeren maddelerin doğru tespit edilme oranını da arttırdığını göstermektedir (Rogers, 1989; Rogers ve Swaminathan, 1993; Swaminathan ve Rogers, 1990, akt: Narayanan ve Swaminathan, 1994). Ayrıca Mazor ve diğ. (1992) yapmış oldukları çalışmada, odak ve referans gruplarına ait yetenek dağılımlarının DMF belirleme tekniklerinin sonuçlarını değiştirdiğini göstermişlerdir (akt: Narayanan ve Swaminathan, 1994).

*Maddelerdeki Değişen Madde Fonksiyonu biçimi.* Maddelerin Tek Biçimli ya da Tek Biçimli Olmayan Değişen Madde Fonksiyonu durumu içermeleri DMF belirleme tekniklerinin sonuçlarını değiştirmektedir. Tek Biçimli Olmayan Değişen Madde Fonksiyonu durumu gösteren maddelerde, alt grupta bir grubun lehine işleyen madde üst grupta diğer grubun lehine işleyebilmekte ve bu durum DMF varlığının saptanamamasına neden olmaktadır (Osterling, 1983).

DMF belirleme teknikleriyle elde edilen bulguları değiştirebilecek bir başka değişkenin ise maddelerin puanlanma şekli olduğu düşünülmektedir. Nitekim sosyal bilimler için her maddenin ve seçeneğin psikolojik uzaydaki karşılıklarının eşit olmasını beklemenin mümkün olmadığı bilinmektedir.

Yapılan araştırmalar, ikili puanlamanın bireylerin kısmi bilgilerini yansıtmakta yetersiz kaldığını; ağırlıklı puanlamanın ise bireylerin kısmi bilgilerini daha iyi yansıttığını ve yetenek boyutunun daha büyük bir ranji hakkında daha fazla ve daha güvenilir bilgi sağladığını göstermektedir (Ostini ve Neiring, 2006; Embretson ve Reise, 2000). Benzer şekilde, Samejima (1975, 1979) da yapmış olduğu çalışmada ağırlıklı puanlanan maddelerin ikili puanlananlara göre daha fazla istatistiksel bilgi sağladığını göstermiştir (akt: Ostini ve Neiring, 2006).

İkili puanlama (1-0), sağlamış olduğu puanlama kolaylığı nedeniyle genellikle çoktan seçmeli test maddelerinin puanlanmasında tercih edilmektedir. Bu puanlama yönteminde doğru cevaplara ‘1’; yanlış, boş, birden fazla cevap işaretlenen ve erişilemeyen maddelere ise ‘0’ puan atfedilmektedir (Gulliksen, 1950). Buradan bu puanlama yönteminde bütün maddelerin eşit şekilde ağırlıklandırıldığı görülmektedir. Diğer ifadeyle bu puanlama yönteminde, maddeyi şansla doğru cevaplayanlar ile maddeyi doğru cevaplayanların ‘tam bilme’; maddeyi kısmen bilenlerin, risk almayarak cevaplama davranışından imtina edenlerin, maddeyi yanlış cevaplayanların, birden fazla seçeneği işaretleyenlerin ve maddeye hiç erişemeyenlerin ise ‘tam yanlış bilme’ durumuna sahip olduğu varsayılmaktadır. Bu durum nedeniyle ikili puanlama yöntemi, bireylerin gerçek bilgi düzeylerini yansıtmakta yetersiz kaldığı şeklinde eleştirilmektedir (Embretson ve Reise, 2000).

Abu Sayf (1979), ikili puanlama durumu için alanyazında yapılan bu eleştirileri psikometrik, pragmatik, ahlaki ve politik olmak üzere 4 ana başlık altında toplamıştır. Psikometrik açıdan yapılan eleştiriler, ikili puanlamanın öğrencileri şans başarısını kullanmaya teşvik etmesi ve bu nedenle hata varyansının büyümesi ile kısmi bilginin ölçülmesindeki yetersizlikler üzerine yoğunlaşmaktadır. Pragmatik açıdan yapılan eleştiriler, ikili puanlamanın risk almayarak maddeyi cevaplamaktan imtina eden bireylerin cezalandırıldığı düşüncesi üzerine yoğunlaşmaktadır. Ahlaki açıdan yapılan eleştiriler, ikili puanlamanın bireyleri şans başarısını kullanmaya itmesi ve şansla maddeyi doğru cevaplayanların ödüllendirilmesi üzerine yoğunlaşmaktadır. Son olarak politik açıdan yapılan eleştiriler ise, ikili puanlamanın bireylerin gerçek yetenek düzeyini yansıtmadaki yetersizliği nedeniyle yapılan sınavlara olan güvenin kaybolması üzerine yoğunlaşmaktadır.

Araştırmacılar bu sıkıntılardan bazılarını gidermek amacıyla çoktan seçmeli maddelerin ya da seçeneklerin farklı şekillerde ağırlıklandırıldığı farklı

ağırlıklı puanlama yöntemleri geliştirmişlerdir. Frary (1989) geliştirilen bu puanlama yöntemlerini; Cevaplayıcı Kararları Yöntemleri (Examinee Judgements Methods) ve Doğrudan Cevaplama Yöntemleri (Direct Response Methods) olmak üzere iki kategoride sınıflamıştır. Bu kategorilerden Cevaplayıcı Kararları Yöntemleri, kendi içinde Güven Testi (Confidence Testing) ve Alt Grup Seçme Yöntemi (Subset-Selection Methods) olmak üzere 2 alt kategoriye; Doğrudan Cevaplama Yöntemleri ise, Seçenek Ağırlıklandırma Yöntemi (Opinion Weighting), Doğruyu Bulana Kadar Cevaplama Yöntemi (Answer-Until-Correct Procedure), Çoklu Doğru Cevap Yöntemi (Multiple Correct Options) ve Madde Tepki Kuramı Yöntemleri (Item Response Theory Methods) olmak üzere 4 alt kategoriye ayrılmaktadır. Bu yöntemlere ilişkin Frary'nin (1989) yapmış olduğu açıklamalara aşağıda kısaca değinilmiştir.

*Cevaplayıcı kararları yöntemlerinde*, bireylerden bir maddeye ilişkin bilgi düzeylerine karar vermeleri beklenmektedir. Bu yöntemlerden *güven testi yönteminde*, bireylerden seçtikleri seçeneğin doğru olduğuna inanma düzeylerini belirtmeleri; *alt grup seçme yönteminde* ise, kesinlikle yanlış olduğuna inandıkları seçeneği belirtmeleri istenmektedir.

*Doğrudan cevaplama yöntemlerinde*, bireylerden en doğru olduğunu düşündükleri seçeneği işaretlemeleri beklenmektedir. Bu yöntemlerden *seçenek ağırlıklandırma yönteminde*, bireyler doğru olduğunu düşündükleri yalnız bir seçeneği işaretlemektedir. Buradan hareketle ise önsel (uzman yargısına dayalı ağırlıklı (UYDA)) ya da deneysel (Deneysel Ağırlıklı (DA)) yöntemlerle madde seçenekleri ağırlıklandırılmaktadır. UYDA durumunda genellikle uzmanların madde seçeneklerine attıkları puanların ortalamaları alınmakta, deneysel ağırlıklandırmada ise bireylerin maddelere verdikleri cevap dağılımı üzerinden hareket edilmektedir. Burada genellikle üst gruptan fazlaca işaretlenen seçeneklerin ağırlığı artırılmakta, alt gruptan fazlaca işaretlenen seçeneklerin ağırlığı ise azaltılmaktadır. *Doğruyu bulana kadar cevaplama yönteminde*,



bireylerden doğru olduğuna inandıkları cevabı, anahtar cevabı buluncaya kadar işaretlemeleri beklenmektedir. *Çoklu doğru cevap yönteminde*, bireylerden farklı sayılarda anahtar cevap içeren ve seçeneklerin her biri aynı zamanda birer doğruyanlış maddesi olan maddelerin her bir seçeneğini incelemeleri ve doğru olan seçenekleri işaretlemeleri beklenmektedir. *Madde Tepki Kuramı'na* dayalı olarak ise ağırlıklı seçeneklere uygun pek çok model geliştirilmiştir. Bu modellerden bazıları Ağırlıklandırılmış Cevap Modeli, Kısmi Puan Modeli, Dereceleme Ölçeği Modeli ve Sınıflamalı Tepki Modeli'dir.

*Ağırlıklandırılmış Cevap Modeli (Graded Response Model) (ACM)*, Samejima (1969) tarafından, 2 Parametrelili Lojistik Modelin bir uzantısı olarak, madde kategorilerinin sıralı ve birikimli olduğu durumlarda kullanılmak üzere geliştirilmiş olup, Madde Tepki Kuramı'nın fark modeli (difference model, indirect) sınıfına girmektedir (Akt: Embretson ve Reise, 2000).

ACM'ye ait özellikler aşağıda maddeler halinde özetlenmiştir:

- 1- Madde kategorilerinin sıralı ve birikimli olduğu durumlarda kullanılmaktadır.
- 2- Madde formatlarının aynı olmasına gerek bulunmamaktadır. Bir madde 5, diğeri 4 kategorili olabilir.
- 3- Her bir madde için ayrı ayrı eğim parametresi ( $\alpha_i$ ) kestirilmektedir.
- 4-  $K_i$ = kategori sayısı olmak üzere  $K_i-1$  tane kategori kesişim (threshold ( $\beta_{ij}$ )) parametresi kestirilmektedir. Kesişim noktası parametreleriyle birbirine komşu iki kategorinin seçilme olasılıklarının eşit olduğu (kesiştigi) noktalar belirlenebilmektedir.
- 5- Bireylerin maddelere verdikleri tepkilerin '0' ve '1' cevap formatına dönüştürebilmesi için; belirli bir kategorinin altında tepki veren bireylere 0, üstünde tepki verenlere ise '1' puan verilmektedir. Bir başka ifadeyle

ikilileştirme (dikotomizasyon); (0-1,2,3), (0,1-2,3), (0,1,2-3) şeklinde gerçekleştirilmektedir (Hambleton ve Swaminathan, 1989; Embretson ve Reise, 2000; Ostini ve Neiring, 2006).

ACM'de olduğu gibi fark modellerinde bireyin bir kategoride tepki verme olasılığı 2 basamakta hesaplanmaktadır (Embretson ve Reise, 2000).

Öncelikle maddedeki kategori sayısı  $K_i = m_i + 1$  olmak üzere her bir madde için ayrı ayrı;  $m_i$  tane İşletim Karakteristik Eğrisi (İKE, Operating Characteristic Curve) hesaplanmaktadır. Bu eğriler bireylerin, maddeye vermiş oldukları tepkilerin kategori kesişim değerinin altında ya da üstünde olma olasılığını vermektedir. İlgili formül eşitlik 16'da verilmiştir.

$$P_{ix}^*(\theta) = \frac{e^{\alpha_i(\theta - \beta_{ij})}}{1 + e^{\alpha_i(\theta - \beta_{ij})}} \quad (16)$$

$P_{ix}^*(\theta)$  : Belirtilen örtük özellik düzeyine sahip bireyin, belirtilen kategorinin üstünde cevap verme olasılığını,

$\alpha_i$  : İlgili maddenin eğim parametresi,

$\beta_{ij}$  : i maddesinin j kategorisine ait 'yer parametresi'. Bu parametre, kesişim (threshold) noktalarının yetenek boyutu üzerindeki (axis) yerini göstermekte olup, her madde ve her kategori için ayrı ayrı hesaplanmaktadır.

Bir sonraki basamakta ise bireyin ilgili kategoride tepki verme olasılıklarını belirten gerçek kategori tepki olasılıkları (Actual Category Response Probabilities), istenilen kategorinin alt sınır olasılığı üst sınır olasılığından çıkartılarak hesaplanmaktadır. Buradan da kategori tepki eğrileri (KTE, Category Response Curve) elde edilmektedir.

4 kategorili bir madde için gerçek kategori tepki olasılıkları eşitlik 17'deki gibi hesaplanmaktadır.

$$\begin{aligned}
 P_{i0}(\theta) &= 1.0 - P_{i1}(\theta) \\
 P_{i1}(\theta) &= P_{i1}(\theta) - P_{i2}(\theta) \\
 P_{i2}(\theta) &= P_{i2}(\theta) - P_{i3}(\theta) \\
 P_{i3}(\theta) &= P_{i3}(\theta) - 0
 \end{aligned} \tag{17}$$

$P_{i0}(\theta)$ : Maddenin 1. Kategorisinin seçilme olasılığı,

$P_{i1}(\theta)$ : Maddenin 2. Kategorisinin seçilme olasılığı,

$P_{i2}(\theta)$ : Maddenin 3. Kategorisinin seçilme olasılığı,

$P_{i3}(\theta)$ : Maddenin 4. Kategorisinin seçilme olasılığı.

*Kısmi Puan Modeli (The Partial Credit Model) (KPM)*, Masters (1982) tarafından yine sıralı kategorilere sahip maddelerin modellenmesi amacıyla geliştirilmiştir (akt: Ostini ve Neiring, 2006). Maddenin doğru cevabına ulaşmak için birden fazla aşamadan geçmenin ve bu aşamaların farklı şekillerde puanlanmasının gerektiği durumlarda kullanılabilir. Bu modelde bütün maddeler için ayrı ayrı kategori sayısının bir eksiği kadar kesişim parametresi kestirilmekle birlikte, maddelerin eşit eğime (ayırt edicilik) sahip olduğu varsayılmaktadır. Bireylerin maddelere verdikleri tepkilerin '0' ve '1' cevap formatına dönüştürebilmesi için; kategoriler ikişer ikişer eşlenmekte ve alt kategoriye 0, üst kategoriye ise 1 puan verilmektedir (Hambleton ve Swaminathan, 1989; Ostini ve Neiring, 2006; Van Der Linden ve Hambleton, 1997).

Bu model Madde Tepki Kuramı'nın 'toplama bölmeli' (divide by total, direct) modeller sınıfına girmektedir. Toplama bölmeli modellerde bireylerin belirtilen bir kategoride tepki verme olasılığı doğrudan  $\frac{(e)}{\text{toplama}(e)}$  formülüyle hesaplanmaktadır.

Modelde bir kişinin,  $i$  maddesinin  $j$  kategorisinde cevap verebilmesi için kendinden önceki bütün kategorileri başarıyla yerine getirmiş olması gerekmektedir. Bu nedenle bireyin  $i$  maddesinin  $j$  kategorisine kadar ulaşabilme olasılığı eşitlik 18’de belirtilen formülle hesaplanmaktadır (Embretson ve Reise, 2000; 105).

$$P_{ix}(\theta) = \frac{\exp \left[ \sum_{j=0}^x (\theta - \delta_{ij}) \right]}{\sum_{r=0}^{m_i} \left[ \exp \sum_{j=0}^r (\theta - \delta_{ij}) \right]} \quad (18)$$

$M_i$ : Kategori sayısı,

$\theta$  : Örtük özellik parametresi,

$x$ : İstenilen kategori,

$\delta_{ij}$  : Madde güçlüğünü ifade eden 2 kategori tepki eğrisinin (category response curve) kesiştiği (intersection) yeri belirten kesişim parametresidir. ACM’deki  $B_{ij}$  parametresine benzemekle birlikte kategorilerin birbirlerine göre görelî güçlüklerini ifade etmektedir.

$r$ : hesaplanacak olan kategori.

*Dereceleme Ölçeği Modeli (Rating Scale Model) (DÖM)*, Andrich (1978) tarafından geliştirilmiştir. Yine bu modelde de bütün maddelerin eşit eğime (ayrıt edicilik) sahip olduğu varsayılmaktadır. Bireylerin maddelere verdikleri tepkilerin ‘0’ ve ‘1’ cevap formatına dönüştürebilmesi için;  $h$  kategorisi seçilmişse ‘1’, seçilmemişse ‘0’ puan verilerek puanlama yapılmaktadır. Aynı cevap formatındaki sıralı tepki kategorilerine sahip maddeler için uygun olan bu modelde kategoriler arası uzaklıkların eşit olduğu varsayılmaktadır. Bu nedenle bütün maddelerin bağıl güçlüğünü belirlemek amacıyla bir yer parametresi, ayrıca aynı seviyedeki kategoriler için birer de ortak kesişim parametresi kestirilmektedir (Hambleton ve Swaminathan, 1989; Embretson ve Reise, 2000; Ostini ve Neiring, 2006).

DÖM’de bireyin belirtilen bir kategoride cevap verme olasılığı eşitlik 19’da belirtilen formülle hesaplanmaktadır (Ostini ve Nering, 2006).

$$Px(\theta) = \frac{\exp\left[\sum_{j=0}^x [\theta - (\lambda i + \delta i)]\right]}{\sum_{x=0}^M \exp\left[\sum_{j=0}^x [\theta - (\lambda i + \delta i)]\right]} \quad (19)$$

$\lambda i$  : Maddenin örtük ölçek üzerindeki yer parametresi olup (her madde için ayrı ayrı hesaplanmaktadır) maddelerin birbirlerine göre görelî güçlüklerinin belirlenmesini sağlamaktadır.

$\delta i$  : Kategori kesişim parametresi (DÖM’de bütün maddeler için eşit olduğu varsayılmaktadır).

*Sınıflamalı Tepki Modeli (STM) (Nominal Response Model)*, Bock (1972) tarafından geliştirilmiş olup, sıralı olmayan iki ya da daha fazla kategoriye sahip maddeleri analiz etmek için kullanılmaktadır (akt: Embretson ve Reise, 2000). Bu tür maddelere verilen standart örnek, seçenekleri arasında herhangi bir sıra bulunmayan çoktan seçmeli maddelerdir. Bu modelde bireylerin maddelere verdikleri tepkilerin ‘0’ ve ‘1’ cevap formatına dönüştürebilmesi için; h kategorisi (seçeneği) seçilmişse ‘1’, seçilmemişse ‘0’ puan verilerek puanlama yapılmaktadır. Bu modelde her bir madde ve bu maddelere ait her bir kategori için ayrı ayrı eğim ve kesişim parametresi kestirilmektedir (Hambleton ve Swaminathan, 1989; Embretson ve Reise, 2000; Ostini ve Nering, 2006).

STM’de bir katılımcının  $i$  maddesinin ‘ $x$ ’ kategorisinde tepki verme olasılığı eşitlik 20’de belirtilen formülle hesaplanmaktadır (Ostini ve Nering, 2006).

$$P_{ix}(\theta) = \frac{\exp(a_{ix}\theta + c_{ix})}{\sum_{x=0}^m \exp(a_{ix}\theta + c_{ix})} \quad (20)$$

$P_{ix}(\theta)$  : Belirli bir düzeyde örtük özelliğe sahip kişinin i maddesinin j kategorisinde tepki verme olasılığını,

$a_{ix}$  : i maddesinin x kategorisine ait eğim parametresini,

$c_{ix}$  : i maddesinin x kategorisine ait kesişim (intercept) parametresini,

$m$  : i maddesinin kategori sayısını temsil etmektedir.

Modelin tanımlanabilmesi için  $a_{ix}$  ve  $c_{ix}$  parametrelerinin kendi içlerinde toplamalarının ayrı ayrı '0' olması gerekmektedir.

İkili ve ağırlıklı puanlama yöntemleri DMF açısından ele alındığında, DMF belirleme tekniklerinin bu yöntemlerden etkileneyeceği ve farklı DMF belirleme teknikleriyle belirlenen maddelerin puanlama yöntemi farklılaştığında değişeceği düşünülmektedir. Bu nedenle bu araştırmanın temel konusunu, farklı DMF belirleme tekniklerinin ikili ve ağırlıklı puanlama durumlarındaki işleyişi oluşturmaktadır.

### 1.3. İlgili Araştırmalar

Narayanan ve Swaminathan (1994) yapmış oldukları çalışmada, M-H ve SIBTEST (Simultaneous Item Bias Procedures) tekniklerinin 1. tip hata oranlarını, 9 farklı örneklem büyüklüğü, 3 farklı yetenek dağılımı ve testte bulunan 2 farklı toplam DMF içeren madde sayısı koşulları altında incelemiştirlerdir. Simülasyon yazılımı kullanılarak 3 parametrelili lojistik modele göre elde edilen veri seti üzerinde (100 replikasyon) yapılan analizler sonucunda, odak ve referans gruplarının yetenek dağılımlarının eşit olduğu durumlarda, kullanılan her iki tekniğin de Tek Biçimli Değişen Madde Fonksiyonu varlığını saptamada eşit

güce sahip olduğu bulunmuştur. Odak ve referans gruplarının yetenek dağılımlarının eşit olmadığı durumlarda ise Tek Biçimli Değişen Madde Fonksiyonu varlığını saptamada SIBTEST tekniğinin daha başarılı olduğu bulunmuştur.

Spray ve Miller (1994) yapmış oldukları çalışmayla, Mantel, Genelleştirilmiş M-H ve Lojistik Diskriminant Fonksiyon Analizi tekniklerinin Tek Biçimli DMF durumunda yüksek derecede uyumlu sonuçlar verdiğini bulmuşlardır.

Kristjansson (2001) yapmış olduğu çalışmayla, çok kategorili puanlama durumlarında Klasik Test Kuramı'na dayalı M-H, Genelleştirilmiş Mantel ve Ordinal Lojistik Regresyon tekniklerini karşılaştırmış ve Tek Biçimli DMF açısından bu tekniklerin benzer sonuçlar verdiğini bulmuştur.

Gelin ve Zumbo (2003) yapmış oldukları çalışmada, bir maddenin farklı şekillerde puanlanmasının DMF sonuçlarını değiştirip değiştirmediğini incelemişlerdir. Cinsiyet değişkeni açısından yürütülen DMF analizleri, 20 maddeden oluşan bir depresyon ölçeğinin (Center For Epidemiologic Studies Depression Scale) 290 bayan ve 310 erkek örneklem grubuna uygulanmasıyla elde edilen veriler üzerinde yapılmıştır. Kullanılan ölçeğin maddeleri depresyon halinin belirtilerini içermektedir ve kategorileri; '0'=nadiren, '1'=ara sıra, '2'=sıklıkla, '3'=genellikle şeklinde puanlanmaktadır. Verilerin analizi için ise maddeler A) '0'=nadiren, '1'=ara sıra, '2'=sıklıkla, '3'=genellikle, B) '0'=nadiren, '1'=ara sıra, '1'=sıklıkla, '1'=genellikle, C) '0'=nadiren, '0'=ara sıra, '1'=sıklıkla, '1'=genellikle, olmak üzere 3 farklı şekilde puanlanmıştır. Farklı puanlama şekilleri kullanılarak elde edilen veriler üzerinde Ordinal Lojistik Regresyon Analizi (OLR) tekniği kullanılmıştır. Araştırma sonucunda, kullanılan puanlama şeklinin DMF sonuçlarını değiştirdiği bulunmuştur. Gelin ve Zumbo (2003), bu durumun depresyon ölçeğinde bulunan 'Çok üzgün olduğunuzda ağlar mısınız?' maddesiyle daha iyi açıklanabileceğini belirtmektedirler. İlgili madde

ikili puanlandığında erkek ve kadınlar arasında DMF oluşmadığı, çok kategorili puanlandığında ise oluştuğu belirtilmiştir. Bunun gerekçesi olarak ise ikili puanlamanın, kullanılan yöntem için bireyler arasında yeterli değişkenlik oluşturmadığı, bir başka ifadeyle bireylerin gerçek depresyon durumunu yansıtmakta yetersiz kaldığı gösterilmiştir (Gelin ve Zumbo, 2003; 72).

Hidalgo ve Lopez (2004) yapmış oldukları çalışmada, LR, MH ve GMH tekniklerinin DMF belirlemedeki etkililiklerini etki büyüklüğü indeksi kullanarak karşılaştırmışlardır. İlgili analizler benzetim tekniğiyle elde edilmiş 2000 kişilik (1000 kişilik odak grup ve 1000 kişilik referans grup) veri üzerinde yapılmıştır. Benzetimlenen veriler, 59 DMF içermeyen ve 16 DMF içeren toplam 75 maddeyi temsil etmektedir. Araştırma sonucunda, LR tekniğinin MH ve GMH tekniğine oranla daha fazla maddede DMF varlığı saptandığı ancak etki büyüklüğü açısından diğer tekniklere oranla daha zayıf kaldığı bulunmuştur.

Kristjansson ve diğ. (2005) yapmış oldukları çalışmada, çoklu puanlanan maddeler için DMF belirlenmesinde kullanılan Mantel, Genelleştirilmiş MH (GMH), Sıralı Lojistik Regresyon (SLR (Ordinal Logistic Regression)) ve Lojistik Diskriminant Fonksiyon Analizi (Logistic Discriminant Function Analysis) (LDFA) tekniklerini karşılaştırmışlardır. Çalışmada 4 kategoride (0, 1, 2, 3) puanlanan 26 maddeye ilişkin 4000 kişilik (3200 kişilik referans, 800 kişilik odak grup) veri, G-KP modeline uygun olarak benzetimlenmiştir. Çalışma sonucunda, kullanılan DMF belirleme tekniğine göre DMF bulgularının değiştiği ve SLR ile GMH tekniklerinin Mantel ve LDFA tekniklerine göre daha duyarlı olduğu sonucuna ulaşılmıştır.

Korkmaz (2005) yapmış olduğu çalışmada, Madde Tepki Kuramı'na dayalı tekniklerle çok kategorili maddelerde madde ve test yanlılığını incelemiştir. Çalışma; Beş Faktör Kişilik Envanteri'nin (5FKE) 5 dereceli (tamamen uygun- hiç uygun değil) puanlanan ve toplam 11 maddeden oluşan Yumuşak Başlılık/Geçimlilik faktörü, 'Uzlaşma' alt ölçeğinin uygulandığı 6000 kişilik veri



seti üzerinde cinsiyet deęişkeni temel alınarak yürütülmüştür. Çalışmada Aęırlıklandırılmış Cevap Modeli kullanılmış ve farklı örneklem genişlikleri (n=500, n=1000 ve n=1500) oluşturularak üç farklı işlev farklılık teknięi (Olabilirlik Oranı Testi Model Karşılaştırma Teknięi (LR), Madde-Test İşlev Farklılık Teknięi (DFIT) ve Parametre Karşılaştırma Teknięi) kullanılmıştır. Araştırma sonucunda, kullanılan DMF belirleme tekniklerine göre DMF bulgularının deęiştiięi, Olabilirlik Oran Testi ve Madde Parametresi Karşılaştırma tekniklerinin, 500 kişilik örneklem genişliğinde işlev farklılıklarını saptamada yetersiz kaldıęı, örneklem sayısının artmasıyla ise (n=1000 ve n=1500) her iki teknięin benzer sonuçlar verdięi ve hem 0.01 hem de 0.05 nominal alfa seviyelerinde gruplar arasında işlev farklılıęı gösteren maddelere karşı hassasiyetlerinin arttıęı gözlenmiştir. DFIT teknięinin ise tüm örneklem genişliklerinde daha tutarlı sonuçlar verdięi gözlenmiştir. Araştırma sonucunda ilgili ölçeęin 2 maddesinin madde işlev farklılıęı gösterdięi bulunmuştur.

Öğretmen (2006) yapmış olduęu çalışmada, Uluslararası Okuma Becerilerinde Gelişim Projesi (PIRLS) 2001 testinin psikometrik özelliklerini (Türkiye-Amerika Birleşik Devletleri Örneęi) incelemiştir. Çalışmada; PIRLS okuma parçalarına ait testlerin ölçtüęü varsayılan psikolojik yapıların kültürlere göre eş deęer olup olmadığı ve test maddelerinin kültürlere göre DMF içerip içermedięi belirlenmeye çalışılmıştır. Çalışma, 5125 kişilik Türk ve 3763 kişilik Amerika Birleşik Devletleri örneklem verisi içermektedir. Yapılan DMF analizleri ise Parametre Karşılaştırma ve Olabilirlik Oran Testi tekniklerini içermektedir. Araştırma sonucunda DMF bulgularının kullanılan DMF belirleme teknięine göre deęiştiięi bulunmuştur. Araştırmada Parametre Karşılaştırma teknięiyle 98 maddenin 36'sında, Olabilirlik Oran Testi teknięiyle ise 49'unda kültürler arası DMF varlıęı saptanmıştır.

Kurnaz (2006) yapmış olduęu çalışmada, Peabody Resim Kelime Testinin madde yanlılıęı açısından incelenmesini içeren bir araştırma yapmıştır.

Araştırma 592 çocuğun cinsiyet ve sosyoekonomik düzey değişkenleri üzerinden Mantel-Haenszel ve Lojistik Regresyon teknikleri kullanılarak yapılmıştır. Ayrıca yanlılık gösterdiği belirlenen maddeler için çeldirici cevap analizi de yapılmıştır. Çalışmada, kullanılan DMF tekniğine göre DMF bulgularının değiştiği sonucuna ulaşılmış, cinsiyet değişkenine göre Mantel-Haenszel tekniğiyle 16, Lojistik Regresyon tekniğiyle 14, sosyoekonomik düzey değişkenine göre Mantel-Haenszel tekniğiyle 33, Lojistik Regresyon tekniğiyle ise 20 maddede DMF varlığı saptanmıştır. Ayrıca Mantel-Haenszel tekniğinin çeldirici yanlılığına daha duyarlı olduğu sonucuna ulaşılmıştır.

Doğan ve Öğretmen (2008) yapmış oldukları çalışmada; DMF belirleme tekniklerinden Ki-Kare, M-H ve Lojistik Regresyon tekniklerini karşılaştırarak, uygulamada ortaya çıkan benzerlik ve farklılıkları belirlemeyi amaçlamışlardır. Çalışmalarını 2003 OKS'nin, 25 maddelik fen alt testi verileri üzerinden yürütmüş ve DMF belirlenmesinde sadece cinsiyet değişkenini temel almışlardır. Çalışma sonucunda; M-H, Ki-Kare ve Lojistik Regresyon tekniklerinin Ki-Kare değerlerinin büyüklüğü bakımından benzer; DMF verdiği belirlenen madde sayısı bakımından ise farklı sonuçlar ürettiği bulunmuştur.

Kalaycıoğlu (2008) yapmış olduğu çalışmada, 599330 lise son sınıf öğrencisinin katıldığı 2005 yılı ÖSS sınavını, madde yanlılığı açısından incelemiştir. Cinsiyet ve okul türü değişkenleri temel alınarak yapılan çalışmada M-H ve LR teknikleri kullanılmıştır. Çalışma sonucunda M-H ve LR tekniklerinin uyumlu sonuçlar verdiği saptanmıştır. Elde edilen bulgular sonucunda cinsiyet ve okul türü değişkenleri açısından 1 adet fen bilimleri maddesinin yanlılık gösterdiği saptanmıştır.

Acar (2008) yapmış olduğu çalışmada, DMF analizinde kullanılan, Genelleştirilmiş Aşamalı Doğrusal Modelleme (GADM), LR ve MTK-OO tekniklerini karşılaştırmıştır. Çalışmada, 2006 yılı Ortaöğretim Kurumları Sınavına katılan ve basit seçkisiz örnekleme tekniğiyle seçilen 10727 kişilik

öğrenci grubunun, 4 seçenekli 25 çoktan seçmeli test maddesine verdiği cevaplar kullanılmıştır. Karşılaştırma çalışmaları cinsiyet ve bölge değişkenleri üzerinden yürütülmüştür. Çalışma sonucunda, tekniklerin DMF’li maddeyi belirleme bakımından aralarındaki ilişkiye bakıldığında; her üç alt testte cinsiyete göre MTK-OO ile LR arasında anlamlı bir ilişki olduğu, bölge alt gruplarına göre GADM ile LR, MTK-OO ile LR ve GADM ile MTK-OO arasında ise farklı alt testlerde anlamlı ilişkilerin olduğu gözlenmiştir.

Abedlazez (2010) yapmış olduğu çalışmada, 54 maddeden oluşan bir başarı testinin 1228 bireye (656 erkek, 624 kadın) uygulanmasıyla elde edilen veriler üzerinde Ki-Kare, ‘b’ Parametresi Faktarı, Alan İndeksi ve Dönüştürülmüş Madde Güçlüğü indeksi tekniklerini uygulamış ve kullanılan DMF belirleme tekniğine göre DMF bulgularının değiştiği sonucuna ulaşmıştır. Ayrıca yapılan çalışmada Ki-Kare ve ‘b’ parametresi faktarı arası uyumun %85, TID ve alan indeksi teknikleri arası uyumun ise %41 olduğunu bulmuştur.

Gök ve diğ. (2010) yapmış oldukları çalışmada, DMF belirlemede kullanılan M-H ve LR tekniklerini karşılaştırmışlardır. Araştırma, 2005 yılı OKS sınavının sayısal bölümünü oluşturan, her biri 25’er maddelik matematik ve fen bilgisi alt testi verileri üzerinde yürütülmüştür. Çalışma, 777.137 kişilik evrenden basit seçkisiz örneklemeyle seçilen 384 kişilik örneklem üzerinde, cinsiyet ve okul türü değişkenleri temel alınarak yapılmıştır. Çalışma sonucunda M-H ve LR teknikleri arasında, matematik ve fen bilgisi alt testlerinde, cinsiyet ve okul türü değişkenleri açısından genel olarak düşük düzeyde bir uyuma rastlanılmıştır.

Ayan (2011) yapmış olduğu çalışmada, Uluslar arası Öğrenci Değerlendirme Programı (PISA) fen okur yazarlığı alt testinde yer alan maddeler üzerinde Lojistik Regresyon ve Mantel Haenszel teknikleriyle cinsiyet değişkenine göre DMF analizi yapmıştır. Araştırma sonucunda her iki yöntemle belirlenen maddelerin birbirinden tamamen farklı olduğu bulunmuştur.

Santelices ve Wilson (2012) yapmış oldukları çalışmada, madde güçlüğü ve maddenin farklı fonksiyonlaşması arasındaki ilişkiyi Madde Tepki Kuramı modelleriyle (1 Parametrelili Lojistik Model ve 3 Parametrelili Lojistik Model) incelemişlerdir. DMF analizleri, ırk değişkeni temel alınarak Scholastic Aptitude Test (SAT) verileri üzerinde yapılmıştır. Araştırma sonucunda her iki Madde Tepki Kuramı modeli için de DMF belirleme teknikleri arasında orta düzeyli ilişki olduğu bulunmuştur.

Atalay ve diğ. (2012) yapmış oldukları çalışmada, farklı DMF belirleme teknikleriyle farklı sonuçlara ulaşıldığını ve MTK-OO tekniğinin Klasik Test Kuramı'na dayalı diğer tekniklerden çok daha duyarlı olduğunu bulmuşlardır.

Ward ve Bennett (2012) yapmış oldukları çalışmayla, M-H ve Standardizasyon teknikleri arasında pek çok farklılık olduğunu ancak bu tekniklerle oldukça benzer sonuçlara ulaşıldığı ve elde edilen sonuçlar arası korelasyonun genel olarak 0.90 ve üzerinde olduğunu bulmuşlardır.

Wetzel ve diğ. (2013) yapmış oldukları çalışmada, NEO-PO-R kişilik envanterinin maddelerine ait kategorilerin farklı şekillerde ifade edilip (extreme response style, non-extreme response style) puanlanmasıyla elde edilen veriler üzerinde cinsiyet değişkeni açısından Madde Tepki Kuramı'na dayalı tekniklerle DMF analizleri yapmışlar ve maddelerin cevap şekline göre DMF bulgularının değiştiğini bulmuşlardır.

Kan, Sünbül ve Ömür (2013) yapmış oldukları çalışmada, DMF belirlemede kullanılan Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı teknikler arasındaki tutarlığı, çeşitli yıllarda uygulanan Seviye Belirleme Sınavlarının alt testleri üzerinde cinsiyet değişkeni açısından incelemişlerdir. Çalışma 121.137 öğrenci üzerinde yürütülmüştür. DMF belirlemede Dönüştürülmüş Madde Güçlüğü, Mantel-Haenszel, Lojistik Regresyon, Lord'un  $\chi^2$  ve Raju'nun Alan Ölçüsü teknikleri kullanılmıştır. Araştırma sonucunda

Klasik Test Kuramı'na dayalı kullanılan tekniklerin kendi içlerinde, Madde Tepki Kuramı'na dayalı tekniklerin de kendi içlerinde benzerlik gösterdikleri bulunmuştur.

#### **I.4. Araştırmanın Amacı ve Önemi**

Madde ve test yanlılığı bir ölçme aracının geçerliğini tehdit eden önemli unsurlardan biridir. Geçerliği sağlanamayan ölçümlere dayalı olarak verilen kararların ise isabetli olması beklenemez.

Bu nedenle araştırmacıların zaman içerisinde farklı değişkenlere dayalı pek çok DMF belirleme tekniği geliştirdiği görülmektedir. Ancak farklı DMF belirleme tekniklerinin farklı sonuçlar ürettiği ve kullanılan hemen her tekniğin güçlü ve zayıf yanlarının olduğu yapılan alanyazın incelemesinden anlaşılmaktadır.

Araştırmacılar farklı DMF belirleme tekniklerinin güçlü ve zayıf yanlarını belirlemek amacıyla bu tekniklerinin birbirleri ile karşılaştırıldığı pek çok çalışma yapmıştır. Ancak bu çalışmalar genel olarak; farklı örneklem büyüklükleri, yetenek dağılımları, DMF içeren madde sayıları/oranı, DMF düzeyi, maddenin puanlanmasında kullanılan kategorilerin sayıları-ifadeleri, vb. değişkenleri kapsamaktadır. Ayrıca bu çalışmaların genel olarak ikili ve farklı kategorilerde puanlanan maddeler üzerinde yapıldığı görülmektedir. Ağırlıklı puanlanan maddeler ve farklı kuramlara dayalı teknikler için ise aynı durum söz konusu değildir. Alanyazında farklı ağırlıklı puanlama durumlarını ve farklı kuramlara dayalı farklı DMF belirleme tekniklerini kapsayan çalışmaların oldukça sınırlı olduğu görülmektedir.

Ülkemiz alanyazını incelendiğinde, DMF belirleme teknikleri ile ilgili çalışmaların genel olarak ikili puanlama durumunda, bu tekniklerin karşılaştırıldığı çalışmalara yoğunlaştığı gözlenmektedir (Yenal, 1995; Öğretmen, 1995; Yurdugül, 2003; Yurdugül, 2010; Öğretmen, 2006; Kan, 2007; Korkmaz,

2005; Kurnaz, 2006; Acar, 2008; Ateşok Deveci, 2008; Doğan ve Öğretmen, 2008; Kalaycıoğlu, 2008; Öğretmen, 2009; Gök ve diğ., 2010; Ayan, 2011; Çepni, 2011; Kan, Sünbül ve Ömür, 2013). Ulaşılabilen çalışmaların hiçbirinde özellikle ikili ve ağırlıklı puanlama yöntemlerinin dikkate alındığı bir uygulamaya rastlanmamıştır. Bu nedenle farklı puanlama durumlarında, farklı kuramlara dayalı, farklı DMF belirleme teknikleri de göz önünde bulundurularak yapılan kapsamlı bir çalışmanın özellikle ülkemiz alanyazınına önemli katkı sağlayacağı düşünülmektedir.

Bu çalışmada, alanyazında yaygın olarak kullanılan Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı Mantel Haenzel/Mantel, Standardizasyon, Lojistik Regresyon ve Olabilirlik Oran Testi tekniklerinin yine alanyazında yaygın olarak kullanılan ikili, UYDA ve DA puanlama durumlarındaki işleyişlerini incelemek amaçlanmıştır.

Bir başka ifadeyle bu çalışmada; Klasik Test ve Madde Tepki Kuramına dayalı DMF belirleme teknikleriyle, farklı puanlama durumlarında;

- DMF belirlenen maddeler arası uyumun,
- DMF içerdiği belirlenen toplam madde sayıları arasında fark olup olmadığının,
- DMF içeren maddeler arasında fark olup olmadığının,
- farklı teknikler kullanılarak DMF içerdiği belirlenen maddelerin ikili ve ağırlıklı puanlama durumlarında farklılık gösterip göstermediğinin,

incelenmesi amaçlanmıştır.

Bu araştırmanın, kullanılan ölçme araçlarının geçerliklerinin sağlanması ve test geliştirme sürecinin iyileştirilmesi çalışmalarına katkı sağlaması açısından önemli olduğu düşünülmektedir.

## **I.5. Problem Cümlesi**

Maddelerin ikili ve ağırlıklı puanlanması durumunda, Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı Mantel Haenzel/Mantel, Standardizasyon, Lojistik Regresyon ve Olasılık Farkı İndeksi teknikleriyle elde edilen DMF'li maddeler ve sayıları farklılık göstermekte midir?

### **I.5.1. Alt Problemler:**

1. Klasik Test Kuramı'na dayalı olarak geliştirilen M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleriyle,
  - a. İkili puanlama durumunda,
  - b. Uzman yargısına dayalı ağırlıklı puanlama durumunda,
  - c. Deneysel ağırlıklı puanlama durumunda
 belirlenen DMF'li maddelerin dağılımı nasıldır?
2. Madde Tepki Kuramı'na dayalı olarak geliştirilen Olabilirlik Oran Testi tekniğiyle,
  - a. İkili puanlama durumunda,
  - b. Uzman yargısına dayalı ağırlıklı puanlama durumunda,
  - c. Deneysel ağırlıklı puanlama durumunda
 belirlenen DMF'li maddelerin dağılımı nasıldır?
3. Klasik Test ve Madde Tepki Kuramı'na dayalı DMF belirleme teknikleriyle ikili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumlarında belirlenen DMF'li maddeler farklılık göstermekte midir?
  - a. Farklı puanlama durumlarında, M-H/Mantel, Standardizasyon, Lojistik Regresyon ve MTK-OO teknikleriyle elde edilen DMF'li maddeler farklılık göstermekte midir?

- b. Farklı teknikler kullanılarak DMF'li olduğu belirlenen maddeler ikili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumlarında farklılık göstermekte midir?
  - c. Klasik Test ve Madde Tepki Kuramı'na dayalı DMF belirleme teknikleriyle, ikili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumlarında belirlenen DMF'li maddelerin uyum yüzdeleri nasıldır?
4. Klasik Test ve Madde Tepki Kuramı'na dayalı DMF belirleme teknikleriyle, ikili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumlarında belirlenen DMF'li madde sayıları arasında fark var mıdır?

## **I.6. Sayılılar**

- 1) Araştırmada kullanılan Sözel Akıl Yürütme Yeteneği Testi, öğrencilerin gerçek yetenek düzeylerini ortaya koymaktadır.

## **I.7. Sınırlılıklar**

Bu araştırma;

- 1) DMF belirleme tekniklerinden Mantel Haenzsel/Mantel, Standardizasyon, Lojistik Regresyon ve MTK-OO testi teknikleri ile,
- 2) Sözel Akıl Yürütme Yeteneği Testi ve bu testin uygulanmasıyla elde edilen veriler ile,
- 3) İkili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumları ile,
- 4) Çalışma grubunda bulunan öğrencilerin yetenek düzeyleri ile,
- 5) Hesaplamalarda kullanılan programlarla,

sınırlıdır.



## BÖLÜM II

### YÖNTEM

#### II.1. Araştırmanın Türü

Bu araştırma Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı DMF belirleme tekniklerinin, farklı puanlama durumlarında, DMF içerdiği belirlenen maddeler ve DMF belirlenen toplam madde sayıları açısından göstereceği benzerlikleri ve farklılıkları incelemek amacıyla yapılmıştır. Araştırmada, örneklem verilerinden hareketle evrene yönelik bir genelleme yapmak yerine, DMF belirleme teknikleri ve puanlama durumları açısından oluşabilecek benzerlikleri ve farklılıkları incelenmek söz konusu olduğundan, bu araştırmanın 'temel araştırma' niteliğinde olduğu söylenebilir.

#### II.2. Çalışma Grubu

Bu araştırma Gözen Çıtak'ın (2007), 2006-2007 eğitim öğretim yılı bahar döneminde, Hacettepe ve Gazi Üniversitelerinin Ankara il merkezinde bulunan çeşitli fakülte ve bölümlerinde öğrenim gören toplam 1593 öğrenciye tek oturumda uyguladığı Sözel Akıl Yürütme Yeteneği Testi'nden (EK 1) elde edilen veriler üzerinden gerçekleştirilmiştir. Araştırma verileri, Gözen Çıtak'ın (2007) izni (Ek 2) ile kullanılmıştır.

Araştırma verilerin elde edilmesinde öğrencilerden cinsiyet, bölüm, fakülte, üniversite ve ÖSS puan türü bilgileri de alınmıştır. Ancak Tablo 3'den de görüleceği üzere üniversite değişkeni dışındaki demografik değişkenlerin DMF değişkeni olarak kullanılması durumunda yüksek oranda kayıp değer ile karşılaşmaktadır. Bu çalışmada ise yanlılık ve yanlılık kaynakları yerine DMF

tekniklerinin farklı puanlama durumlarındaki işleyişlerini incelemek amaçlandığından Gözen Çıtak'ın (2007) araştırma verileri yalnızca üniversite değişkenine göre dikkate alınmış ve odak grubun Hacettepe Üniversitesi, referans grubun ise Gazi Üniversitesi öğrencileri olarak seçilmesi uygun görülmüştür. Buna göre, verilerin elde edildiği çalışma grubuna ilişkin bazı bilgiler Tablo 3'de yer almaktadır.

**Tablo 3:** Verilerin çeşitli değişkenlere göre elde edilen bazı betimsel değerleri

<b>Değişken</b>			<b>f</b>	<b>%</b>	<b>Ort.</b>	<b>S.S.</b>
Üniversite	Hacettepe	İkili	631	39.6	9.30	3.87
		UYDA			58.40	7.37
		DA			58.70	7.29
	Gazi	İkili	962	61.4	9.60	2.73
		UYDA			58.80	5.59
		DA			57.30	5.52
Toplam Öğrenci Sayısı			1593			
Bölüm	Sayısal	İkili	729	74.6	9.03	3.24
		UYDA			57.81	6.53
		DA			58.19	6.44
	Sözel	İkili	248	25.4	9.75	2.86
		UYDA			59.16	5.59
		DA			59.50	5.55
Toplam Öğrenci Sayısı			977			
Cinsiyet	Erkek	İkili	288	32.7	9.22	3.2
		UYDA			58.39	6.32
		DA			58.80	6.13
	Kadın	İkili	591	67.3	9.91	3.24
		UYDA			59.71	6.07
		DA			60.07	5.96
Toplam Öğrenci Sayısı			879			

Tablo 3 incelendiğinde Gazi Üniversitesinden 962 öğrenci (%61.4), Hacettepe Üniversitesinden ise 631 öğrenci (%39.6) olmak üzere toplam 1593 öğrencinin çalışmaya katıldığı görülmektedir.

### II.3. Ölçme Aracı

Gözen Çıtak'ın (2007) doktora tez çalışması kapsamında geliştirdiği Sözel Akıl Yürütme Yeteneği Testi, her biri farklı düzeylerde doğruluk içeren ve ağırlıklı olarak da puanlanabilen 4'er seçenekli, 18 çoktan seçmeli maddeden oluşmaktadır. Teste ilişkin bazı istatistikler Tablo 4'de verilmiştir.

**Tablo 4:** Teste ilişkin elde edilen betimsel istatistikler

	Test Formu
Madde Sayısı	18
Cevaplayıcı Sayısı	1593
Aritmetik Ortalama	9.49
Varyans	10.47
Standart Sapma	3.24
Ortanca	10.00
Mod	10.00
En Düşük Puan	1.00
En Yüksek Puan	18.00
Testin Ortalama Güçlüğü	0.53
Çarpıklık	-0.06
Çarpıklığın Standart Hatası	0.061
Basıklık	-0.19
Basıklığın Standart Hatası	0.123

Tablo 4 incelendiğinde; ortanca ve modun aynı olduğu, aritmetik ortalamasının da bu iki ölçüye yakın bir değer aldığı görülmektedir. Ayrıca dağılımın çarpıklık katsayısı (-0.06) standart hatasına (0.061) oranlandığında ve basıklık katsayısı da (-0.19) yine standart hatasına (0.123) oranlandığında elde

edilen deęerler sırasıyla -0.98 ve -1.54 olarak bulunmuştur. Bu bulgular, araştırma verilerinin dağılımının normal olduęu şeklinde yorumlanabilir.

Araştırma kapsamında incelenen test maddelerinin her biri farklı düzeylerde doęruluk içermekte ve '1'-'4' puan arasında puanlanabilmektedir. '4' puan maddenin en doęru cevabını, '1' puan ise yine doęru olan ancak en doęru cevaba en uzak ifadeyi içermektedir. Bu ağırlıklar ise Gözen Çıtak (2007) tarafından DA ve UYDA seçenek ağırlıklandırma yöntemleriyle belirlenmiştir.

Maddelerin (seçeneklerin) UYDA ve DA puanları Tablo 5'de sunulmuştur.

**Tablo 5:** Maddelerin (seeneklerin) ağırlıklı puan dağılımları

Madde No	Seenekler	Puanlar		Madde No	Seenekler	Puanlar	
		UYDA	DA			UYDA	DA
1	A	1	1	10	A	2	2
	B	4	4		B	1	1
	C	3	2		C	3	3
	D	2	3		D	4	4
2	A	3	3	11	A	4	4
	B	4	4		B	3	3
	C	1	2		C	1	2
	D	2	1		D	2	1
3	A	2	2	12	A	1	1
	B	1	1		B	2	2
	C	4	4		C	3	3
	D	3	3		D	4	4
4	A	4	4	13	A	2	2
	B	2	3		B	1	1
	C	3	2		C	4	4
	D	1	1		D	3	3
5	A	2	2	14	A	3	3
	B	4	4		B	1	1
	C	1	1		C	2	2
	D	3	3		D	4	4
6	A	4	4	15	A	2	1
	B	3	3		B	3	3
	C	2	2		C	1	2
	D	1	1		D	4	4
7	A	1	1	16	A	3	2
	B	3	3		B	4	4
	C	4	4		C	1	1
	D	2	2		D	2	3
8	A	1	1	17	A	3	2
	B	3	3		B	4	4
	C	4	4		C	2	3
	D	2	2		D	1	1
9	A	2	2	18	A	4	4
	B	4	4		B	1	2
	C	1	1		C	2	1
	D	3	3		D	3	3

(Gözen Çıtak'dan (2007) alınmıştır)

Gözen Çıtak (2007) testin güvenilirliğinin, '1-0' puanlama için 0.64 (KR-20), UYDA puanlama için 0.68 (Cronbach  $\alpha$ ) ve DA puanlama için 0.69 (Cronbach  $\alpha$ ) olduğu belirtmiştir.

Gözen Çıtak (2007) testin geçerliğine kanıt elde etmek için ise, sayısal ve sözel bölümden öğrenci alan programlara kayıtlı öğrencilerin test puanı ortalamaları arasındaki farkı t testiyle incelemiş ve her üç puanlama yöntemi için de ortalamalar arasında 0.01 düzeyinde anlamlı fark olduğunu ifade etmiştir. Ayrıca her üç puanlama yöntemi için de test verileri ile öğrencilerin Türk Dili Sözlü Anlatım dersi notları arasında hesaplanan Pearson Momentler Çarpımı korelasyon katsayısı ‘1-0’ puanlama için 0.55; UYDA puanlama için 0.52 ve DA puanlama için ise 0.52 olarak bulunmuştur.

#### **II.4. Araştırmanın Değişkenleri**

Bu çalışmada incelenen bağımlı değişkenler, DMF içerdiği belirlenen maddeler ve DMF içeren toplam madde sayısıdır. Araştırmanın bağımsız değişkenleri ise, Klasik Test Kuramı ve Madde Tepki Kuramı’na dayalı DMF belirleme teknikleri ile maddelerin puanlanma yöntemleridir.

#### **II.5. Verilerin Analizi**

Araştırma kapsamında kullanılan verinin DMF analizleri M-H/Mantel, Standardizasyon, Lojistik Regresyon ve MTK-OO teknikleriyle yapılmıştır.

Bilindiği gibi M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleri parametrik tekniklerin karşılaması gereken sayıtları içermemektedir. Ancak MTK-OO, Madde Tepki Kuramına dayalı tekniklerden biri olduğundan, kullanılan verinin MTK’nın temel sayıtları olan tek boyutluluk ve yerel bağımsızlık sayıtlarını karşılaması gerekmektedir (Hambleton ve Swaminathan, 1989; Embretson ve Reise, 2000; Ostini ve Neiring, 2006). Bu nedenle, veri analizinin ilk aşamasında bu sayıtların karşılanıp karşılanmadığı kontrol edilmiştir. Gözen Çıtak (2007), Madde Tepki kuramının temel sayıtlarından biri olan tek boyutluluk sayıtlısını maddeler arası tetrakorik korelasyon matrisine dayalı temel bileşenler analizinden yararlanarak incelemiş ve testte yer alan 18 maddeden ise 17’sinin birinci bileşen altında toplandığını belirlemiştir. Bu bulgu,

tek boyutluluk ve dolayısıyla da yerel bağımsızlık sayıtlarının karşılandığı şekilde yorumlanabilir (Lord, 1980).

Veri analizinin ikinci aşamasında, Madde Tepki Kuramına dayalı analizlerin yapılabilmesi için, model-veri uyumu incelenmiştir. Gözen Çıtak (2007), verilerin bir parametrelili lojistik model için elde edilen  $-2 \log$  olabirlik (likelihood) değerini 36418.73 olarak; ağırlıklı puanlama durumlarında ise  $-2 \log$  olabirlik değerini ağırlıklandırılmış cevap modeline göre 61159.35 olarak belirlemiştir.  $\chi^2$  dağılımı gösteren bu istatistiğin örneklem büyüklüğüne karşı oldukça hassas olması ve büyük örneklerde model-veri uyumunun hemen her model için sağlanamaması nedeniyle; model veri uyumu için  $-2 \log \text{olabilirlik} / (S-1) - 2n(r-1) \leq 3.00$  koşulu dikkate alınmıştır. Burada ‘S’ cevap deseni (response pattern) sayısını,  $n$  madde sayısını,  $r$  ise cevap kategorisi sayısını ifade etmektedir. Bu araştırmanın olası cevap deseni sayısı madde sayısına ve cevap kategorisi sayısına bağlı olarak  $4^{18}$ ’dir. Bock (1997),  $-2 \log \text{olabilirlik} / (S-1) - 2n(r-1) \leq 3.00$  koşulunu sağlayan uyum değerlerinin, model-veri uyumu için yeterli olduğunu belirtilmektedir (akt: Gözen Çıtak, 2007). Bu bulgulardan hareketle, verinin ‘1-0’ puanlama için bir parametrelili lojistik modeli ile; ağırlıklı puanlama modelleri için ise ağırlıklandırılmış cevap modeli ile uyumlu olduğu sonucuna ulaşılmıştır.

Veri analizinin üçüncü aşamasında, araştırma verileri her alt problem için ayrı ayrı analiz edilmiştir. Yapılan analizler her alt problem için aşağıda ayrıntılı olarak açıklanmıştır.

Araştırmanın birinci alt problemini cevaplamak amacıyla, ilk olarak öğrencilerin 18 maddeye vermiş oldukları cevaplar ikili, UYDA ve DA olarak puanlanmıştır.

Araştırmada M-H/Mantel ve Standardizasyon tekniklerine dayalı analizler için Easy-DIF yazılımından, Lojistik Regresyon tekniğine dayalı

analizler için ise, Zumbo'nun yazmış olduğu sentakstan (syntax) (Ek 3) yararlanılmıştır.

M-H/Mantel ve Standardizasyon Teknikleriyle ikili ve ağırlıklı puanlama durumunda DMF'li maddelerin dağılımını ortaya koymak üzere Easy-DIF yazılımında aşağıdaki düzenlemeler gerçekleştirilmiştir.

- 1) Easy-DIF yazılımı, ağırlıklı puanlama durumları için 1, 2, 3 ve 4 olarak belirlenen puanlamayı kabul etmediğinden, madde puanları yazılım uyumluluğu için 0, 1, 2 ve 3 puan formatına dönüştürülmüştür.
- 2) Elde edilen veri matrisinin ilk sütununa bireylerin odak grupta olmaları durumunda '1', referans grupta olmaları durumunda ise '2' sayısı eklenmiştir. Böylece veri matrisinin ilk sütununda bireylerin grubu; 2-19. sütunlarda ise maddelerden alınan puanlar yer almıştır. Bu işlem her üç puanlama durumu için de tekrarlanmıştır.
- 3) Sonraki aşamada, ilgili veri Easy-DIF yazılımına aktarılmış, ikili puanlama için yazılım üzerinde bulunan 'dichotomous' seçeneği; ağırlıklı puanlama için ise 'ordinal' seçeneği işaretlenerek program çalıştırılmıştır.

Easy-DIF yazılımı M-H/Mantel ve Standardizasyon teknikleri için eş zamanlı olarak 'M-H/Mantel, p, Alpha, Delta, S.E.' 'SPD-SMD, Delta, Error' ve odak-referans grupları için madde ortalamalarını verebildiğinden, M-H/Mantel tekniği için M-H/Mantel ve p değerleri; Standardizasyon tekniği için ise SPD (ikili puanlama) ya da SMD (ağırlıklı puanlama) değerleri dikkate alınmıştır.

M-H/Mantel tekniği için elde edilen M-H/Mantel değeri tek biçimli DMF için 1 serbestlik derecesinde  $\chi^2$  dağılımı gösterdiğinden, buradaki 'p' değerinin anlamlı bulunduğu durumlarda ilgili maddenin DMF gösterdiği (Angoff, 1993) kabul edilmiştir.



Standardizasyon tekniğinde her bir madde için ayrı ayrı elde edilen SPD veya SMD değeri dikkate alınmıştır.

Lojistik Regresyon tekniğiyle ise ikili ve ağırlıklı puanlama durumunda DMF’li maddelerin dağılımını ortaya koymak üzere aşağıdaki düzenlemeler gerçekleştirilmiştir.

- 1) İkili puanlama durumu için elde edilen veriler ilk sütunda ‘group’ (odak, referans) diğer sütunlarda ise ‘item’ (madde puanı) ve ‘scale’ (toplam puan) yer alacak şekilde bilgisayar ortamına aktarılmıştır.
- 2) Ağırlıklı puanlama durumlarında Lojistik Regresyon analizlerinin yapılabilmesi için madde puanları 1-0 formatına dönüştürülmüştür. Bu süreçte, maddeden 1 veya 2 puan alanlara ‘0’, 3 veya 4 puan alanlara ise ‘1’ puan atfedilmiştir.

Bu düzenlemelerden sonra, Zumbo’nun yazmış olduğu sentaks (Nagelkerke-R-Square) çalıştırılmıştır<sup>1</sup>. Daha sonra elde edilen çıktı dosyasından ‘Nagelkerke R Square’ değerleri arası fark alınarak  $\Delta R^2$  değerleri hesaplanmış ve elde edilen değerler Tablo 2’de Gierl ve ark tarafından önerilen ölçütler yardımıyla değerlendirilmiştir.

Araştırmanın ikinci alt problemi Madde Tepki Kuramına dayalı incelemeleri içermektedir. Bu nedenle, MTK-OO tekniğine dayalı analizler IRTLRDIF yazılımıyla yapılmıştır<sup>2</sup>. IRTLRDIF yazılımı, olabilirlik oran testini 2

---

<sup>1</sup> Sentaksın aktif hale getirilebilmesi için verilerin bulunduğu dosya masaüstünde açık durumda iken, sentaks dosyası açılmalı (tutulup sürüklenilerek veri dosyasının üstüne bırakılmalı) daha sonra açılan pencereden ‘run’ ve ‘all’ butonlarıyla yazılım çalıştırılmalıdır.

<sup>2</sup> IRTLRDIF yazılımı, IRTLRDIF201.exe dosyasıyla çalıştırılmalıdır. Daha sonra yazılıma; birey sayısı, madde sayısı, dosya adı, kullanılacak MTK modeli (her bir madde için ayrı ayrı), verilerin bulunduğu dosya adı, grup bilgisi, maddelerin bulunduğu sütunların bilgisi, çıktı dosyası adı ve ankor madde numaraları girilmeli (girilmediği durumda incelenilen maddenin dışında kalan diğer maddeler yazılım tarafından ankor olarak atanmaktadır) daha sonra ‘enter’ tuşuna basılarak yazılım çalıştırılmalıdır.

parametrelı lojistik model, 3 parametrelı lojistik model ve ağırlıklandırılmıř cevap modeli doęrultusunda gerekleřtirebilmektedir. İkinici alt probleme iliřkin analizler ařaęıdaki ařamalar izlenerek gerekleřtirilmiřtir.

- 1) İgili puanlama iin elde edilen veriler öncelikle ilk sütünunda grup (odak ve referans) bilgisi, dięer sütünlarda da madde puanları olmak üzere metin belgesi (.txt) formatına dönüřtürölmüř ve veriler yazılıma aktarılarak yazılım alıřtırılmıřtır.
- 2) Yazılım ıktılarından elde edilen  $G^2$  deęerleri parametre sayısı kadarki serbestlik derecesinin  $\chi^2$  tablo deęerleriyle karřılařtırılmıř ve anlamlı bulunan deęerlerde DMF varlıęı olduęu kabul edilmiřtir.
- 3) MTK-OO testiyle hesaplanan  $G^2$  deęerleri, kestirilen parametre sayısı kadarki serbestlik derecesinde  $\chi^2$  daęılımı gösterdięinden, buradaki  $\chi^2$  daęılımının kritik deęerleri, ikili puanlama durumu iin (2 PLM) 5.99 ( $\alpha = 0.05$ ,  $df=2$ ); UYDA ve DA puanlama durumları iin (ACM) ise 7.81 ( $\alpha = 0.05$ ,  $df=3$ ) ve 9.48 ( $\alpha = 0.05$ ,  $df=4$ ) olarak alınmıřtır (Diři, 2012).

Arařtırmanın üüncü alt problemine iliřkin verilerin analizinde farklı tekniklerle belirlenen DMF'li maddelerin farklılık gösterip göstermedięi Cochran's Q ve McNemar testleriyle incelenmiřtir. Cochran's Q testi; ikili, UYDA ve DA puanlama durumlarında, M-H/Mantel, Standardizasyon, Lojistik Regresyon ve MTK-OO teknikleriyle belirlenen DMF'li maddelerin birbirlerinden anlamlı derecede farklılařıp farklılařmadıklarını test etmek amacıyla; McNemar testi ise, Cochran's Q testiyle anlamlı fark bulunduęu takdirde bu farkın hangi tekniklerden kaynaklandıęını incelemek amacıyla

---

kullanılmıştır. Farklı puanlama yöntemlerine dayalı olarak belirlenen belirlenen DMF'li maddelerin farklılık gösterip göstermediği de, benzer şekilde, Cochran's Q ve McNemar testleriyle test edilmiştir. Cochran's Q ve McNemar testlerinin kullanılmasında, DMF belirleme teknikleri ve puanlama yöntemleriyle ortak olarak saptanan maddeler (aynı maddeler) dikkate alınmıştır.

Cochran's Q testinin formülü eşitlik 21'de verilmiştir (Stamidis, 2001).

$$Q = \frac{(k-1)(k \sum c^2 - T^2)}{KT - \sum R^2} \quad (21)$$

k: Grup sayısı

R: Sütun toplamı

C: Satır toplamı

T: Satır ve sütun toplamı

McNemar testinin formülü eşitlik 22'de verilmiştir (Lawless ve Heymann, 1998).

**Tablo 6:** McNemar testi için kullanılan Ki-Kare Testi tablosu

		DMF tekniği a	
		0	1
DMF tekniği b	0	a	b
	1	c	c

$$McNemar \chi^2 = \frac{(|b-c|-1)^2}{b+c} \quad (22)$$

Farklı teknikler ve farklı puanlama yöntemlerine dayalı olarak belirlenen DMF'li maddeler arası uyum ise basit uyum katsayısıyla (yüzdesi) incelenmiştir. Basit uyum katsayısı (yüzdesi), uyumlu madde sayısının toplam madde sayısına oranlanmasıyla elde edilmiştir (Erkuş, 2006).

Araştırmanın dördüncü alt problemine ilişkin verilerin analizinde ise, DMF içerdiği belirlenen toplam madde sayıları arasındaki farkın manidarlığı Ki-Kare testiyle incelenmiştir. Ki-Kare Testinin formülü eşitlik 23'de verilmiştir (Placket, 1983).

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (23)$$

$O_i$ : Gözlenen frekans

$E_i$ : Beklenen frekans

n: Tablodaki hücre sayısı

## BÖLÜM III

### BULGULAR

Bu bölümde, araştırma verilerinin problem ve alt problemlere yönelik analizleri sonucunda elde edilen bulgular yer almaktadır.

Araştırma bulguları sunulmadan önce verilerin bazı betimsel istatistikleri hesaplanmış ve elde edilen değerler Tablo 7’de sunulmuştur.

**Tablo 7:** Odak ve Referans Grupları açısından araştırma verilerine ait bazı betimsel istatistikler

Puanlama Şekli	Odak Grup						Referans Grup					
	n	$\bar{X}$	Sx	Maks.	Min.	Ranj	n	$\bar{X}$	Sx	Maks.	Min.	Ranj
(1-0)	631	9.3	3.87	18	1	17	962	9.6	2.73	17	1	16
UYDA	631	58.4	7.37	72	33	39	962	58.8	5.59	71	34	37
DA	631	58.7	7.29	72	35	37	962	57.3	5.52	71	33	38

Not: Odak Grup (Hacettepe Üniversitesi), Referans Grup (Gazi Üniversitesi)

Tablo 7 incelendiğinde; odak grubun (Hacettepe Üniversitesi Öğrencileri) 631, referans grubun (Gazi Üniversitesi Öğrencileri) ise 962 kişiden oluştuğu ve verilerin toplamda 1593 kişiden elde edildiği görülmektedir. Odak grubun puan ortalaması; ikili puanlama durumu için 9.3 (SS: 3.87), UYDA puanlama durumu için 58.4 (SS: 7.37) ve DA puanlama durumu için ise 58.7 (SS: 7.29) olarak hesaplanmıştır. Referans grubun puan ortalaması ise; ikili puanlama durumu için 9.6 (SS: 2.73), UYDA puanlama durumu için 58.8 (SS: 5.59) ve DA puanlama durumu için ise 57.3 (SS: 5.52) olarak hesaplanmıştır. Odak grup için elde edilen

en yüksek puan; ikili puanlama için 18.00 üzerinden 18.00, UYDA puanlama durumu için 72.00 üzerinden 72.00 ve DA puanlama durumu için 72.00 üzerinden 72.00 olarak bulunmuştur. En düşük puan ise; ikili puanlama durumu için 1.00, UYDA puanlama durumu için 33.00 ve DA puanlama durumu için 35.00 olarak bulunmuştur. Benzer şekilde referans grup için elde edilen en yüksek puan; ikili puanlama için 18.00 üzerinden 17.00, UYDA puanlama durumu için 72.00 üzerinden 71.00 ve DA puanlama durumu için 72.00 üzerinden 71.00 olarak bulunmuştur. En düşük puan ise; ikili puanlama durumu için 1.00, UYDA puanlama durumu için 34.00 ve DA puanlama durumu için 33.00 olarak bulunmuştur.

### **III.1. Alt Problemler Doğrultusunda Elde Edilen Bulgular**

Aşağıda araştırma problemi doğrultusunda elde edilen bulgular alt problemlere göre ayrı ayrı açıklanmıştır.

#### **III.1.1. Birinci Alt Problem Doğrultusunda Elde Edilen Bulgular:**

**1-a.** *Klasik Test Kuramına dayalı olarak geliştirilen M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleriyle, ikili puanlama durumunda belirlenen DMF'li maddelerin dağılımı nasıldır?*

Bu alt probleme cevap aramak amacıyla odak ve referans gruplarına ilişkin madde puanı ortalamaları  $MH\chi^2$  ve  $p$  değerleri, SPD değerleri ve  $\Delta R^2$  değerleri hesaplanmıştır. Elde edilen bulgular Tablo 8'de verilmiştir.

**Tablo 8:** İkili puanlama durumunda, M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleriyle elde edilen bulgular

		Odak Grup Ort.	Ref. Grup Ort.	Klasik Test Kuramı			
				M-H/Mantel		Std.	LR
				$MH \chi^2$	p	SPD	$\Delta R^2$
İkili Puanlama	Madde 1	0.66	0.71	0.23	0.62	0.004	0.001
	Madde 2	0.61	0.64	0.10	0.74	0.008	0.001
	Madde 3	0.75	0.72	6.01	<b>0.01</b>	<b>-0.052</b>	0.008
	Madde 4	0.72	0.80	1.52	0.22	0.021	0.004
	Madde 5	0.48	0.52	0.85	0.36	0.020	0.001
	Madde 6	0.36	0.33	0.57	0.44	-0.018	0.001
	Madde 7	0.48	0.47	0.01	0.91	-0.002	0.001
	Madde 8	0.71	0.73	0.40	0.52	-0.005	0.001
	Madde 9	0.40	0.43	0.91	0.34	0.034	0.005
	Madde 10	0.49	0.41	8.30	<b>0.00</b>	<b>-0.077</b>	0.012
	Madde 11	0.51	0.57	1.62	0.20	0.031	0.002
	Madde 12	0.42	0.48	2.51	0.11	0.040	0.002
	Madde 13	0.35	0.34	0.22	0.63	0.008	0.012
	Madde 14	0.34	0.30	0.00	0.98	0.009	0.004
	Madde 15	0.41	0.44	3.57	0.05	0.040	0.004
	Madde 16	0.57	0.65	1.79	0.18	0.020	0.006
	Madde 17	0.43	0.42	0.00	0.99	-0.009	0.005
	Madde 18	0.61	0.54	14.21	<b>0.00</b>	<b>-0.079</b>	0.016
<b>DMF Belirlenen Toplam Madde Sayısı</b>					3	3	0

Tablo 8 incelendiğinde ikili puanlama durumunda; M-H/Mantel tekniğine göre 3., 10. ve 18. maddelerde anlamlı düzeyde DMF olduğu;

Standardizasyon tekniğine göre; 3., 10. ve 18. maddelerde ihmal edilemeyecek düzeyde DMF olduğu; Lojistik Regresyon tekniğine göre ise maddelerde DMF olmadığı görülmektedir.

Buradan hareketle M-H/Mantel tekniğine göre toplam 3, Standardizasyon tekniğine göre toplam 3, Lojistik Regresyon tekniğine göre ise 0 maddede DMF varlığı görülmektedir.

**1-b.** *Klasik Test Kuramına dayalı olarak geliştirilen M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleriyle, uzman yargısına dayalı ağırlıklı puanlama durumunda belirlenen DMF'li maddelerin dağılımı nasıldır?*

Bu alt probleme cevap aramak amacıyla ikili puanlama durumunda olduğu gibi odak ve referans gruplarına ilişkin madde puanı ortalamaları  $MH \chi^2$  ve  $p$  değerleri, SMD değerleri ve  $\Delta R^2$  değerleri hesaplanmıştır. Elde edilen bulgular Tablo 9'da verilmiştir.



**Tablo 9:** UYDA puanlama durumunda M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleriyle elde edilen bulgular

		Odak Grup Ort.	Ref. Grup Ort.	Klasik Test Kuramı			
				M-H/Mantel		Std.	LR
				M-H	p	SMD	$\Delta R^2$
Uzman Yargısına Dayalı Ağırlıklı Puanlama	Madde 1	2.37	2.49	2.52	0.11	<b>0.061</b>	0.003
	Madde 2	2.28	2.43	5.70	<b>0.01</b>	<b>0.095</b>	0.013
	Madde 3	2.67	2.61	2.85	0.09	<b>-0.058</b>	0.004
	Madde 4	2.47	2.66	7.82	<b>0.00</b>	<b>0.079</b>	0.009
	Madde 5	2.26	2.29	0.48	0.48	0.002	0.001
	Madde 6	2.08	2.04	0.39	0.52	-0.029	0
	Madde 7	2.15	2.11	0.91	0.33	-0.048	0.002
	Madde 8	2.45	2.54	0.14	0.70	0.011	0.001
	Madde 9	2.22	2.20	1.89	0.16	-0.038	0.008
	Madde 10	2.44	2.31	18.63	<b>0.00</b>	<b>-0.139</b>	0.028
	Madde 11	2.28	2.47	11.67	<b>0.00</b>	<b>0.109</b>	0.020
	Madde 12	2.23	2.34	6.50	<b>0.01</b>	<b>0.106</b>	0.006
	Madde 13	1.77	1.79	0.14	0.69	0.014	0.002
	Madde 14	1.95	1.90	0.19	0.65	-0.001	0.001
	Madde 15	2.24	2.27	0.28	0.59	0.013	0.001
	Madde 16	2.23	2.31	0.12	0.72	0.024	0.001
	Madde 17	1.95	1.92	0.00	0.95	-0.004	0.003
	Madde 18	2.26	2.08	15.08	<b>0.00</b>	<b>-0.198</b>	0.022
<b>DMF Belirlenen Toplam Madde Sayısı</b>					6	8	0

(Kategoriler analiz için 1,2,3,4 formatından 0,1,2,3 formatına dönüştürülmüştür (LR tekniği hariç))

Tablo 9 incelendiğinde UYDA puanlama durumunda; M-H/Mantel tekniğine göre 2., 4., 10., 11., 12. ve 18. maddelerin anlamlı düzeyde DMF içerdiği; Standardizasyon tekniğine göre 1., 2., 3., 4., 10., 11., 12. ve 18.

maddelerin ihmal edilemeyecek düzeyde DMF içerdiği; Lojistik Regresyon tekniğine göre ise maddelerin DMF içermediği görülmektedir.

Buradan hareketle M-H/Mantel tekniğine göre toplam 6, Standardizasyon tekniğine göre toplam 8, Lojistik Regresyon tekniğine göre ise 0 maddede DMF varlığı görülmektedir.

**1-c.** *Klasik Test Kuramına dayalı olarak geliştirilen M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleriyle, deneysel ağırlıklı puanlama durumunda belirlenen DMF'li maddelerin dağılımı nasıldır?*

Bu alt probleme cevap aramak amacıyla UYDA puanlama durumunda olduğu gibi yine odak ve referans gruplarına ilişkin madde puanı ortalamaları  $MH\chi^2$  ve  $p$  değerleri, SMD değerleri ve  $\Delta R^2$  değerleri hesaplanmıştır. Elde edilen bulgular Tablo 10'da verilmiştir.

**Tablo 10:** DA puanlama durumunda M-H/Mantel, Standardizasyon ve Lojistik Regresyon teknikleriyle elde edilen bulgular

		Odak Grup Ort.	Ref. Grup Ort.	Klasik Test Kuramı			
				M-H/Mantel		Std.	LR
				M-H	p	SMD	$\Delta R^2$
<b>Deneysel Ağırlıklı Puanlama</b>	Madde 1	2.35	2.45	0.91	0.33	0.018	0.001
	Madde 2	2.26	2.42	7.49	<b>0.00</b>	<b>0.091</b>	0.011
	Madde 3	2.67	2.61	3.21	0.07	<b>-0.061</b>	0.005
	Madde 4	2.40	2.63	9.18	<b>0.00</b>	<b>0.096</b>	0.013
	Madde 5	2.26	2.29	0.19	0.65	-0.004	0.001
	Madde 6	2.08	2.04	0.53	0.46	-0.036	0.001
	Madde 7	2.15	2.11	0.78	0.37	<b>-0.055</b>	0.001
	Madde 8	2.45	2.54	0.11	0.72	0.016	0.001
	Madde 9	2.22	2.20	3.27	0.07	-0.050	0.011
	Madde 10	2.44	2.31	18.72	<b>0.00</b>	<b>-0.141</b>	0.032
	Madde 11	2.31	2.46	6.98	<b>0.00</b>	<b>0.069</b>	0.017
	Madde 12	2.23	2.34	6.00	<b>0.01</b>	<b>0.112</b>	0.008
	Madde 13	1.77	1.79	0.12	0.71	0.019	0.002
	Madde 14	1.95	1.90	0.41	0.52	-0.007	0.003
	Madde 15	2.24	2.29	1.61	0.20	0.038	0.001
	Madde 16	2.18	2.34	2.73	0.09	<b>0.071</b>	0.007
	Madde 17	2.23	2.25	0.35	0.55	0.015	0.002
	Madde 18	2.40	2.21	21.05	<b>0.00</b>	<b>-0.190</b>	0.024
<b>DMF Belirlenen Toplam Madde Sayısı</b>					6	9	0

Tablo 10 incelendiğinde DA puanlama durumunda; Mantel-Haenszel tekniğine göre 2., 4., 11., 10., 12 ve 18. maddelerin anlamlı düzeyde DMF içerdiği; Standardizasyon tekniğine göre 2., 3., 4., 7., 10., 11., 12., 16. ve 18. maddelerin

ihmal edilemeyecek düzeyde DMF içerdiği; Lojistik Regresyon tekniğine göre ise maddelerin DMF içermediği görülmektedir.

Buradan hareketle M-H/Mantel tekniğine göre toplam 6, Standardizasyon tekniğine göre toplam 9, Lojistik Regresyon tekniğine göre ise 0 maddede DMF varlığı görülmektedir.

### **III.1.2. İkinci Alt Problem Doğrultusunda Elde Edilen Bulgular:**

**2-a.** *Madde Tepki Kuramına dayalı olarak geliştirilen Olabilirlik Oran Testi tekniğiyle, ikili puanlama durumunda belirlenen DMF'li maddelerin dağılımı nasıldır?*

Daha önceden de belirtildiği gibi MTK-OO tekniğinde, odak ve referans gruplar için kestirilen madde parametreleri arasında manidar farklılık olup olmadığı incelenmekte ve parametreler arası fark olmadığına dair yokluk hipotezi test edilmektedir. MTK-OO testiyle hesaplanan  $G^2$  değeri, kestirilen parametre sayısı kadar serbestlik derecesinde  $\chi^2$  dağılımının kritik değeri ile karşılaştırılmakta ve anlamlı bulunan değer DMF varlığını göstermektedir.

Bu alt probleme cevap aramak amacıyla da ikili puanlama durumunda odak ve referans gruplarına ilişkin madde puanı ortalamaları ve  $G^2$  değerleri hesaplanmıştır. Elde edilen bulgular Tablo 11'de verilmiştir.

**Tablo 11:** İkili puanlama durumunda Olabilirlik Oran Testi tekniğiyle elde edilen bulgular

		Odak Grup Ort.	Ref. Grup Ort.	Madde Tepki Kuramı
				Olabilirlik Oran Testi
				$G^2$
<b>İkili Puanlama</b>	Madde 1	0.66	0.71	0.62
	Madde 2	0.61	0.64	0.52
	Madde 3	0.75	0.72	<b>8.92</b>
	Madde 4	0.72	0.80	5.82
	Madde 5	0.48	0.52	1.82
	Madde 6	0.36	0.33	1.92
	Madde 7	0.48	0.47	0.62
	Madde 8	0.71	0.73	3.32
	Madde 9	0.40	0.43	5.42
	Madde 10	0.49	0.41	<b>15.12</b>
	Madde 11	0.51	0.57	2.92
	Madde 12	0.42	0.48	3.82
	Madde 13	0.35	0.34	<b>13.02 (TBO)</b>
	Madde 14	0.34	0.30	2.52
	Madde 15	0.41	0.44	4.02
	Madde 16	0.57	0.65	2.32
	Madde 17	0.43	0.42	2.32
	<b>Madde 18</b>	0.61	0.54	<b>23.92 (TBO)</b>
<b>DMF Belirlenen Toplam Madde Sayısı</b>				4

(TBO: Tek Biçimli Olmayan Değişen Madde Fonksiyonu durumu)

Tablo 11'den ikili puanlama durumunda MTK-OO tekniğiyle elde edilen  $G^2$  değerleri incelendiğinde; 3., 10., 13. ve 18. maddelerin ihmal edilemeyecek

düzeyde DMF içerdiği görülmektedir. Ayrıca 13. ve 18. maddelerde TBO DMF durumu varlığı görülmektedir.

**2-b.** *Madde Tepki Kuramına dayalı olarak geliştirilen Olabilirlik Oran Testi tekniğiyle uzman yargısına dayalı ağırlıklı puanlama durumunda belirlenen DMF'li maddelerin dağılımı nasıldır?*

Bu alt probleme cevap aramak amacıyla yine odak ve referans gruplarına ilişkin madde puanı ortalamaları ve  $G^2$  değerleri hesaplanmıştır. Elde edilen bulgular Tablo 12'de verilmiştir.

**Tablo 12:** UYDA puanlama durumunda Olabilirlik Oran Testi tekniğiyle elde edilen bulgular

	Odak Grup Ort.	Ref. Grup Ort.	Madde Tepki Kuramı	
			Olabilirlik Oran Testi	
			G <sup>2</sup>	
<b>Uzman Yargısına Dayalı Ağırlıklı Puanlama</b>	Madde 1	2.37	2.49	8.24
	Madde 2	2.28	2.43	<b>23.54</b>
	Madde 3	2.67	2.61	6.74
	Madde 4	2.47	2.66	<b>25.34</b>
	Madde 5	2.26	2.29	3.54
	Madde 6	2.08	2.04	3.44
	Madde 7	2.15	2.11	3.24
	Madde 8	2.45	2.54	7.84
	Madde 9	2.22	2.20	<b>10.44</b>
	Madde 10	2.44	2.31	<b>17.14</b>
	Madde 11	2.28	2.47	<b>32.74</b>
	Madde 12	2.23	2.34	<b>13.34</b>
	Madde 13	1.77	1.79	<b>13.04 (TBO)</b>
	Madde 14	1.95	1.90	<b>13.64 (TBO)</b>
	Madde 15	2.24	2.27	7.74
	Madde 16	2.23	2.31	6.54
	Madde 17	1.95	1.92	3.54
	Madde 18	2.26	2.08	<b>25.14</b>
<b>DMF Belirlenen Toplam Madde Sayısı</b>			9	

Tablo 12'den UYDA puanlama durumunda MTK-OO tekniğiyle elde edilen G<sup>2</sup> değerleri incelendiğinde; 2., 4., 9., 10., 11., 12., 13., 14. ve 18.

maddelerin ihmal edilemeyecek düzeyde DMF içerdiği görülmektedir. Ayrıca 13. ve 14. maddelerde TBO DMF durumu varlığı görülmektedir.

**2-c.** *Madde Tepki Kuramına dayalı olarak geliştirilen Olabilirlik Oran Testi tekniğiyle, deneysel ağırlıklı puanlama durumunda belirlenen DMF'li maddelerin dağılımı nasıldır?*

Bu alt probleme cevap aramak amacıyla yine UYDA puanlama durumunda olduğu gibi odak ve referans gruplarına ilişkin madde puanı ortalamaları ve  $G^2$  değerleri hesaplanmıştır. Olabilirlik Oran Testi tekniğiyle elde edilen bulgular Tablo 13'de verilmiştir.



**Tablo 13:** DA puanlama durumunda Olabilirlik Oran Testi tekniğiyle elde edilen bulgular

		Odak Grup Ort.	Ref. Grup Ort.	Madde Tepki Kuramı
				Olabilirlik Oran Testi
				G <sup>2</sup>
<b>DeneySEL Ağırlıklı Puanlama</b>	Madde 1	2.35	2.45	6.04
	Madde 2	2.26	2.42	<b>24.44</b>
	Madde 3	2.67	2.61	5.14
	Madde 4	2.40	2.63	<b>26.14</b>
	Madde 5	2.26	2.29	3.84
	Madde 6	2.08	2.04	2.94
	Madde 7	2.15	2.11	3.44
	Madde 8	2.45	2.54	7.84
	Madde 9	2.22	2.20	<b>11.44</b>
	Madde 10	2.44	2.31	<b>17.14 (TBO)</b>
	Madde 11	2.31	2.46	<b>32.54</b>
	Madde 12	2.23	2.34	<b>13.54</b>
	Madde 13	1.77	1.79	<b>13.64 (TBO)</b>
	Madde 14	1.95	1.90	<b>13.14 (TBO)</b>
	Madde 15	2.24	2.29	9.14
	Madde 16	2.18	2.34	<b>25.54</b>
	Madde 17	2.23	2.25	8.14
	Madde 18	2.40	2.21	<b>21.64</b>
<b>DMF Belirlenen Toplam Madde Sayısı</b>				10

Tablo 13'den DA puanlama durumunda MTK-OO tekniğiyle elde edilen G<sup>2</sup> değerleri incelendiğinde 2., 4., 9., 10., 11., 12., 13. 14., 16. ve 18. maddelerin

ihmal edilemeyecek düzeyde DMF içerdiği görülmektedir. Ayrıca 10., 13. ve 14. maddelerde TBO DMF durumu varlığı görülmektedir.

### III.1.3. Üçüncü Alt Problem Doğrultusunda Elde Edilen Bulgular:

*Klasik Test ve Madde Tepki Kuramına dayalı DMF belirleme teknikleriyle ikili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumlarında belirlenen DMF’li maddeler farklılık göstermekte midir?*

**3a.** *Farklı puanlama durumlarında, M-H/Mantel, Standardizasyon, Lojistik Regresyon ve MTK-OO teknikleriyle elde edilen DMF’li maddeler farklılık göstermekte midir?*

3a alt problemine cevap bulmak amacıyla; ikili, UYDA ve DA puanlama durumlarında, DMF’li maddeler açısından Klasik Test Kuramı ve Madde Tepki Kuramı’na dayalı DMF belirleme teknikleri arasında fark olup olmadığı, Cochran’s Q testiyle incelenmiştir. Yapılan analizlere hiçbir maddede DMF bulamaması nedeniyle Lojistik Regresyon tekniği dâhil edilmemiştir. Elde edilen bulgular Tablo 14’de sunulmuştur.

**Tablo 14:** İkili, UYDA ve DA puanlama durumlarında, belirlenen Değişen Madde Fonksiyonu içeren maddeler açısından Klasik Test Kuramı ve Madde Tepki Kuramı’na dayalı Değişen Madde Fonksiyonu belirleme teknikleri arası farka ilişkin Cochran’s Q testi sonuçları

Puanlama Durumları	M-H/Mantel, SPD/SMD ve MTK-OO	
	Cochran’s Q	p
İkili Puanlama	2.00	0.37
UYDA	2.80	0.24
DA	4.33	0.11

Tablo 14 incelendiğinde; Cochran’s Q testi sonuçlarına göre, ikili puanlama durumunda teknikler arasında anlamlı fark olmadığı ( $p=0.37$ ,  $p>0.05$ );

UYDA puanlama durumunda teknikler arasında anlamlı fark olmadığı ( $p=0.24$ ,  $p>0.05$ ); DA puanlama durumunda teknikler arasında anlamlı fark olmadığı ( $p=0.11$ ,  $p>0.05$ ) görülmektedir.

**3b.** *Farklı teknikler kullanılarak DMF’li olduğu belirlenen maddeler ikili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumlarında farklılık göstermekte midir?*

3b alt problemine cevap bulmak amacıyla; Klasik Test Kuramı ve Madde Tepki Kuramı’na dayalı DMF belirleme teknikleri kullanıldığında, belirlenen DMF’li maddeler açısından ikili, UYDA ve DA puanlama durumları arasında fark olup olmadığı Cochran’s Q ve McNemar testleriyle incelenmiştir. Yapılan analizlere hiçbir maddede DMF bulamaması nedeniyle Lojistik Regresyon tekniği dâhil edilmemiştir. Elde edilen bulgular Tablo 15’de sunulmuştur.

**Tablo 15:** Klasik Test Kuramı ve Madde Tepki Kuramı’na dayalı Değişen Madde Fonksiyonu belirleme teknikleri kullanıldığında, belirlenen Değişen Madde Fonksiyonu içeren maddeler açısından ikili, UYDA ve DA puanlama durumları arası farka ilişkin McNemar ve Cochran’s Q testi sonuçları

DMF Belirleme Teknikleri	İkili Puanlama, UYDA ve DA		İkili Puanlama ve UYDA	İkili Puanlama ve DA	UYDA ve DA
	Cochran’s Q	p	McNemar p	McNemar p	McNemar p
M-H/Mantel	3.60	0.16	-	-	-
SPD/SMD	8.85	0.01	0.06	0.03	1.00
MTK-OO	7.75	0.02	0.12	0.07	1.00

Tablo 15 incelendiğinde; Cochran’s Q testi sonuçlarına göre, M-H/Mantel tekniğiyle, puanlama durumları arasında anlamlı fark olmadığı ( $p=0.16$ ,  $p>0.05$ );

Standardizasyon tekniđiyle, puanlama durumları arasında anlamlı fark olduđu ( $p=0.01$ ,  $p<0.05$ ); MTK-OO tekniđiyle, puanlama durumları arasında anlamlı fark olduđu ( $p=0.02$ ,  $p<0.05$ ) görölmektedir.

Elde edilen bu farkın hangi puanlama durumlarından kaynaklandığını bulmak amacıyla ise McNemar testi yapılmış ve puanlama durumları birbirleri ile ikili olarak karşılaştırılmıştır. Bu doğrultuda Tablo 15 incelendiğinde McNemar testi sonuçlarına göre;

Standardizasyon tekniđi kullanıldığında;

-İkili ve UYDA puanlama durumları arasında anlamlı fark olmadığı ( $p=0.06$ ,  $p>0.05$ ),

-İkili ve DA puanlama durumları arasında anlamlı fark olduđu ( $p=0.03$ ,  $p<0.05$ ),

-UYDA ve DA puanlama durumları arasında anlamlı fark olmadığı ( $p=1.00$ ,  $p>0.05$ ),

MTK-OO tekniđi kullanıldığında;

-İkili ve UYDA puanlama durumları arasında anlamlı fark olmadığı ( $p=0.12$ ,  $p>0.05$ ),

-İkili ve DA puanlama durumları arasında anlamlı fark olmadığı ( $p=0.07$ ,  $p>0.05$ ),

-UYDA ve DA puanlama durumları arasında anlamlı fark olmadığı ( $p=1.00$ ,  $p>0.05$ ),

görölmektedir.

**3c.** *Klasik Test ve Madde Tepki Kuramı'na dayalı DMF belirleme teknikleriyle, ikili, uzman yargısına dayalı ve deneysel ağırlıklı puanlama durumlarında belirlenen DMF'li maddelerin uyum yüzdeleri nasıldır?*

3c. alt problemine cevap bulabilmek amacıyla Klasik Test ve Madde Tepki Kuramına dayalı DMF belirleme teknikleriyle, ikili, UYDA ve DA puanlama durumlarında belirlenen DMF'li maddeler arası basit uyum katsayıları (yüzdeleri) hesaplanmıştır.

Elde edilen bulgular Tablo 16'da verilmiştir.

**Tablo 16:** Klasik Test ve Madde Tepki Kuramına dayalı Değişen Madde Fonksiyonu belirleme teknikleriyle, ikili, UYDA ve DA puanlama durumlarında belirlenen Değişen Madde Fonksiyonu içeren maddeler arası uyum katsayıları (yüzdeleri)

<b>Teknikler Arası Uyum</b>	<b>M-H/Mantel ve Std.</b>	<b>M-H/Mantel ve MTK-OO</b>	<b>Std. ve MTK-OO</b>	<b>M-H/Mantel, Std. ve MTK-OO</b>
İkili Puanlama	%100	%94	%94	%94
UYDA Puanlama	%89	%83	%78	%72
DA Puanlama	%83	%78	%78	%67
<b>Puanlamalar Arası Uyum</b>	<b>İkili ve UYDA</b>	<b>İkili ve DA</b>	<b>UYDA ve DA</b>	<b>İkili, UYDA ve DA</b>
M-H/Mantel	%72	%78	%100	%72
Standardizasyon	%72	%67	%83	%61
LR	%100	%100	%100	%100
MTK-OO	%61	%56	%94	%56

Tablo 16'dan teknikler arası uyum kısmı incelendiğinde,

-İkili puanlama durumunda, en yüksek uyumun M-H/Mantel ve SPD teknikleri arasında (%100) olduğu, diğer teknikler arası uyumun ise M-H/Mantel ve SPD arası uyumdan düşük ve birbirleri ile aynı olduğu (%94),

-UYDA puanlama durumunda, en yüksek uyumun M-H/Mantel ve SPD teknikleri arasında olduğu (%89) ancak bu uyumun ikili puanlama durumunda sağlanan uyumdan daha düşük olduğu, ayrıca M-H/Mantel ile MTK-OO, SPD ile MTK-OO ve her üç teknik arası uyumun ikili puanlama durumundakinden daha düşük olduğu,

-DA puanlama durumunda ise, en yüksek uyumun yine M-H/Mantel ve SPD teknikleri arasında olduğu (%83), ancak diğer puanlama durumlarına oranla teknikler arası uyumda genel bir düşüş olduğu,

görülmektedir.

Bu kısımda; Lojistik Regresyon tekniğiyle araştırma kapsamında kullanılan bütün puanlama durumlarında DMF içeren madde saptanamadığından, teknikler arası uyuma Lojistik Regresyon tekniğine ait bulgular dâhil edilmemiştir.

Tablo 16'dan puanlamalar arası uyum kısmı incelendiğinde,

-M-H/Mantel tekniğinde en yüksek uyumun UYDA ve DA puanlama durumları arasında (%100) olduğu, ayrıca ikili ve UYDA (%72) ile ikili ve DA (%78) puanlama durumları arası uyumun UYDA ve DA puanlama durumları arası uyumdan düşük olduğu,

-Standardizasyon ve tekniğinde en yüksek uyumun UYDA ve DA puanlama durumları arasında (%83) olduğu, ikili ve UYDA (%72) ile ikili ve DA (%67) puanlama durumları arası uyumun UYDA ve DA puanlama durumları arası uyumdan düşük olduğu, ayrıca standardizasyon tekniğiyle farklı puanlama durumlarında elde edilen uyumun M-H/Mantel tekniğiyle elde edilenden yüksek olduğu,

-Lojistik Regresyon tekniđiyle bütn puanlama durumlarında hibir maddede DMF varlıđı saptanamadıđından puanlama durumları arası uyumun %100 olarak belirlendiđi,

-MTK-OO tekniđinde ise yine en yksek uyumun UYDA ve DA puanlama durumları arasında (%94) olduđu, ikili ve UYDA (%61) ile ikili ve DA (%56) puanlama durumları arası uyumun UYDA ve DA puanlama durumları arası uyumdan dřk olduđu, ayrıca ikili ve ađırlıklı puanlama durumları iin sađlanan en dřk uyumun MTK-OO tekniđiyle elde edildiđi,

grlmektedir.

#### **III.1.4. Drdnc Alt Problem Dođrultusunda Elde Edilen Bulgular:**

*Klasik Test ve Madde Tepki Kuramına dayalı DMF belirleme teknikleriyle, ikili ve ađırlıklı puanlama durumlarında belirlenen DMF'li madde sayıları arasında fark var mıdır?*

Klasik Test ve Madde Tepki Kuramına dayalı DMF belirleme teknikleriyle, ikili ve ađırlıklı puanlama durumlarında belirlenen DMF'li madde sayıları Tablo 17'de sunulmuřtur.

Bu kısımda; Lojistik Regresyon tekniđiyle arařtırma kapsamında kullanılan btn puanlama durumlarında DMF ieren madde saptanamadıđından, yapılan analizlere Lojistik Regresyon tekniđine ait bulgular dhil edilmemiřtir.

**Tablo 17:** Klasik Test ve Madde Tepki Kuramına dayalı Değişen Madde Fonksiyonu belirleme teknikleriyle, ikili, UYDA ve DA puanlama durumlarında belirlenen Değişen Madde Fonksiyonu içeren madde sayıları

Puanlama Durumları	DMF'li Madde Sayısı			
	M-H/Mantel	SPD/SMD	MTK-OO	Toplam
İkili Puanlama	3 (2.5)	3 (3.8)	4 (3.9)	10
UYDA	6 (5.8)	8 (8.2)	9 (9.0)	23
DA	6 (6.6)	10 (9.3)	10 (10.1)	26
Toplam	15	21	23	<b>59</b>

Tablo 17'de görüleceği üzere toplam 18 maddeden,

-İkili puanlama durumunda M-H/Mantel ve SPD tekniğiyle 3'ü, MTK-OO tekniğiyle 4'ü,

-UYDA puanlama durumunda M-H/Mantel tekniğiyle 6'sı, SMD tekniğiyle 8'i, MTK-OO tekniğiyle 9'u,

-DA puanlama durumunda M-H/Mantel tekniğiyle 6'sı, SMD tekniğiyle 10'u ve MTK-OO tekniğiyle de 10'u,

DMF'li olarak saptanmıştır.

Ayrıca ikili puanlama durumundan ağırlıklı puanlama durumuna gidildikçe genel olarak DMF içerdiği belirlenen madde sayılarında bir artış olduğu görülmektedir.

Bu bilgiler doğrultusunda, dördüncü alt probleme cevap bulmak amacıyla Tablo 17'de bulunan veriler üzerinde  $\chi^2$  testi yapılmak istenmiş ancak Tablo 17'den de anlaşılacağı üzere beklenen değeri 5'in altında olan hücre sayısı, toplam hücre sayısının %20'sinin üzerine (%33) bulunmuştur.



Analizin tamamlanabilmesi ve ilgili alt probleme cevap bulunabilmesi amacıyla Klasik Test Kuramı'na dayalı DMF belirleme tekniklerine (M-H/Mantel, SPD/SMD ve Lojistik Regresyon) ve ağırlıklı puanlama durumlarına ait hücrelerin kendi içlerinde birleştirilmesine karar verilmiştir. Birleştirme işlemi yapılırken her üç teknikle de ve her iki ağırlıklı puanlama durumunda da DMF içerdiği belirlenen maddelerin 1 kez sayılmasına dikkat edilmiş, bir başka ifadeyle kümelerin birleşimi alınmıştır.

Yapılan birleştirme işlemleri sonucunda elde edilen yeni değerler Tablo 18'de sunulmuştur.

**Tablo 18:** Klasik Test ve Madde Tepki Kuramına dayalı Değişen Madde Fonksiyonu belirleme teknikleriyle, ikili ve ağırlıklı puanlama durumlarında belirlenen Değişen Madde Fonksiyonu içeren madde sayılarına ilişkin birleştirilmiş küme

Puanlama Durumları	DMF'li Madde Sayısı		Toplam
	Klasik Test Kuramı	Madde Tepki Kuramı	
İkili Puanlama	3 (3.4)	4 (3.6)	7
Ağırlıklı Puanlama	10 (9.6)	10 (10.4)	20
Toplam	13	14	27

Dördüncü alt probleme cevap bulmak amacıyla Tablo 18'de bulunan veriler üzerinde Fisher'in Kesin  $\chi^2$  Testi yapılmış ve kullanılan DMF belirleme teknikleri ile puanlama durumlarının birbirinden bağımsız olduğu bulunmuştur ( $p=0.55$ ,  $p>0.05$ ).

## BÖLÜM IV

### TARTIŞMA VE YORUM

Yapılan bu araştırmada farklı kuramlara dayalı, farklı DMF belirleme tekniklerinin, farklı puanlama durumlarındaki işleyişlerini incelemek amaçlanmıştır.

Bu amaç doğrultusunda, geliştirilmiş bir ölçme aracının maddeleri ikili, UYDA ve DA olarak puanlanmış ve elde edilen veriler üzerinde Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı tekniklerle DMF analizleri yapılmıştır.

Bu bölümde de yapılan analizler sonucunda elde edilen bulgular tartışılmış-yorumlanmış ve ulaşılan sonuçlara yönelik bazı önerilerde bulunulmuştur.

İkili puanlama durumunda, Klasik Test Kuramı'na dayalı M-H/Mantel ve Standardizasyon teknikleriyle belirlenen DMF'li maddelerin ve DMF içeren toplam madde sayılarının aynı olduğu, iki teknik arası uyumun %100 olduğu ve bu tekniklerle belirlenen maddeler arasında anlamlı fark olmadığı görülmüştür. Lojistik Regresyon tekniğiyle ise hiçbir maddede DMF varlığı saptanamamıştır. Ayrıca Madde Tepki Kuramı'na dayalı MTK-OO tekniğiyle, M-H/Mantel ve Standardizasyon tekniğiyle belirlenen maddeler yine DMF'li olarak saptanmış ancak bunlara ek olarak 1 maddede daha DMF bulunduğu ve M-H/Mantel, Standardizasyon ve MTK-OO teknikleri arası uyumun %94 olarak belirlendiği görülmüştür. Yapılan Cochran's Q testi sonucunda da ikili puanlama durumunda M-H/Mantel, Standardizasyon ve MTK-OO teknikleriyle belirlenen DMF'li maddeler arasında anlamlı fark olmadığı görülmüştür.

Elde edilen bu bulgulardan, belirlenen DMF'li maddeler için ikilli puanlama durumunda, kullanılan teknikler arasında her ne kadar anlamlı fark bulunmasa da Klasik Test Kuramı'na dayalı tekniklerin kendi içlerinde (LR tekniği hariç) Madde Tepki Kuramı'na dayalı teknikle sağlanandan daha fazla uyum sağladığı sonucuna ulaşılmıştır. Bu durumun en büyük nedeni olarak MH/Mantel ve Standardizasyon tekniklerinin kontinjensi tablosu üzerinden hesaplanmaları ve aynı kurama dayalı olmaları gösterilebilir. MTK-OO tekniğiyle ise M-H/Mantel ve Standardizasyon tekniklerine ek olarak 1 maddede daha DMF saptanması ve bu maddenin TBO DMF durumu göstermesi, MTK-OO tekniğinin TBO DMF durumu saptamada daha duyarlı olduğu sonucunu doğurmaktadır. Bu bulguların alanyazında Atalay ve diğ.'nin (2012) yapmış olduğu çalışmanın bulgularıyla paralellik gösterdiği söylenebilir. Bu durumun M-H ve Standardizasyon tekniklerinin tek biçimli olmayan DMF varlığını yakalamadaki yetersizliklerinden, Madde Tepki Kuramı'nın sağladığı genel avantajlardan ve MTK-OO tekniğinde eğim parametrelerinin de hesaplama katılmasından kaynaklandığı düşünülmektedir. Lojistik Regresyon tekniğiyle ise hiçbir maddede DMF saptanamamasının, alanyazında bu teknik için DMF varlığı saptamada kullanılan ölçütlerden kaynaklandığı düşünülmektedir.

Alanyazında farklı tekniklerle farklı sonuçlara ulaşıldığı ancak buna ek olarak Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı tekniklerin kendi içlerinde birbirlerine oranla daha fazla uyumlu olduğu sonucunu gösteren pek çok çalışma bulunmaktadır (Kan, Sünbül ve Ömür ,2013; Spray ve Miller, 1994; Abedlazez, 2010; Ward ve Bennett, 2012; Doğan ve Öğretmen, 2008). Bu araştırmadan elde edilen sonuçların ise bu çalışmalar ile genel olarak uyumlu olduğu görülmektedir.

UYDA puanlama durumunda ise, M-H/Mantel ve Standardizasyon teknikleriyle belirlenen ve DMF içeren maddelerin aynı olduğu ancak standardizasyon tekniğiyle, M-H/Mantel tekniğine ek olarak 2 maddede daha

DMF saptandığı ve iki teknik arası uyumun %89 olarak belirlendiği, M-H/Mantel ve Standardizasyon teknikleriyle belirlenen ve DMF içeren maddeler arasında anlamlı fark olmadığı görülmüştür. Lojistik Regresyon tekniğiyle ise ikili puanlama durumunda olduğu gibi yine hiçbir maddede DMF saptanamamıştır. Madde Tepki Kuramı'na dayalı MTK-OO tekniğiyle ise, M-H/Mantel tekniğine ek olarak 3, Standardizasyon tekniğine ek olarak 1 maddede daha DMF bulunduğu ve M-H/Mantel, Standardizasyon ve MTK-OO teknikleri arası uyumun %72 olarak belirlendiği görülmüştür. Bu bulguların alanyazında Abedlazez'in (2010) yapmış olduğu çalışmanın bulgularıyla paralellik gösterdiği söylenebilir.

Standardizasyon tekniğiyle M-H/Mantel tekniğine ek olarak 2 maddede daha DMF varlığı saptanması ve belirlenen toplam DMF'li madde sayıları açısından Standardizasyon ve MTK-OO tekniklerinin birbirine daha yakın sonuçlar vermesi akla ilk başta Standardizasyon ve MTK-OO arası uyumun M-H/Mantel ve MTK-OO arası uyumdan daha yüksek olduğu düşüncesini getirmektedir. Ancak M-H/Mantel ve MTK-OO teknikleri arası uyumun %83, Standardizasyon ve MTK-OO teknikleri arası uyumun %78, M-H/Mantel ve Standardizasyon teknikleri arası uyumun ise %89 olarak belirlenmesi DMF içerdiği belirlenen maddeler açısından M-H/Mantel ve MTK-OO arası uyumun Standardizasyon ve MTK-OO arası uyumdan daha yüksek olduğunu göstermiştir. Buradan M-H/Mantel tekniğinin MTK-OO tekniği ile daha uyumlu olduğu sonucuna ulaşılmaktadır. Bu duruma M-H/Mantel tekniğinin odds oranları, Standardizasyon tekniğinin ise ' $p$ ' değerleri üzerinden hesaplanmasının neden olduğu düşünülmektedir. Alanyazında da DMF göstergesi olarak yalnızca ' $p$ ' değerlerinin kullanılmasının; ' $p$ ' değerlerinin ortalama grup farklılıklarından (geçerlik kanıtı olan gerçek fark) ve madde ayırt ediciliğinden ( $r_{jk}$ ) etkilenmesi nedeniyle DMF bulguları üzerinde sıkıntı oluşturduğundan bahsedilmektedir (Angoff, 1982; Hunter, 1975; Lord, 1977, akt: Camilli ve Shepard, 1994).

Ayrıca yapılan Cochran's Q testi sonucunda, UYDA puanlama durumunda M-H/Mantel, Standardizasyon ve MTK-OO teknikleriyle belirlenen ve DMF içeren maddeler arasında anlamlı fark olmadığı görülmüştür. Elde edilen bu bulgulardan, belirlenen DMF'li maddeler için UYDA puanlama durumunda da ikili puanlama durumundakine benzer şekilde, kullanılan teknikler arasında her ne kadar anlamlı fark bulunmasa da Klasik Test Kuramı'na dayalı tekniklerin kendi içlerinde (LR tekniği hariç) Madde Tepki Kuramı'na dayalı teknikle sağlanandan daha fazla uyum sağladığı sonucuna ulaşılmıştır. İkili puanlama durumunda olduğu gibi yine bunun en büyük nedeni olarak M-H/Mantel ve Standardizasyon tekniklerinin kontinjensi tablosu üzerinden hesaplanmaları ve aynı kurama dayalı olmaları gösterilebilir.

MTK-OO tekniğiyle ise M-H/Mantel ve Standardizasyon tekniklerinden daha fazla maddede DMF saptanması ve bu maddelerin TBO DMF durumu göstermesi, MTK-OO tekniğinin TBO DMF durumu saptamada UYDA puanlama durumunda da daha duyarlı olduğu sonucunu doğurmaktadır.

Ayrıca UYDA puanlama durumunda ikili puanlama durumundakine oranla kullanılan bütün teknikler için belirlenen DMF'li madde sayılarında 2 katın üzerinde bir artış olduğu görülmüştür. Örneğin M-H/Mantel tekniğinde ikili puanlama durumunda 3, UYDA puanlama durumunda 6; Standardizasyon tekniğinde ikili puanlama durumunda 3, UYDA puanlama durumunda 8; MTK-OO tekniğinde ise ikili puanlama durumunda 4, UYDA puanlama durumunda 9 maddede DMF varlığı saptanmıştır. Yapılan Cochran's Q testi sonucunda da Standardizasyon ve MTK-OO teknikleri için ikili ve UYDA puanlama durumları arasında DMF içerdiği belirlenen maddeler açısından anlamlı fark olduğu görülmüştür. M-H/Mantel tekniği için ise bu fark anlamlı bulunmamıştır. Bu bulguların alanyazında Atalay ve diğ.'nin (2012) bulgularıyla paralellik gösterdiği söylenebilir.

Elde edilen bu bulgulardan, farklı DMF belirleme teknikleriyle ikili ve UYDA puanlama durumlarında aynı test için elde edilen DMF bulgularının değiştiği sonucuna ulaşılmaktadır.

Bu durumun maddelerin ağırlıklı puanlanmasının bireylerin kısmi bilgilerini daha iyi yansıtmasından, yetenek boyutunun daha büyük bir ranjı hakkında bilgi sağlanmasından ve daha fazla-daha güvenilir istatistiksel bilgi elde edilmesinden kaynaklandığı düşünülmektedir (Ostini ve Neiring, 2006; Embretson ve Reise, 2000).

DA puanlama durumunda ise, M-H/Mantel ve Standardizasyon teknikleriyle belirlenen ve DMF içeren maddelerin aynı olduğu ancak standardizasyon tekniğiyle M-H/Mantel tekniğine ek olarak 3 maddede daha DMF varlığı saptandığı ve iki teknik arası uyumun %83 olarak belirlendiği, M-H/Mantel ve Standardizasyon teknikleriyle belirlenen DMF'li maddeler arasında ise anlamlı fark olmadığı görülmüştür. Lojistik Regresyon tekniğiyle ise yine hiçbir maddede DMF saptanamamıştır. Buradan Lojistik Regresyon tekniğinin bu araştırma kapsamında kullanılan bütün puanlama durumlarında en az sayıda DMF saptayan teknik olduğu sonucuna ulaşılmıştır. Bu bulgu Doğan ve Öğretmen'in (2008) yapmış olduğu çalışmanın bulgularıyla örtüşmektedir. Madde Tepki Kuramı'na dayalı MTK-OO tekniğiyle ise, M-H/Mantel tekniğine ek olarak 4, Standardizasyon tekniğine ek olarak ise 1 maddede daha DMF bulunduğu ve M-H/Mantel, Standardizasyon ve MTK-OO teknikleri arası uyumun %67 olarak belirlendiği görülmüştür. Ayrıca M-H/Mantel ve MTK-OO teknikleri arası uyumun %78, M-H/Mantel ve Standardizasyon teknikleri arası uyumun %83, Standardizasyon ve MTK-OO teknikleri arası uyumun ise yine %78 olarak belirlendiği görülmüştür. Yapılan Cochran's Q testi sonucunda ise DA puanlama durumunda M-H/Mantel, Standardizasyon ve MTK-OO teknikleriyle belirlenen ve DMF içeren maddeler arasında anlamlı fark olmadığı görülmüştür.

DA puanlama durumunda da UYDA puanlama durumundakine benzer sonuçlara ulaşılması; ikili, UYDA ve DA puanlama durumlarının tamamında Klasik Test Kuramı'na dayalı tekniklerin kendi içlerinde (LR tekniği hariç) Madde Tepki Kuramı'na dayalı tekniklerle sağlanandan daha fazla uyum sağladığı ve MTK-OO tekniğinin çok daha fazla maddede DMF durumu saptadığı sonucunu doğurmaktadır.

Ayrıca elde edilen bulgulardan, Lojistik Regresyon tekniği hariç araştırma kapsamında kullanılan DMF belirleme tekniklerinin birbirleri ile gösterdikleri uyumun ikili puanlama durumundan UYDA ve DA puanlama durumuna gidildikçe bozulduğu sonucuna ulaşılmıştır. UYDA ve DA puanlama durumlarında belirlenen maddeler arası uyumun ise kullanılan bütün teknikler için oldukça yüksek olduğu ve bu puanlama durumları arasında anlamlı fark oluşmadığı görülmüştür. Standardizasyon tekniği için DMF içerdiği belirlenen maddeler açısından ikili ve UYDA puanlama durumları arasında ise anlamlı fark olduğu görülmüştür. Bu bulgulardan ağırlıklı puanlama durumlarının kendi içlerinde ikili puanlama durumundakine oranla çok daha uyumlu sonuçlar verdiği sonucu çıkarılmaktadır.

Ayrıca ikili ve ağırlıklı puanlama durumlarında Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı tekniklerle belirlenen toplam madde sayılarının anlamlı derecede birbirinden bağımsız olduğu görülmüştür. Bu bulguların alanyazında Abedlazez (2010) yaptığı çalışmanın bulgularıyla paralellik gösterdiği söylenebilir.

Elde edilen bu bulgulardan, aynı maddeler için, aynı DMF belirleme teknikleriyle ikili ve ağırlıklı puanlama durumlarında elde edilen DMF bulgularının değiştiği görülmektedir. Bu bulgulardan DMF'nin yalnızca maddenin değil aynı zamanda kullanılan puanlama yönteminin de bir fonksiyonu olduğu sonucuna ulaşılmaktadır. Bu durumun gerekçesi olarak ise ikili puanlamanın, kullanılan teknikler için bireyler arasında yeterli değişkenlik

oluşturamaması ve bireylerin gerçek yetenek düzeylerini yansıtmakta yetersiz kalması gösterilebilir. Bu bulgular Wetzel ve ark. (2013) ile Gelin ve Zumbo'nun (2003) yapmış olduğu çalışmanın bulgularıyla örtüşmektedir.

Ayrıca bu araştırmada, kullanılan bütün puanlama durumları için en fazla sayıda DMF varlığı MTK-OO tekniğiyle saptanmıştır. Daha önceden de bahsedildiği gibi bu durumun M-H ve Standardizasyon tekniklerinin TBO DMF varlığını yakalamadaki yetersizliklerinden, Madde Tepki Kuramı'nın sağladığı genel avantajlardan ve MTK-OO tekniğinde eğitim parametrelerinin de hesaplamaya katılmasından kaynaklandığı düşünülmektedir. Benzer şekilde Embretsson ve Reise da (2000) Madde Tepki Kuramı'nın genel olarak DMF içeren maddeleri yakalamada Klasik Test Kuramı'na dayalı tekniklerden daha avantajlı olduğunu belirtmiştir.

#### **IV.1. Öneriler**

Bu araştırmada, aynı maddelere yönelik farklı puanlama durumlarının ve farklı DMF belirleme tekniklerinin kullanılmasıyla DMF bulgularının değiştiği görülmüştür. Bu nedenle DMF analizlerinde birden fazla teknik kullanmasının, farklı tekniklerin birbirlerinin eksik yanlarını tamamlaması ve daha doğru sonuçlara ulaşılması adına önemli olduğu ve önerilebileceği düşünülmektedir.

DMF analizi yapılacak verilerin Madde Tepki Kuramı sayıtlılarını karşılayabildiği durumlarda MTK-OO tekniği, ikili ve ağırlıklı puanlama durumlarında M-H/Mantel ve Standardizasyon tekniklerinden daha tutarlı sonuçlar vermesi nedeni ile önerilebilir.

DMF analizi yapılacak verilerin Madde Tepki Kuramı sayıtlılarını karşılayamadığı durumlarda ise odds oranları üzerinden hesaplanması nedeniyle ikili ve ağırlıklı puanlama durumlarında M-H/Mantel tekniğinin kullanılabileceği önerilebilir.



Alanyazından ve yapılan bu arařtırmadan, ikili puanlama durumunun DMF aısından bilgi kaybına neden olduėu gzlendiėinden, mmkn olduėunca aėırlıklı puanlama modellerinin kullanımı nerilebilir.

Bu alıřmada Klasik Test Kuramı'na dayalı M-H/Mantel, Standardizasyon, Lojistik Regresyon ile Madde Tepki Kuramı'na dayalı MTK-OO teknikleri kullanılmıřtır. Bu nedenle bu arařtırmanın farklı tekniklerle de yapılması alanyazına katkı saėlayacaktır.

Bu arařtırma Szel Akıl Yrtme Yeteneėi Testi verileri zerinde gerekleřtirilmiřtir. Benzer alıřmalar farklı disiplinlere ynelik farklı lme araları zerinde de yapılmalıdır. Bunlara ek olarak duyuřsal zelliklerin ve sergilenen davranıřların llmesine ynelik geliřtirilmiř lme araları zerinde de bu tr alıřmaların yapılması nerilebilir.

Ayrıca arařtırma verilerinin tek boyutluluėunun her ne kadar faktr analizi alıřmalarıyla saėlandıėı ifade edilmiř olsa da 3 farklı puanlama yntemine dayalı olarak elde edilen gvenirlik katsayılarının ok yksek olmaması tek boyutluluk durumunu tartıřmalı hale getirmektedir. Ancak DMF varlıėı, tanımı gereėi zaten ortamda bařka bir faktrn varlıėına iřaret etmektedir. Dolayısıyla gvenirliėin dřk olması bir taraftan da DMF varlıėının gstergesi olarak dřnlebilir. Bu nedenle arařtırma bulgularının genellenebilmesi iin benzer alıřmaların farklı gvenirlikteki lme araları yardımıyla da tekrarlanması nerilir.

## KAYNAKÇA

- Abedlazez, N. (2010). Exploring DIF:comparison of CTT and IRT methods. [International Journal of Sustainable Development, Vol. 1, No. 7, pp. 11-46.](#)
- Abu Sayf, F.K. (1979). The scoring for Multiple-Choice Tests: A closer look. *Educational Technology.*
- Acar, T. (2008). *Maddenin Farklı Fonksiyonlaşmasını belirlemede kullanılan Genelleştirilmiş Aşamalı Doğrusal Modelleme, Lojistik Regresyon ve Olabilirlik Oranı tekniklerinin karşılaştırılması.* Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı. Doktora Tezi, Ankara.
- Atalay, K., Gök, B., Kelecioğlu, H., Arslan, N. (2012). Değişen Madde Fonksiyonunun belirlenmesinde kullanılan farklı yöntemlerin karşılaştırılması: bir simülasyon çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi.* 43: 270-281.
- Aiken, L. R. (2000). *Psychological Testing and Assessment.* Allyn and Bacon
- Anastasi, A. ve Urbina, S. (1997). *Psychological Testing.* Prentice Hall.
- Angoff, W.H. (1993). Perspectives On Differential Item Functioning Methodology. Holland ve Wainer (Ed.), *Differential Item Functioning.* Lawrence Erlbaum Associates, Publishers, New Jersey.
- Ateşok Deveci, N. (2008). *Üniversitelerarası Kurul Yabancı Dil Sınavı'nın Madde Yanlılığı bakımından incelenmesi.* Ankara Üniversitesi Eğitim Bilimleri Enstitüsü Eğitimde Psikolojik Hizmetler Anabilim Dalı. Yayınlanmamış Doktora Tezi, Ankara.

- Ayan, C. (2011). *PISA 2009 fen okuryazarlığı alt testinin değişen madde fonksiyonu açısından incelenmesi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı. Yüksek Lisans Tezi, Ankara.
- Bao, H., Dayton, C.M., Hendrickson, A. B. (2009). Differential Item Functioning Amplification and Cancellation in a reading test. *Practical Assessment, Research & Evaluation*. 14;19.
- Baykul, Y. (2000). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. ÖSYM Yayınları, Ankara.
- Camili, G., Shepard, L.A. (1994). *Methods for identifying biased test items*. Sage Publication, London.
- Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. USA: Rinehart and Winston Inc.
- Çepni, Z. (2011). *Değişen Madde Fonksiyonlarının SIBTEST, Mantel Haenzsel, Lojistik Regresyon ve Madde Tepki Kuramı Yöntemleriyle İncelenmesi*. Doktora Tezi, Hacettepe Üniversitesi, Ankara.
- Dişçi, R. (2012). *Temel ve Klinik Biyoistatistik*. İstanbul Tıp Kitabevi.
- Doğan, N. ve Öğretmen, T. (2008). Değişen Madde Fonksiyonunu belirlemede Mantel-Haenzsel, Ki-Kare ve Lojistik Regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, Cilt:33, Sayı:148.
- Dorans, N.J. ve Holland, P.W. (1993). DIF Detection and Description: Mantel-Haenzsel and Standardization. Holland ve Wainer (Ed.), *Differential Item Functioning*. Lawrence Erlbaum Associates, Publishers, New Jersey.
- Embretson, S.E., ve Reise, S.P. (2000). *Item Response Theory for Psychologists*. Lawrence Erlbaum Associates, Publishers, London.

- Erkuş, A. (2003). *Psikometri Üzerine Yazılar*, Psikologlar Derneği Yayınları, Ankara.
- Erkuş, A. (2006). *Sınıf Öğretmenleri İçin Ölçme ve Değerlendirme: Kavramlar ve Uygulamalar*, Ekinoks, Ankara.
- Frary, R.B. (1989). Partial-Credit scoring methods for Multiple-Choice tests. *Applied Measurement In Education*, 2(1), 79-96.
- Gelin, M.N. ve Zumbo, B.D. (2003). Differential item functioning results may change depending on how an item is scored: an illustration with the center for epidemiologic studies depression scale. *Educational and Psychological Measurement*, 63:65, DOI: 10.1177/0013164402239317.
- Gierl, M.J., Jodoin, M.G., Ackerman, T.A. (2000). Performance of Mantel-Haenszel, Simultaneous Item Bias Test, and Logistic Regression When the Proportion of DIF Items is Large. Annual Meeting of the American Educational Research Association (AERA).
- Gonzales, A., Padilla, J.L., Dolores, H., Gomez-Benito, J., Benitez, I. (2010). EASY-DIF: Software for Analyzing Differential Item Functioning Using the Mantel-Haenszel and Standardization Procedures. *Applied Psychological Measurement*. doi:10.1177/0146621610381489.
- Gök, B., Kelecioğlu, H. ve Doğan, N. (2010). Değişen Madde Fonksiyonunu belirlemede Mantel-Haenszel ve Lojistik Regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, Cilt:35, Sayı:156.
- Gözen Çıtak, G. (2007). *Klasik Test Ve Madde-Tepki Kuramlarına Göre Çoktan Seçmeli Testlerde Farklı Puanlama Yöntemlerinin Karşılaştırılması*. Doktora Tezi, Ankara Üniversitesi, Ankara.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.

- Hambleton, R.K. ve Swaminathan, H. (1989). *Item Response Theory: Principles and Applications*. Kluwer-Nijhoff Publishing, Boston.
- Hidalgo, D.M. ve 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: 903, DOI: 10.1177/0013164403261769.
- Holland, P.W. ve Wainer, H. (Editörler). (1993). *Differential Item Functioning*. Lawrence Erlbaum Associates, Publishers, New Jersey.
- Kalaycıođlu, D.B. (2008). *Öđrenci Seçme Sınavı'nın Madde Yanlılıđı açısından incelenmesi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı. Doktora Tezi, Ankara.
- Kan, A. (2007). Test Yansızlıđı: H.Ü. Yabancı Dil Muafiyet Sınavının Cinsiyete ve Bölümlere Göre DMF Analizi. *Eurasian Journal of Educational Research*, 29; 45-58.
- Kan, A., Sünbül, Ö., Ömür, S. (2013). 6.- 8. Sınıf Seviye Belirleme Sınavları alt testlerinin çeşitli yöntemlere göre deđişen madde fonksiyonlarının incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, Cilt 9, Sayı 2, ss. 207-222.
- Korkmaz, M. (2005). *Madde Cevap Kuramına Dayalı Olarak Çok Kategorili Maddelerde Madde ve Test Yanlılıđının (İşlevsel Farklılıđın) İncelenmesi*. Doktora Tezi, Ege Üniversitesi, İzmir.
- Kristjansson, E. (2001). *Detecting DIF In Polytomous Items: An Empirical Comparison Of The Ordinal Logistic Regression, Logistic Discriminant Function Analysis, Mantel And Generalized Mantel Haenszel Procedure*. National Library Of Canada, Ottawa.

- Kristjansson, E., Aylesworth, R., McDowell, I., Zumbo, B.D. (2005). A comparison of four methods for detecting differential item functioning in ordered response items. *Educational and Psychological Measurement*. 65: 935, DOI: 10.1177/0013164405275668.
- Kurnaz, F.B. (2006). *Peabody Resim Kelime Testinin Madde Yanlılığı açısından incelenmesi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı. Yüksek Lisans Tezi, Ankara.
- Lawless, H.Ht, Heymann, H. (1998). *Sensory Evaluation Of Food: Principles and Practices*. Kluwer Academic/Plenum Publishers, New York.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Magnusson, D. (1967). *Test Theory*. Massachusetts. Addison-Wesley.
- Murphy, K.R. ve Davidshofer, C.O. (2005). *Psychological Testing: principles and applications*. Pearson/Prentice Hall.
- Narayanan, P., Swaminathan, H. (1994). Performance of the Mantel-Haenszel and Simultaneous Item Bias Procedures for Detecting Differential Item Functioning, *Applied Psychological Measurement*, Volume: 18, Number:4.
- Padilla, J.L., Hidalgo, J.L., Benitez, I., Gomez-Benito, J. (2012). Comparison of three software programs for evaluating DIF by means of the Mantel-Haenszel procedure; EASY DIF, DIFAS and EZDIF, *Psicologica*, 33,135-156.
- Plackett, R.L. (1983). *Karl Pearson and the Chi-Squared Test*. *International Statistical Review*.51 (1): 59-72.
- Osterling, S.J. (1983). *Test Item Bias*. Sage Publication, London.

- Ostini, R. ve Neiring, M.L. (2006). *Polytomous Item Response Theory Models*. Sage Publications, Thousand Oaks, California.
- Öğretmen, T. (1995). *Differential Item Functioning analysis of the verbal ability section of the first stage of the University Entrance Examination in Turkey*. Ortadoğu Teknik Üniversitesi Sosyal Bilimler Enstitüsü. Yüksek Lisans Tezi, Ankara.
- Öğretmen, T. (2006). *Uluslararası okuma becerilerinde gelişim projesi (PIRLS) 2001 testinin psikometrik özelliklerinin incelenmesi: Türkiye-Amerika Birleşik Devletleri örneği*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı. Yayınlanmamış Doktora Tezi, Ankara.
- Öğretmen, T. (2009). Değişen Madde ve Test Fonksiyonunun Belirlenmesinde Madde Tepki Kuramı'na Dayalı Parametrik Yöntemlerin Karşılaştırılması. *Eğitim ve Bilim*, (34), 152; 113-125.
- Santelices, M.V. ve Wilson, M. (2012). On the relationship between differential item functioning and item difficulty: an issue of methods? item response theory approach to differential item functioning. *Educational and Psychological Measurement* 72(1) 5–36.
- Spray, J., ve Miller, T. (1994). *Identifying nonuniform DIF in polytomously scored test items*. American College Testing Research Report Series 94-1. Iowa City, IA: American College Testing Program.
- Stamatis, D.H. (2001) *Six Sigma and Beyond: Statistics and Probability*. St. Lucie Press.
- Swaminathan, H., Rogers, H.J. (2005). Detecting Differential Item Functioning Using Logistic Regression Procedures. *Journal of Educational Measurement*. 27:4;361-370.

- Thorndike, R.L. (1982). *Applied Psychometrics*. Houghton Mifflin Company, Boston.
- Thissen, D. (2001). *IRTLRDIF (v.2.0b Sürümü)* [Computer Software]. Chapel Hill: L. L. Thurstone Psychometric Laboratory, University of North Carolina at Chapel Hill.
- Van der Linden, W. J. ve Hambleton, R. K. (Editörler). (1997). *Handbook of Modern Item Response Theory*. New York: Springer.
- Ward, W.C., Bennett, R.E. (2012). *Construction Versus Choice in Cognitive Measurement: Issues in Constructed Response, Performance Testing, and Portfolio Assessment*. Routledge, Taylor ve Francis Group, London and New York.
- Wetzel, E., Böhnke, J.R., Carstensen, C.H., Zeigler, M., Ostendorf, F. (2013). Do individual response styles matter? Assessing differential item functioning for men and women in the NEO-PI-R. *Journal of Individual Differences*, Vol 34(2), 2013, 69-81. doi: [10.1027/1614-0001/a000102](https://doi.org/10.1027/1614-0001/a000102).
- Yenal, E. (1995). *Differential Item Functioning analysis of the quantitative ability section of the first stage of the University Entrance Examination in Turkey*. Ortadoğu Teknik Üniversitesi Sosyal Bilimler Enstitüsü. Yayınlanmamış Yüksek Lisans Tezi, Ankara.
- Yurdugül, H. (2003). *Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı'nın madde yanlılığı açısından incelenmesi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı. Yayınlanmamış Doktora Tezi, Ankara.
- Yurdugül, H. (2010). Farklı madde puanlama yöntemlerinin ve farklı test puanlama yöntemlerinin karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1, 1-8.



- Zwick, R. (2012). A Review of ETS Differential Item Functioning Assessment Procedures: Flagging Rules, Minimum Sample Size Requirements, and Criterion Refinement. Research Report. Educational Testing Service.
- 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 ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

## EKLER

## EK:1

## SÖZEL YETENEK TESTİ

**AÇIKLAMA:** Bu test, sözel malzeme ile akıl yürütme yeteneğinizi ölçmek amacıyla hazırlanmıştır. Testte, her biri dört seçenekli 18 çoktan seçmeli soru yer almaktadır. Her soruyu okuduktan sonra doğru bulduğunuz cevabı, test formu ile birlikte dağıtılan cevap kağıdı üzerine kurşun kalemle işaretleyiniz. Lütfen test formu üzerinde işaretleme yapmayınız. Yanıt bulmakta güçlük çektiğiniz sorularda, size en doğru görünen yanıtı işaretleyiniz. Yanıtlama süresi 15 dakikadır.

Araş. Gör. Gökkuş GÖZEN  
A.Ü. Ölçme ve Değerlendirme Anabilim Dalı

1. - 3. sorularda, cümledeki veya parçadaki boşluğu en uygun biçimde tamamlayan seçeneği bulunuz.

1. Dil, insan hem ————— hem ————— yapar.  
A) zengin-yoksul B) rezil-vezir  
C) güçlü-zayıf D) dürüst-yalancı
2. Amacı olmayan gemiyi hiçbir —————.  
A) kılavuz yol göstermez  
B) rüzgar fayda etmez  
C) yolcu binmez  
D) deniz feneri ışık tutmaz
3. Fethi Naci: Ekonomi dünyasından gelip eleştirilenliğin başkôşesine oturmuş bir kişi. İyiye iyi kötüye kötü. Hiç kimseye bir borcu yok. Belli ki yalnız ————— gücüne ağınyor ve yalnız ————— yaptığını düşünüyor.  
A) bileğinin-sorumluluklarını  
B) gazetesinin-gereğini  
C) kaleminin-işini  
D) bilgisinin-doğruluğu
4. "Turnuvada bu akşam karşılaşacak iki takımın da birer galibiyeti bulunuyor." cümlesinden kesin olarak çıkarılabilecek yargı aşağıdakilerden hangisidir?  
A) Turnuvada bu maçtan önce en az iki maç oynamıştır.  
B) İki takım eş sayıda maç yapacaktır.  
C) Bu maç bu iki takım ilk kez karşı karşıya getirecektir.  
D) Bu turnuvadaki maçlar genelde akşam saatlerinde oynanmaktadır.
5. "Şiir öylesine bir dilsel düzenleme değildir ki sözcüklerden biri çıkarılır ya da birinin yeri değiştirilirse şiirin tüm büyüğü bozulur." cümlesinde anlatılmak istenene anlamca en yakın olan ifade aşağıdakilerden hangisidir?  
A) Şiirdeki dilsel düzenlemeler kendine özgüdür.  
B) Şiirde her sözcük, bulunduğu yerde güzeldir.  
C) Şiirin en önemli öğesi sözcüktür.  
D) Şiiri güzelleştiren sözcüklerin uyumudur.
6. Ölim şurada dursun, hayat burada Ben hayattan yana Hayat, üstüne bir şev giymeden uzan yamama Bu dizelerdeli altı çizili sözün yerine aşağıdakilerden hangisinin kullanılması en uygun olur?  
A) doğallığına B) tüm gerçekliğine  
C) yalnızlığına D) sakinliğine
7. Bizim yazımızda deneme türü oldukça cılızdır. Bu durum, dünya yazın için de geçerlidir. Hemen belirtelim ki denememizin cılızlığı nicelikselidir. Yani denememiz az, deneme türünde yazılmış yapıtların sayısı sınırlıdır. Buna karşın nitelik bakımından dünya yazınındaki seçkin deneme örnekleriyle rahatça boy ölçişebilecek yetkinliktedir. Komu yönünden de insanlığı tümüyle kuşatan bir çeşitlilik gösterir. Aynı şeyi, öteki yazın türleri için, örneğin roman için, tiyatro için söyleyemem. Bu parçada, yazımızdaki deneme türünün en çok hangi yönü üzerinde durulmaktadır?  
A) Diğer yazın türlerinden üstünlüğü  
B) Yazınların sayısı azlığı  
C) Niteliksel açıdan zenginliği  
D) Dünya yazınındaki yeri
8. Edebiyatta, düşüncenin ardında bir düşünce daha olması, onun ardında gene bir düşünce olması iyidir. Böyle olmazsa edebiyat bir plastik masaya benzer; dümdüz olur. Oysa bir tahta masaya benzemesi gerekir; tahtanın altında başka damarlar, başka biçimde görümler vardır. Tahtanın derinine indikçe insan başka biçimlere rastlar. Bu paragraf için en uygun başlık aşağıdakilerden hangisidir?  
A) Edebiyatta Anlam  
B) Edebiyatta Düşüncenin Gücü  
C) Edebiyatta Derinlik  
D) Edebiyatın Görünmeyen Yüzü

9. "Ağrıların, türkülerin komuları beni yüreğimden vuran komular oluyor. Bir duruma döneminden sonra, kişileri, olayı titizce inceliyorum. Duygularımı beni ne denli zorlarsa zorlasın, bir kez bile, ağrının, türkülerin kişisini ya da olayımı, omu yetiştiren, oluşurana toprağıyla, tarihiyle öğrenmeden, doğrudan şiirlerime almıyorum."
- Bu parçada özellikle anlatılmak istenen aşağıdakilerden hangisidir?**
- A) Şiirde duyguların önemli bir yeri olduğu  
B) Şairin araştırmacı bir kişiliğe sahip olması gerektiği  
C) Acıklı olayların şairleri derinden etkilediği  
D) Her ağrının, türkülerin bir hikayesi olduğu
10. Geleneklerimizde göre konuk kusmesiyle gelir. Ev sahibi olanlar ölçüsünde elinden geleni yapar. Bu arada dertler paylaşılır, anılar tazelenir, dostluklar pekiştirilir. Kısaca konuk ağırlamak Türk toplumunda dayanışmanın önemli bir göstergesidir. **Parçada konuk ağırlamanın en çok hangi işlevi üzerinde durulmaktadır?**
- A) Sevgi ve saygı ortamı yaratıldığı  
B) Anılara bağlı kalmayı sağladığı  
C) İnsan ilişkilerini güçlendirdiği  
D) Paylaşmayı ve özveriye geliştirdiği
11. O, unutulmuş yazarları hatırlatmayı, onların yapıtlarını yeniden gün ışığına çıkarmayı değil, günümüz yapıtlarını, birer birer, okurlara tanıtmayı amaçlamıştı.
- Aşağıdakilerden hangisi bu cümleye anlamca en yalındır?**
- A) Yapıtları değerlendirirken, geçmişin ürünlerini bugüne taşınma yerine, yeni yazılmış olanları ele almayı isterdi.  
B) Geçmişte ortaya konan yapıtların öneminden değil, bugün nitelikli yapıtların okurla buluşmasından söz ederdi.  
C) Yıllar önce ortaya konmuş ürünlerin, bugün için büyük bir değer taşımadığına inanırdı.  
D) Unutulmuş yazarların tanıtılmasını gerekli olduğunu düşünürdü.
12. "Dağıtmak" sözcüğü aşağıdaki cümlelerde kaç değişik anlamda kullanılmıştır?  
I. Görevliler kalabalığı dağıttı.  
II. Bir yumruktaki beyını dağıttı.  
III. Evi niye böyle çok dağıtmış?  
IV. Yine kendini dağıtmışsın, bir şeye mi üzüldün?  
V. Anlaşamayınca kooperatifi dağıtmak zorunda kaldılar.
- A) 2 B) 3 C) 4 D) 5
13. Bugün eskisi kadar ne mektup yazıyor ne de mektup alıyorum. Okumanın yerini televizyon ekranlarının, mektuplaşmanın yerini telefonum aldığı bir gerçek. Yine de mektup yazarken verilen emeği ve mektuptaki kalıcılığı önemsiyorum. Mektubum aynı zamanda yazınsal bir tür olduğumu unutmamaya gerek var mı? Çocuklarına, gençlerin birbirlerine mektup yazmalarını, telefonla konuşmalarından daha önemli, daha geliştirici buluyorum.
- Bu parçada özellikle aşağıdakilerden hangisi vurgulanmaktadır?**
- A) Eskisine göre daha az mektuplaşıldığı  
B) Mektup yazarken emek harcanıldığı  
C) Gençlerin birbirine mektup yazması gerektiği  
D) Mektubum salt haberleşme aracı sayılmaması doğru bulunmadığı
14. Gösteriş için herkesten farklı giyinmek, farklı kılıklara girmek nasıl pasuruklksa konuşmada bitimmedik, duyulmadık cümleler aramak da bir medreseli çocuk gayretidir. "Ah, keşke Paris'in zerrizatı çarşısında kullanılan sözcüklerle konuşabilsem!" diyor Montaigne. **Bu parçada Montaigne'nin konuşmayla ilgili olarak asıl vurgulanmak istediği aşağıdakilerden hangisidir?**
- A) Halktan kopuk bir dilin yapıldığı  
B) Esnaf rahatlığıyla konuşabilmek  
C) Farklı olmaya çalışmanın gereksizliği  
D) Günlük konuşulan dille konuşabilmek
15. Televizyonda sağlık konusunda konuşan bir doktor şöyle diyordu: "Bence, hazır gıdalardan uzak durarak sebze, meyve, balık gibi doğal ürünlerle beslenmek sağlığımızı korumak için dikkat edilmesi gereken önemli bir noktadır."
- Bu parçada asıl vurgulanmak istenen hangisidir?**
- A) Hazır gıdalardan çok yenememek gerekir.  
B) Doğal gıdalar hazır gıdalardan daha yararlıdır.  
C) Sebze, meyve, beyaz et gibi ürünler hazır gıda sayılmaz.  
D) Sağlığı korumak ancak doğal gıdalarla olasıdır.
16. Hoyratır bu akşam üzerleri daima!  
Gün saltanatıyla gitti mi bir defa  
Yalnızlığımızla doldurup her yeri  
Bir renk çiböğün içinde bahçemizden,  
Bir el çıkarmaya başlar boşcazandan  
Lavanta çiçeği kokan kederleri  
Hoyratır bu akşam üstileri daima!
- A. Muhip Durana'nın bu şiirinde asıl dile getirilmek istenen aşağıdakilerden hangisidir?**
- A) Yaşamın gerçekleri B) Yalnızlık  
C) Doğa D) Umursuzluk

17. Dilimizdeki pek çok yabancı sözcük gibi "Pantik"ın de bir hikayesi var. Eski Yunan mitolojisinde bir tanrı olan "Pan" gravürlerde pan flüt çalarken resmedilir hep. Bu flütün hem mucidi hem de "isim babası"dır. Pan'ın bir başka özelliği ise düşman ordularının korkup kaçmasına neden olan "haykırışı"dır. Dahası Antik Atinalılar Marathon Savaşı'nda Perslerin Pan'ın sesiyle panikleyip kaçıklarına inanırlar. Tüm bu güvültü, toz-dümen ve kaçışmadan geriye dilimizde de kullanılan "panik" kahr.

**Bu parçada asıl anlatılan aşağıdakilerden hangisidir?**

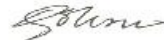
- A) Eski Yunan tanrılarında olan Pan'ın öyküsü  
 B) "Panik" sözcüğünün dilimize yabancı dilden girdiği  
 C) Dilimize giren her kelimenin bir hikayesinin olduğu  
 D) Perslerin Marathon Savaşı'nda Pan'ın haykırışından korktukları
18. Bir toplumdaki:  
 -trafik polisi yoksa kırmızı ışıkta geçilmesi,  
 -piknik yapılan yerlerde, çöplerin çöp kumusuna alınarak açıkta bırakılması,  
 -toplu taşıma araçlarında gençlerin, yaşlılara yer vermemesi  
 gibi davranışların yaygın olması aşağıdakilerden hangisinin en güçlü göstergesidir?
- A) Bazı kuralların gereği gibi benimsenmemiş olduğunun  
 B) Toplumsal değerlerin değişim gösterdiğinin  
 C) Bazı toplumlarda yasaların yetersiz kaldığının  
 D) Eğitim düzeyinin düşük olduğunun

TEST BİTTİ.

EK:2

Sayın İlgili,

"KLASİK TEST VE MADDE-TEPKİ KURAMLARINA GÖRE ÇOKTAN SEÇMELİ TESTLERDE FARKLI PUANLAMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI" başlıklı doktora tez çalışmam kapsamında kullandığım araştırma verilerimin, Mersin Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı doktora programı öğrencisi Hüseyin SELVİ'nin "KLASİK TEST KURAMI VE MADDE TEPKİ KURAMINA DAYALI DMF BELİRLEME YÖNTEMLERİNİN İKİLİ, ÇOKLU VE METRİK PUANLAMA DURUMUNDA İNCELENMESİ" başlıklı tez çalışması kapsamında kullanmasını onaylıyorum. Veriler tarafımdan ilgili kişiye gönderilmiştir. Bilgilerinizi saygılarımla arz ederim.



Yrd.Doç.Dr. Gökse GÖZEN  
Mimar Sinan Güzel Sanatlar Üniversitesi  
Fen Edebiyat Fakültesi  
Eğitim Bilimleri Bölümü

**EK: 3**

GET

FILE='binary.sav'.

EXECUTE .

compute item= item1.

compute total= scale.

compute grp= group.

\* 2 df Chi-squared test and R-squared for the DIF (note that this is a simultaneous test of uniform .

\* and non-uniform DIF).

LOGISTIC REGRESSION VAR=item

/METHOD=ENTER total /METHOD=ENTER grp\*total

/CONTRAST (grp)=Indicator

/CRITERIA PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .

execute.

\* 1 df Chi-squared test and R-squared for uniform DIF.

\* This is particularly useful if one wants to determine the incremental R-squared

\* attributed to the uniform DIF.

LOGISTIC REGRESSION VAR=item

/METHOD=ENTER total /METHOD=ENTER grp

/CONTRAST (grp)=Indicator

/CRITERIA PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .

execute.

## ÖZGEÇMİŞ

Bu tezin yazarı olan Hüseyin SELVİ, 28.10.1984 tarihinde Mersin/Erdemli’de doğmuştur.

İlk ve orta öğrenimini yine bu ilçede tamamlayan yazar, Lisans öğrenimini Fen Bilgisi Öğretmenliği Programı üzerine 2006 yılında Abant İzzet Baysal Üniversitesinde tamamlamıştır. Yüksek Lisans öğrenimini 2009 yılında yine aynı üniversitenin Sosyal Bilimler Enstitüsü, Eğitimde Ölçme ve Değerlendirme Bilim Dalında tamamlayan yazar, Doktora öğrenimini ise 2013 yılında Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitimde Ölçme ve Değerlendirme Bilim Dalında tamamlamıştır.