

**T.C.**  
**Mersin Üniversitesi**  
**Eđitim Bilimleri Enstitüsü**  
**Eđitim Bilimleri Anabilim Dalı**  
**Eđitimde Ölçme Ve Deđerlendirme Bilim Dalı**

**MADDELERİ GÜÇLÜKLERİNE GÖRE FARKLI SIRALAMANIN BİREY TEPKİLERİNE  
ETKİSİNİN DEĐİŐEN MADDE FONKSİYONUyla İNCELENMESİ**

**YÜKSEK LİSANS TEZİ**

**Ebru BALTA**

**MERSİN, 2016**

**T.C.**  
**MERSİN ÜNİVERSİTESİ**  
**EĞİTİM BİLİMLERİ ENSTİTÜSÜ**  
**EĞİTİM BİLİMLERİ ANABİLİM DALI**  
**EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI**

**MADDELERİ GÜÇLÜKLERİNE GÖRE FARKLI SIRALAMANIN BİREY TEPKİLERİNE**  
**ETKİSİNİN DEĞİŞEN MADDE FONKSİYONUyla İNCELENMESİ**

**YÜKSEK LİSANS TEZİ**

**Ebru BALTA**

**TEZ DANIŞMANI**  
**YRD. DOÇ. DR. SEÇİL ÖMÜR SÜNBÜL**

## KABUL VE ONAY

Mersin Üniversitesi Eğitim Bilimleri Enstitüsü Müdürlüğü'ne,

Ebru BALTA tarafından hazırlanan "Maddeleri Güçlüklerine Göre Farklı Sıralamanın Birey Tepkilerine Etkisinin Değişen Madde Fonksiyonuyla İncelenmesi" başlıklı bu çalışma, jürimiz tarafından Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Başkan .....

Doç. Dr. Cem Oktay GÜZELLER

Üye.....

Doç. Dr. Devrim ALICI

Üye.....

Yrd. Doç. Dr. Seçil ÖMÜR SÜNBÜL (Danışman)

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.



## TEŐEKKÜR

Arařtırma konumun belirlenmesinde, planlanmasında ve bir ürün olarak ortaya ıkmasına kadar geen süre boyunca yanımda bulunan, alıřmamın her ařamasını özen ve titizlikle takip eden hiçbir zaman desteęini esirgemeyen, yol gösteren danıřmanım Yrd. Do. Dr. Seil ÖMÜR SÜNÖÜL'e ok teőekkür ederim.

Yüksek lisans eęitimim boyunca ve tez sürecinde büyük yardımlarını gördüğüm, bilgi ve deneyimlerinden yararlandığım Yrd. Do. Dr. Önder SÜNÖÜL'e, yüksek lisans eęitimim boyunca ders aldığım ve arařtırmamla ilgili görüş, dönüt ve düzeltmeleri ile alıřmama katkıda bulunan deęerli hocalarım Do. Dr. Devrim ÖZDEMİR ALICI, Yrd. Do. Dr. Bilge Uzun BAŐUSTA' ya ve tezimde jüri üyesi olarak katkıda bulunan Do. Dr. Cem Oktay GÜZELLER'e teőekkürlerimi sunarım.

Yüksek lisans eęitimine başlamam için beni teővik eden ve bu süreçte desteęini esirgemeyen farklı Őehirlerde olsak ta her koşulda yanımda hissettiğim aileme ve bu anlamlı süreçte bana destek olan deęerli dostum Reyyan KARA'ya ve emeęi geen tüm arkadařlarıma sonsuz teőekkür ederim.

## ÖZET

# MADDELERİ GÜÇLÜKLERİNE GÖRE FARKLI SIRALAMANIN BİREY TEPKİLERİNE ETKİSİNİN DEĞİŞEN MADDE FONKSİYONUyla İNCELENMESİ

Ebru BALTA

Yüksek Lisans Tezi

Eğitim Bilimleri Enstitüsü

Yrd. Doç. Dr. Seçil ÖMÜR SÜNBUİL

Mersin, 2016

74 Sayfa

Bu arařtırmada, test ierisinde yer alan maddelerin glk dzeylerine gre farklı Őekilde sıralandıęı (kolaydan zora ve zordan kolaya) test formlarının verilmesinin, testte yer alan maddelerde DeęiŐen Madde Fonksiyonu (DMF)'na neden olup olmadıęının ve kullanılan DMF belirleme yntemleri arasında uyumun olup olmadıęının belirlenmesi amalanmıŐtır.

alıŐmada c adet Matematik baŐarı testi kullanılmıŐtır. Bu testlerden biri, ęrencilerin bilgi ve becerileri aısından yetenek dzeylerinin belirlendięi test, dięer ikisi ise maddelerin glk dzeylerine gre kolaydan zora ve zordan kolaya doęru sıralandıęı testlerdir. Testlerde yer alan maddelerin DMF ierip iermedięi Lojistik Regresyon(LR) ve Mantel-Haenszel (MH) yntemleriyle belirlenmiŐtir. Bu testlerden elde edilen veriler R-3.2.0 programı kullanılarak analiz edilmiŐtir.

AraŐtırmanın sonucunda, glk dzeylerine gre kolay maddelerin zor maddelerden sonra yer almasının, maddelerin doęru cevaplama olasılıklarında artıŐa neden olduęu grlmŐtr. Ayrıca orta glkteki maddelerin kolay ya da zor maddeden sonra geldięi ve glk dzeyi zor olan maddelerin hem testin baŐında hem de testin sonunda yer aldıęı durumlarda, maddelerin doęru cevaplanma olasılıęının deęiŐtięi ve zor maddelerin test formunun sonunda yer almasının, maddelerin doęru cevaplanma olasılıęındaki farkın artmasına neden olduęu sonucuna ulaŐılmıŐtır.

Böylece bu çalışmada, maddelerin güçlük düzeylerine göre farklı şekilde (kolaydan zora ve zordan kolay) sıralanmasının farklı gruplarda yer alan bireylerin, maddelere doğru cevap verme olasılıklarını etkilediği ve LR ve MH yöntemlerinin, DMF miktarlarındaki büyüklük sıralamalarında benzer belirledikleri DMF'li maddeler bakımından farklı sonuçlar ürettiği sonucuna ulaşılmıştır.

**Anahtar kelimeler:** Değişen Madde Fonksiyonu, Madde Sıralamaları, Mantel-Haenszel, Lojistik Regresyon, Moodle



## ABSTRACT

### AN INVESTIGATION OF EFFECT OF ORDERING ITEMS DIFFERENTLY DEPENDING ON THEIR DIFFICULTY LEVEL UPON INDIVIDUALS' RESPONSE BY USING DIFFERENTIAL ITEM FUNCTIONING

Ebru BALTA

Master thesis

Institute of Education Sciences

Assist. Prof. Dr. Seçil ÖMÜR SÜNBUİL (Advisor)

MERSİN, 2016

74 Pages

This study aims to supply test forms in which items in the test are ordered differently depending on their difficulty level (from easy to difficult or difficult to easy); to determine whether the items in the test form result in Differential Item Functioning (DMF) and to determine whether if there exists a consistency between the methods for detecting DMF.

Three mathematics achievement test has been used in the study. One of these tests intends for measuring students' competence levels in terms of their knowledge and skill. The others are the tests in which the items' difficulty is ordered from easy to difficult and difficult to easy. Methods of Logistic Regression (LR) and Mantel Haenszel (MH) have been taken into consideration to identify whether the items in the tests involve DMF. The data obtained from tests have been statistically analyzed by using R-3.2.0. computing programme.

As a result of the study, it is found out that the if easy items precede the difficult ones, it leads to an increase in the probability of answering the items correctly. In addition, the probability of the items being answered correctly changes in situations in which the items with medium difficulty take place before and after easy and difficult item; and difficult items take place both in the beginning and at the

end of test. Also, it is concluded that the placement of difficult items to the end of test form leads to an increase in the difference in the probability of the items being answered correctly.

Therefore, results of this study can be summarized with the following findings: “ordering the items differently depending on their difficulty level (from easy to difficult and difficult to easy) affects the probability of individuals in various groups answering the items correctly; and also LR and MH methods produce different results with respect to the items with DMF which they have identified similar in terms of magnitude order in the amount of DMF.

**Keywords:** Differential Item Functioning, Item Orderings, Mantel-Haenszel, Logistic Regression, Moodle





## İÇİNDEKİLER

TEŞEKKÜR .....	i
ÖZET.....	ii
ABSTRACT.....	iv
İÇİNDEKİLER .....	v
TABLolar DİZİNİ .....	viii
ŞEKİLLER DİZİNİ .....	x
<b>BÖLÜM I: GİRİŞ .....</b>	<b>1</b>
I.1. Değişen Madde Fonksiyonu (DMF) ve Yanlılık .....	6
I. 1. 1. Tek Biçimli Değişen Madde Fonksiyonu.....	9
I. 1. 2. Tek Biçimli Olmayan Değişen Madde Fonksiyonu .....	10
I.2. Değişen Madde Fonksiyonu Belirleme Yöntemleri .....	11
I. 2. 1. $\chi^2$ Analizi ile DMF Belirleme Yöntemleri.....	12
I.3. İlgili Araştırmalar.....	19
I.3. 1. Yurt İçinde Yapılan Çalışmalar.....	19
I.3. 2. Yurt Dışında Yapılan Çalışmalar .....	21
I.4. Problem .....	36
I.4. 1. Alt Problemler .....	37
I.5. Amaç .....	37
I.6. Önem.....	38
I.7. Sayıtlar .....	38
I.8. Sınırlılıklar.....	39
<b>BÖLÜM II: YÖNTEM .....</b>	<b>40</b>
II.1. Araştırmanın Türü .....	40
II.2. Araştırma Grubu .....	40
II.3. Veri Toplama Araçları.....	40
II.3. 1. Deneme Uygulaması İçin Testlerin Hazırlanması .....	41
II.3. 2. Testlerin Deneme Uygulamaları .....	42
II.3. 3. Asıl Uygulama İçin Testlerin Hazırlanması ( Moodle Sınav Yönetimi) .....	47

II.3. 4.Uygulama.....	48
II.4.Verilerin Analizi.....	50
<b>BÖLÜM III: BULGULAR.....</b>	<b>55</b>
III.1.a. Mantel-Haenszel yöntemiyle yapılan analizde, odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular .....	55
III.1.b. Mantel-Haenszel yöntemiyle yapılan analizde, odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular .....	58
III.1.c. Mantel-Haenszel yöntemiyle yapılan analizlerde, değişen madde fonksiyonu gösteren maddelerin farklılık gösterip göstermediğine ilişkin bulgular .....	60
III.2.a. Lojistik Regresyon yöntemiyle yapılan analizde, odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular	61
III.2.b. Lojistik Regresyon yöntemiyle yapılan analizde, odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular	64
III.2.c. Lojistik Regresyon yöntemiyle yapılan analizlerde, değişen madde fonksiyonu gösteren maddelerin farklılık gösterip göstermediğine ilişkin bulgular .....	67
III.3.a. Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle yapılan analizlerde, odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelerin uyum gösterip göstermediğine ilişkin bulgular .....	68
III.3.b. Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle yapılan analizlerde, odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelerin uyum gösterip göstermediğine ilişkin bulgular .....	69
<b>BÖLÜM IV. TARTIŞMA VE YORUM .....</b>	<b>70</b>
IV.1. Mantel-Haenszel yöntemiyle yapılan analiz sonuçlarının değerlendirmesi.....	70
IV.2. Lojistik Regresyon yöntemiyle yapılan analiz sonuçlarının değerlendirmesi .....	72
IV.3. Mantel-Haenszel ve Lojistik Regresyon yöntemiyle yapılan analiz sonuçlarının karşılaştırması	72
<b>ÖNERİLER .....</b>	<b>74</b>
<b>KAYNAKÇA .....</b>	<b>75</b>
<b>EKLER.....</b>	<b>85</b>
EK: 1: Uzman Görüşleri Örnek Madde Bilgi Formu .....	85
Ek 2: Uzmanlar Arası Uyum Veri Toplama Formu .....	87

Ek 3: 40 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testinin (Test 1) Madde Analizleri .....	88
Ek 4: 20 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testi (Test 1) .....	89
Ek 5: 40 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testlerinin (Test 2-Test 3) Madde Analizleri .....	92
Ek 6: 20 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testi (Test 2) .....	93
Ek 7: 20 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testi (Test 3) .....	95
Ek 8:Yönerge Sayfasının Ekran Görüntüsü .....	97
Ek 9: Odak Grubunun Test 2' yi, Referans Grubunun Test 3' ü Alması Durumuna Göre MH Analiz Sonuçları.....	98
Ek 10: Odak Grubunun Test 3' ü, Referans Grubunun Test 2'yi Alması Durumuna Göre MH Analiz Sonuçları.....	100
Ek 11: Odak Grubunun Test 2' yi, Referans Grubunun Test 3' ü Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları .....	102
Ek 12: Odak Grubunun Test 3'ü, Referans Grubunun Test 2' yi Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları .....	104
Ek 13: Odak Grubunun Test 2'yi, Referans Grubunun Test 3' ü Alması Durumuna Göre Tek Biçimli Olmayan DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları.....	106
Ek 14: Odak Grubunun Test 3'ü, Referans Grubunun Test 2' yi Alması Durumuna Göre Tek Biçimli Olmayan DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları.....	108
<b>ÖZGEÇMİŞ.....</b>	<b>110</b>

## TABLolar DİZİNİ

Tablo 1: Mantel-Haenszel Yönteminde Kullanılan Veri Düzeni.....	13
Tablo 2: j Yetenek Seviyesindeki Bireyler İçin i Maddesinin Olasılık Tablosu .....	14
Tablo 3: Mantel Haenszel DMF İstatistiğinin Yorumlanması.....	16
Tablo 4: Lojistik Regresyon Tekniğinde Değişen Madde Fonksiyonu Düzeyi İçin $R^2$ Ölçütleri .....	19
Tablo 5: 20 Maddelik Test 1'in Madde İstatistikleri .....	43
Tablo 6: 20 Maddelik Test 1'in Test İstatistikleri.....	44
Tablo 7: 20 Maddelik Test 2 ve Test 3'ün Madde İstatistikleri.....	45
Tablo 8: 20 Maddelik Test 2 ve Test 3'ün Test İstatistikleri .....	46
Tablo 9: Testlerin Uygulanma Şekline Ait Desen .....	49
Tablo 10: Testlere İlişkin Tek Boyutluluk Analiz Sonuçları.....	50
Tablo 11: Odak ve Referans Grubuna Uygulanan Test 1'e İlişkin Betimsel İstatistikler.....	51
Tablo 12: Test 3'ün Uygulandığı Odak Ve Referans Gruplarına Yapılan Mann Whitney U Testi Sonuçları .....	53
Tablo 13: Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Mantel-Haenszel Yöntemine İlişkin Analiz Sonuçları .....	56
Tablo 14: Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Mantel-Haenszel Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar .....	57
Tablo 15: Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Mantel-Haenszel Yöntemine İlişkin Analiz Sonuçları .....	58
Tablo 16: Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Mantel-Haenszel Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar .....	59
Tablo 17: Mantel- Haenszel Yöntemleriyle Yapılan Analiz Sonucunda DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar.....	60
Tablo 18: Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Lojistik Regresyon Analizine İlişkin Sonuçlar.....	62
Tablo 19: Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Lojistik Regresyon Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar.....	63
Tablo 20: Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Lojistik Regresyon Analizine İlişkin Sonuçlar.....	65
Tablo 21: Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Lojistik Regresyon Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar.....	66

Tablo 22: Odak Grubunun Test 2' yi, Referans Grubunun Test 3'ü Alması Durumuna Göre LR ve MH Yöntemleriyle DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar ..... 68

Tablo 23: Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre LR ve MH Yöntemleriyle DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar ..... 69



## ŞEKİLLER DİZİNİ

Şekil 1. Tek Biçimli DMF (Osterlind ve Everson, 2009: 12).....	9
Şekil 2. Tek Biçimli Olmayan DMF (Osterlind ve Everson, 2009: 12) .....	10
Şekil 3. Test 2 'de yer alan 17. maddenin ekran görüntüsü .....	48
Şekil 4. Odak Grubunda Yer Alan Öğrencilerin Matematik Puanı Dağılım Grafiği.....	52
Şekil 5. Referans Grubunda Yer Alan Öğrencilerin Matematik Puanı Dağılım Grafiği .....	52
Şekil 6. Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Mantel-Haenszel Analizine İlişkin Madde Dağılımları.....	57
Şekil 7. Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Mantel-Haenszel Analizine İlişkin Madde Dağılımları.....	59
Şekil 8. Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı .....	63
Şekil 9. Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Tek Biçimli Olmayan DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı .....	64
Şekil 10. Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı .....	66
Şekil 11. Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Tek Biçimli Olmayan DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı. ....	67

## BÖLÜM I: GİRİŞ

Günümüzde eğitim girdi, süreç, çıktı ve kontrol öğeleri olan insan davranışlarını geliştiren ve değiştiren bir sistem olarak görülmektedir. Eğitim sisteminin kontrolü, öğelerin iyi işleyip işlemediği, varsa işlemeyen yönlerinin ortaya konulması ve sistemin onarılmasının sağlanması ölçme ve değerlendirme ile gerçekleştirilmektedir. Değerlendirme, bir karar verme, yargıda bulunma sürecidir ve sistemin sağlıklı işlemesi açısından oldukça önemlidir. Öğretim sürecinde ise, eğitim programlarının sağlam olup olmadığını anlama, öğretimde başvurulan teknik ve yöntemlerin etkililik derecesini saptama, öğrencilerin öğrenme düzeylerini ve güçlüklerini belirleme, öğrenci başarısını saptama gibi amaçlarla ölçme ve değerlendirme çalışmaları yapılır (Baykul, 2010). Eğitimde ölçme konusu olan ve doğrudan gözlenemeyen, yetenek, başarı, tutum, zeka, ilgi vb gibi değişkenler, operasyonel olarak tanımlanmasında güçlük çekilen psikolojik değişkenlerdir. Eğitimde ölçme ve değerlendirme çalışmalarının çoğunda öğrencilerin bu psikolojik özelliklerinin ortaya çıkarılması amaçlanır. Bu psikolojik özellikler ve yapılar doğrudan gözlenemezler ancak dolaylı olarak bir ölçme aracıyla gözlenebilirler. Ölçme araçları; bireyleri tanımak, bireyler arasındaki farklılıkları ortaya çıkarmak, bireyleri yetenek düzeylerine göre sıralamak ve ayırmak gibi çeşitli amaçlar için geliştirilebilmektedir. Ölçme araçlarından elde edilen sonuçlara göre bireyler hakkında çeşitli kararlar alınmakta ve bu alınan kararlar kişisel, sosyal ve politik öneme sahip olabilmektedir. Bireyler hakkında bu kadar ciddi kararların verildiği ölçme ve değerlendirme uygulamalarının da sağlıklı olması gerekir. Bunun için öncelikle ölçme araçlarının nitelikli olması, bir başka ifadeyle ölçme araçlarının yeterli düzeyde güvenilir ve geçerli olduklarına dair kanıtların toplanmış olması gerekmektedir. Baykul'a (2010) göre, "ölçülen nitelik hakkında doğru ölçme sonuçları vermeyen, amaca hizmet etmeyen ölçme araçlarından elde edilen sonuçlara bakılarak alınan kararların yerinde ve isabetli olma olasılığı oldukça düşüktür".

Nitelikli bir ölçme aracında bulunması gereken özelliklerden biri güvenilirliktir. Güvenirlik, bir nitelik ile ilgili olarak aynı bireyler üzerinden yapılan ölçmelerin benzer şartlar oluşturulduğunda tekrarlanabilir olmasıdır (Crocker ve Algina, 1986). Bir testin güvenilir olabilmesi için, farklı zaman ve ölçümlerde aynı sonucun elde edilmesi gerekmektedir. Bir başka ifadeyle elde edilen sonuçlar kararlı ve tutarlı olmalıdır. Ölçme yaparken ölçülmek istenilen özellik hakkında doğru bilgi elde edebilmek ve elde edilen sonuçlara dayanarak doğru kararlar verebilmek için ölçme sonuçlarına karışan hatanın da az olması gerekmektedir. Ancak her zaman ölçme sonuçlarına bir miktar hata karışmaktadır. Ölçmeye karışan çeşitli hatalar nedeniyle gözlenen bir özelliğin gerçek değeri doğrudan elde edilemez. Klasik Test Kuramı'na (KTK) göre ölçülen özelliğin gerçek değeri, ölçülen özelliğe ilişkin sonsuz sayıda ölçmelerden elde edilen puanların ortalaması, yani gerçek puan şeklinde tanımlanmaktadır (Crocker ve Algina, 1986). Gerçek değer, gözlenen değer yardımıyla kestirilmeye çalışılır. Hata, ölçülmek istenen özelliğin gözlenen değeri ile o özelliğin gerçek değeri arasındaki farktır (Crocker ve Algina, 1986). Ölçmelere karışan hatalar, kaynağı, miktarı ve yönüne göre sabit, sistematik ve seçkisiz (random) hatalar

olarak sınıflanmaktadır. Burada belirtilen hata seçkisiz (random) hata olup, bir bireyin ölçülmek istenilen özelliğine ilişkin gözlenen değeri ile gerçek değeri arasındaki fark olarak da tanımlanabilir. Ölçme sonuçlarına karışan seçkisiz (random) hatalar, ölçme sonuçlarına gelişigüzel karışan, kaynağı, miktarı ve yönü kesin olarak bilinmeyen hatalardır. Bu tür hatalar, ölçmenin yapıldığı ortam, ölçmenin yapıldığı grup, ölçmeyi yapan kişi, ölçme aracı gibi birçok hata kaynağından kaynaklanan bir hata türüdür (Baykul, 2010; Turgut, 1995). Sabit hatalar, bir ölçmeden diğerine miktarı değişmeyen, miktarı, yönü ve kaynağı belli olan hatalardır. Sistemik hatalar ise ölçülen yapı ya da yeterlikle ilişkili olmayan gruplar arasındaki tutarlı farklılıklar olarak tanımlanmaktadır (Osterlind ve Everson, 2009). Sistemik hatalar, ölçülen büyüklüğe, ölçmeciye veya ölçme koşullarına ve belli bir duruma bağlı olarak ölçmeden ölçmeye miktarı değişen hatalardır. Sabit hata da olduğu gibi sistemik hatada da ölçme sonuçlarına karışan hataların miktarı, yönü ve kaynağı belli olduğundan ölçme sonuçlarına karışmasının engellenmesi ya da düzeltilmesi mümkün olabilmektedir. Eğitim alanında yapılan ölçmelerde ölçme sonuçlarına dayanarak bir karara varılması amaçlanır. Kararların isabetlilik derecesi yapılan ölçmenin doğruluğuna bağlı olduğundan, ölçümlere karışan bu hataların tanımlanması ve hata kaynaklarının belirlenerek giderilmesi gerekmektedir (Turgut, 1995).

Bir ölçme aracında bulunması gereken diğer yapısal özellik ise geçerliktir. Bir ölçme aracının ölçmek istediği şeyi ölçüp ölçmediği hakkında karar verebilmek için birçok kanıtı ihtiyacı duyulmaktadır. Geçerlik; testi kullanan kişinin, test puanlarından çıkarılabilecek yordamaları desteklemek için topladığı kanıtlardır (Cronbach, 1990). Geçerlik, bir ölçme aracının ölçmek istediği değişkene ait özellikleri ne derecede ölçtüğünü ifade eder (Anastasi'den akt. Baykul, 2010). Diğer bir ifade ile geçerlik, bir testin sadece o testle ölçülmek istenilen değişkeni ölçmesi, başka değişkenlerle karıştırmamasıdır. Bu anlamda geçerlik tek bir tanımdan ziyade, testin kullanım amacına hizmet ettiğini gösteren kanıtların toplanması olarak düşünülebilir. Bu kanıtların toplanması, testin neyi ölçtüğüne ilişkin anlayış geliştirmek ve bireyler ya da nesnelere ilişkin alınacak kararların doğruluğuna destek sağlayacak bilgi elde etmek adına önemli görülmektedir (Atılgan, 2014). Geçerlik, güvenilirliği de kapsayan daha geniş bir kavram olup, ölçülmek istenilen değişkenin yapısı, ölçme aracının ilgili konu alanını kapsama düzeyi, kullanım amacı, yanlılık gibi pek çok değişkenden etkilenmektedir (Crocker ve Algina, 1986; Magnusson, 1967; Aiken, 2000). Seçkisiz hata kaynakları ölçme sonuçlarının güvenilirliğini etkileyen bir sorun olarak ortaya çıkarken; sistemik hata kaynakları geçerliği etkileyen bir sorun olarak ortaya çıkmaktadır. Geçerliği etkileyen değişkenler incelendiğinde ise, yanlılığın bu değişkenler arasında önemli bir yer tuttuğu söylenebilir.

Eğitim alanında yapılan ölçme ve değerlendirme çalışmalarında ölçmelere karışan sistemik hatalar genel olarak “yanlılık” (bias) kapsamında ele alınmaktadır. Bir ölçme aracı geliştirildiğinde “ölçme aracı uygulanan bireylerde aynı özelliği ölçer” kabulü ile hazırlanır. Dolayısıyla ölçme aracından



alınan puanların farklı alt gruplara bağımlı olmadığı kabul edilir. Ölçme sonuçlarının geçerliğine ilişkin kanıtlar toplanırken, ölçme aracını cevaplayan grup üyeleri için değişmez olması beklenir. Bir ölçme aracı geliştirilirken aracın farklı gruplar için geçerli olabileceği bir çerçeve sunulmalıdır. Bir başka ifadeyle, testten elde edilen puanların farklı alt gruplara yanlı davranmaması gerekir. Eğer, ölçme aracının ölçtüğü özellikler farklı alt gruplar için geçerli değilse ve yanlılık gösteriyorsa, bu ölçme aracından elde edilen verilerle yapılan karşılaştırmaların, yorumların, alınan kararların geçerliği sorgulanabilir. Çünkü yapılan ölçmeye hata karışmıştır. Hatalı yapılan ölçme sonuçlarına bağlı olarak yapılan değerlendirmeler de hatalı olur. Bu nedenlerden ötürü, ölçme aracının testi cevaplayan çeşitli gruplara haksız bir şekilde avantaj veya dezavantaj sağlamadığından emin olunması gerekir (Crocker ve Algina, 1986). Ölçme aracının aynı yetenek düzeyindeki gruplardan birine avantaj sağlama durumu, yetenek düzeyleri aynı olan bireylerin test puanlarının, ölçülmesi amaçlanan yapıyla ilişkili olmayan bir veya daha fazla değişkenlik kaynağına bağlı olarak farklılaşmasıdır. Bu tür bir farklılaşma, Değişen Madde Fonksiyonu (DMF) ve yanlılık analizleri kapsamında yürütülen süreçlerle belirlenmeye çalışılır.

Eğitimde ve psikolojide çeşitli ölçme araçlarından biri olan testlerden yararlanılmaktadır. Öğrencilerin belirli bir ders veya konu alanı için belirlenen bilgi ve becerilere sahip olma derecelerini belirlemek amacıyla kullanılan ölçme araçlarına test denilmektedir (Airasian, 1997). Bazı testler, bireylere kazandırıldığı düşünülen bilgi ve becerilerin ne kadar kazanıldığını belirlemek; bazıları, bireylerin şimdiki durumlarından gelecekteki durumlarını yordamak; bazıları ise bireylerin psikolojik yapıları hakkında bilgi toplamak amacıyla geliştirilir (Erkuş, 2003). Psikolojik testler arasında en yaygın olarak kullanılan testlerden biri olan başarı testleri ise, kişinin eğitim süreci içinde veya belirli çevre koşulları altında ne kadar öğrendiğini ölçmek amacıyla düzenlenir (Tan, 2013). Eğitim ve psikoloji alanında en çok çoktan seçmeli testlerden yararlanılmaktadır. Bunun nedenleri arasında çoktan seçmeli test maddelerinin puanlanmasının kolay ve nesnel olması, ayrıca sınav süresi içinde çok sayıda maddenin sorulabilmesi gösterilebilir.

Çoktan seçmeli testlerde testin planlanması hazırlanması ve uygulanması aşamalarında bazı hususlara dikkat edilmesi gerekmektedir. Crocker ve Algina'ya (1986) göre test geliştirme süreci, test puanlarının kullanılacağı amacın belirlenmesi, davranışın temsil ettiği yapının açıklanması, teste uygun madde türünün seçilmesi, özelliğe uygun madde türünün seçilmesi, test formlarının hazırlanması, testin uygulanması, madde analizi yapılarak maddelerin seçilmesi, seçilen maddelerden oluşturulan nihai testin psikometrik bazı özelliklerinin incelenmesi gibi adımlardan oluşmaktadır.

Test geliştirme sürecinin ilk adımı testin (test puanlarının) hangi amaçla kullanılacağı belirlenmesidir. Bu amaçlar; seçme yerleştirme, öğrenciler hakkında eğitim kararlarının verilmesi, öğrencilerin öğrenme güçlüklerinin belirlenmesi ve öğrenci başarısının saptanması olabilmektedir.

Testin içeriği testin kullanılma amacına göre belirlendiğinden her amaç için testte yer alan maddeler birbirinden farklı olacaktır (Crocker ve Algina, 1986).

Test geliştirme sürecinin ikinci aşaması ise, davranışın ve davranışın temsil ettiği psikolojik yapıların tanımlanmasıdır. Bu aşamada yapıyı ya da alanı temsil eden davranışlar belirlenip belirtke tablosu oluşturulmaktadır. Test geliştirme sürecinin üçüncü aşamasında ise teste uygun madde türünün seçilmesi yer almaktadır. Belirtke tablosunda yer alan bilişsel alanın her bir basamağındaki davranışlar için test maddesi yazılırken o davranışı ölçebilecek en uygun madde tipinin seçilmesi gerekmektedir (Atılgan, 2014). Başarı, yetenek gibi performansı ölçmeye yönelik testlerde genellikle doğru-yanlış, eşleştirmeli ve çoktan seçmeli madde türleri kullanılmaktadır (Crocker ve Algina, 1986). Test geliştirmenin beşinci aşaması ise denemelik olarak hazırlanan maddeler gözden geçirildikten sonra denemelik test formunun hazırlanmasıdır. Denemelik test formunun hazırlanması; maddelerin test formu içerisine dağıtılması, test yönergesinin yazılması ve maddelerin yazılması gibi adımları içermektedir. Maddelerin test formu içerisine dağıtılmasında aynı davranışı yoklayan maddelerin art arda gelmemesine, testin başına kolay maddelerin konulmasına ve maddelerin konulara göre gruplanıp bu doğrultuda maddelerin test içerisine tamamen seçkisiz ya da sistematik dağıtılmasına özen gösterilmelidir. Maddelerin test içerisinde hangi sırayla alacağına belirlenmesinden sonra testin başına konulacak test hakkında bilgi veren, cevaplamanın nasıl yapılacağını açıklayan yönerge yazılır. Test geliştirmenin altıncı aşamasında ise testin deneme uygulamasının yapılması yer almaktadır. Madde analizlerinin yapılabilmesi ve madde seçimi için denemelik test maddelerinden oluşturulan denemelik test formunun bir gruba uygulanması gereklidir. Madde analiziyle elde edilen madde istatistikleri ve bunlara dayalı olarak kestirilecek test istatistiklerinin asıl uygulama grubundan elde edilecek istatistiklere yakın olması deneme grubunun asıl grubu temsil gücüne bağlı olduğundan deneme grubunun nihai testin uygulanacağı asıl grubu yansıtacak nitelikte olmasına dikkat edilmelidir. Deneme uygulaması sonucunda elde edilen veriler ile madde ve test istatistikleri hesaplanarak teste alınacak maddeler belirlenir. Madde seçimi tamamlandıktan sonra test geliştirmenin son aşamasında nihai testin ortalaması, standart sapması, güvenilirliği ve geçerliği gibi bazı psikometrik özellikleri ortaya konur (Baykul, 2010; Atılgan, 2014).

Deneme uygulamasının son aşamasında yapılan madde analiziyle elde edilen madde istatistikleri, test geliştirme ve soru bankası oluşturma işlemlerinde, maddelerin test ya da soru bankasına alınması, atılması veya düzeltilip yeniden kullanılması gibi kararların alınmasında kullanılmaktadır. Bu nedenle test geliştirme sürecinin beşinci aşamasında yer alan, test formlarının düzenlenmesi aşaması, ölçme araçlarının bazı psikometrik özelliklerinin belirlenmesi açısından önem teşkil etmektedir. Test formları düzenlenirken testin güvenilirliğini ve geçerliğini arttırmak için çeşitli önlemler alınmaya

çalışılır. Burada dikkat edilmesi gereken noktalardan bir tanesi de maddelerin test formu içerisindeki yerinin ve sırasının belirlenmesidir.

Madde sırası, maddenin test içindeki yerinin değiştirilmesi anlamına gelir. Özellikle çoktan seçmeli testlerin uygulandığı sınavlarda, güvenilirliği artırmak adına öğrencilerin kopya çekmelerini engellemek amacıyla, testlerde yer alan maddelerin yerleri değiştirilerek farklı test formları hazırlanır. Bu test formlarında yer alan maddelerin sıralanması, maddelerin güçlük düzeyine (kolaydan zora, zordan kolayla ya da seçkisiz), konuların içeriğine ve bilişsel güçlük düzeyine (bilgi, kavrama, uygulama, analiz, sentez, değerlendirme) göre yapılabilmektedir. Leary ve Dorans (1982), madde sıralamaları ile ilgili yaptıkları literatür taramasında 21 çalışma sonucunda maddelerin test içerisinde sıralanmasının; maddelerin sıralanmasında 1) tek faktörün incelendiği, 2) etkileşim etkilerinin incelendiği, 3) madde sıralamasının test türüne bağlı olduğu ve 4) Madde Tepki Kuramı (MTK) kapsamında ele alındığı çalışmalar olmak üzere dört ana başlık altında toplamışlardır. Leary ve Dorans (1985), yaptıkları sınıflandırmanın dışında yeni çalışmalar olduğunu, buna göre maddelerin test içerisinde, *konuya göre sıralandığı, madde güçlük düzeylerine göre (kolaydan zora, zordan kolayla ve seçkisiz olarak) sıralandığı, belirli bir sistematığe göre sıralanmadığı, hem madde güçlük düzeylerine göre hem de bilişsel güçlük düzeyine göre sıralandığı ve madde sıralamalarıyla ilgili literatür taramaları ve meta-analitik çalışmalar* olmak üzere beş başlık altında toplamışlardır.

Test maddelerinin konuya göre sıralandığı çalışmalar, test maddelerini konuların sınıf içinde işleniş sırasına göre sıralayan çalışmalardır. Bu çalışmalarda maddeler öncelikle, konuların derste işleniş sırasına göre gruplanmakta sonra bu gruplar içindeki maddeler madde güçlük düzeylerine göre kolaydan zora, zordan kolayla ve seçkisiz olarak sıralanmaktadır.

Test maddelerinin madde güçlük düzeylerine göre kolaydan zora, zordan kolayla ve seçkisiz olarak sıralandığı çalışmalar ise, testteki tüm maddelerin madde güçlük düzeylerine göre sıralandığı ve testte kullanılan madde türüne göre maddelerin önce gruplanıp ardından her gruptaki maddelerin madde güçlük düzeylerine göre kolaydan zora, zordan kolayla ve seçkisiz olarak sıralandığı çalışmalardan oluşmaktadır.

Test maddelerinin belirli bir sistematığe göre sıralanmadığı çalışmalar ise; test maddelerinin testin bölümleri içinde seçkisiz olarak sıralandığı çalışmalardan oluşmaktadır. Testteki maddelerin hem madde güçlük düzeylerine hem de bilişsel güçlük düzeyine göre sıralandığı çalışmalarda, maddeler test içinde hem madde güçlük düzeylerine göre hem de maddenin bilişsel güçlük düzeyi dikkate alınarak sıralanmaktadır.

Klasik Test Teorisi ve uygulamalarında, bireylerin maddelere verdikleri tepkilerin maddelerin test içerisindeki yerlerinden bağımsız olduğu varsayılmaktadır (Lord ve Novick, 1968: 252). Yani test

maddelerinin madde içeriği değiştirilmeden test formlarında farklı sıralarda olmasının öğrencilerin performansını etkilemeyeceği kabul edilmektedir. Fakat maddelerin farklı şekillerde sıralandığı alternatif formları cevaplayan bireylerin maddelere verdikleri cevaplar dolayısıyla testten aldıkları puanlar değişebilmektedir. Bireylerin maddelere verdiği tepkilerin, maddenin, test içerisindeki sırasından etkilenmesi sıra etkisi (position effect) olarak tanımlanmaktadır (Kingston ve Dorans, 1984; Yen, 1980). Sıra etkisi bireyin test performansını çeşitli şekillerde etkilemektedir: Madde güçlüğü açısından kolay ve zor maddelerin testin başında ya da sonunda yer almasına bağlı olarak öğrencilerin test boyunca motivasyonları artıp ya da azalmaktadır (Towle ve Merrill, 1975). Ayrıca test maddelerinin farklı sıralarda alınması öğrencilerin kaygı düzeylerinin artmasına böylece testi olduğundan daha zor ya da kolay algılamasına neden olarak testin güvenilirliğinin düşmesine neden olabilmektedir (Leary ve Dorans, 1982). Bunların dışında, maddelerin farklı şekillerde sıralandığı farklı test formlarını cevaplayan bireylerin maddelere verdikleri cevaplar dolayısıyla testten aldıkları puanlar değişebilmekte ve bu durum da madde ve test istatistiklerini etkileyebilmektedir (Flaughner, Melton ve Meyers, 1968; Hambleton ve Traub, 1974; Barciovski ve Olsen, 1974; Kleinke, 1980).

Bu bilgiler ışığında testlerde maddelerin çeşitli özelliklere göre sıralanmasının bireylerin maddelere verdikleri tepkilerde değişikliğe neden olup olmadığının, eğer neden oluyorsa bunun sebeplerinin araştırılması önemli görülmektedir. Bu çalışmada, maddelerin madde güçlük düzeylerine göre test içerisinde farklı sıralarda yerleştirildiği farklı test formlarının verilmesinin, bireylerin maddelere verdikleri tepkileri etkileyip etkilemediği Değişen Madde Fonksiyonu'yla (DMF) incelenecektir. Aşağıdaki kısımda Değişen Madde Fonksiyonu konusundan ve çalışma kapsamında kullanılan DMF belirleme yöntemlerinden kısaca bahsedilecektir.

### **I.1. Değişen Madde Fonksiyonu (DMF) ve Yanlılık**

Eğitimde kullanılan ölçme araçlarının, test maddelerinin dolayısıyla ölçümlerin geçerliği, ölçme alanının temel sorunlarından biridir. Geçerliği olumsuz yönde etkileyen faktörlerden biri de “yanlı” maddelerdir (biased items). Yanlılık, ölçme sonuçlarının sistematik hata içermesi anlamına gelmektedir (Camilli ve Shephard, 1994). Yanlılık; testlerin önemli psikometrik özelliklerinden olan geçerliği düşüren bir durumdur. Bu nedenle yanlılık çalışmaları geçerlik çalışmaları içerisinde yer alır. Yanlılık çalışmalarında üzerinde durulan hata türü sistematik hatadır ve ölçme sonuçlarına karışan sistematik hata kaynaklarının belirlenmesi geçerlik açısından oldukça önemlidir.

Test maddelerinin, aynı yetenek düzeyinde olan fakat farklı alt gruplarda yer alan bireylerin cevapları açısından farklı özellikler göstermemesi; grupların ölçülen özellik bakımından doğru tanımlanması gerekmektedir. Madde yanlılığı; aynı yetenek düzeyinde olan fakat farklı gruplarda yer

alan bireylerin bir maddeyi doğru cevaplama olasılıklarının aynı olmaması durumu olarak tanımlanmaktadır (Osterlind, 1983; Hambleton ve Rogers, 1989; Zumbo, 1999). Bu tanıma göre, madde yanlılığı analizleri her bir test maddesinin aynı evrenden elde edilen farklı alt gruplarda benzer şekilde işleyip işlemediği sorusu üzerine odaklanmaktadır.

Yanlılık kavramının doğru bir biçimde anlaşılabilmesi için, istatistiksel yanlılık (statistical bias), test yanlılığı (test bias) ters etki (adverse impact), madde etkisi (item impact), değişen madde fonksiyonu (DMF) ve madde yanlılığı (item bias) kavramlarının ayırt edilmesi gerekmektedir. İstatistiksel yanlılık, bir veya daha çok sayıdaki parametreye dair tahminin, olması gerektiğinden çok düşük veya yüksek olması durumunu belirtir (Camilli ve Shepard, 1994). Genel bir ifadeyle, istatistiksel yanlılık, eşit olması gereken iki parametre arasındaki sistematik farktır. Test yanlılığı, bir testin geçerli olmama hali veya belli bir grubun üyelerine dair ölçmelerdeki sistematik hata olarak tanımlanır (Camilli ve Shepard, 1994:8). Test yanlılığı sıklıkla ırksal, etnik ve cinsiyet farklılıklarının ortaya konulduğu çalışmalarda bir problem olarak görülür ve farklı çalışma alanlarında, sosyoekonomik düzey, yaş, din, şehirde ya da kırsal bölgede yaşama gibi kimi grup farklılıklarının incelendiği çalışmalarda da önem teşkil etmektedir (Camilli ve Shepard, 1994:8). Bir testin yanlılığı genel olarak madde bazında yapılan analizlerle belirlenir. Testi oluşturan maddelerin yanlılık içerip içermeme durumuna göre testin de yanlı olup olmadığına karar verilir. Test puanlarının geçerliğinin düşmesinin veya test yanlılığının artmasının nedeni, test maddelerine bir grubun diğerinden daha fazla doğru cevap vermesi, grubun biri lehine adaletsizce işleyen yanlı maddeler olması şeklinde tanımlanmaktadır (Holland ve Wainer, 1993). Testteki yanlı maddeler belirlenebilirse, testten çıkarılarak test yansız hale getirilebilmektedir. Bu nedenle, yanlılık araştırmaları, madde bazında yapılan yanlılık çalışmalarına dayalı olarak yürütülür (Atılğan, 2014). Etki, iki grubun bir test veya maddeye ait ortalama puanları arasındaki farkı tanımlamak için kullanılan bir kavramdır. Ters etki, gruplar arasında test performans farkının orantısız bir seçme veya başka test-yanlılığı kararlarına yol açma durumlarını ifade eden kavramdır. Ters etki ve gruplar arasında fark her zaman test yanlılığına işaret etmez. Sadece gruplardan birinin diğerinden düşük ortalamaya sahip olması, testte yanlılık olduğuna dair karar verebilmek için yeterli değildir. Bu nedenden, testlerdeki yanlılığı bulma amacımız, gruplar arasındaki farklılığın gerçek yetenek farklarından kaynaklanıp kaynaklanmadığını bulmaya yöneliktir (Camilli ve Shepard, 1994:7-8). Madde etkisi, farklı alt gruplarda bulunan bireylerin bir maddeyi doğru cevaplama olasılıklarının, söz konusu madde ile ölçülmek istenen özellik boyutunda farklılaştığında oluşur (Zumbo, 1999). Madde etkisinde gruplar arasındaki fark, grupların ölçülen özelliğe yönelik yetenek düzeylerinin farklılaşmasından kaynaklanır. Yani madde etkisi; farklı alt gruplardan aynı testi alanların bir maddeye doğru cevap verme olasılıklarının, ölçülen özellikteki gerçek farklılıktan kaynaklandığını ifade eder. Alt gruplardan herhangi birinin bir madde üzerindeki performansları farklılık gösteriyorsa, bunun nedeni gruplar arasındaki gerçek yetenek farklılıkları olabileceği gibi madde yanlılığından da kaynaklı olabilir. Yani

DMF'nin ortaya çıkmasının iki nedeni olabilir. Bunlar alt gruplar arasındaki gerçek farklılık (true difference) ve madde yanlılığıdır (item bias) (Camilli ve Shepard,1994). Bunun kaynağını belirlemek için ilk olarak, grup farklılıklarını ortaya çıkaran çalışmalar yapılır; daha sonraki aşamada ise bu farklılıkların gerçekte grupların yetenek düzeyleri arasındaki farktan mı yoksa ölçme sürecindeki ölçülmesi amaçlanan yapıyla ilişkili olmayan bir bozulmadan mı kaynaklandığı belirlenmeye çalışılır. Gruplar arasında farklılıklar bulunması ölçme yanlılığına işaret etmeyebilir. Madde yanlılığı, maddelere verilen tepkilerdeki farklılığın hem istatistiksel yönden araştırılmasını hem de farklılığın kaynağının ne olduğunun belirlenmesi süreçlerini içerir. İstatistiksel olarak maddelere verilen tepkilerin farklılaşmasına “Değişen Madde Fonksiyonu (DMF)” adı verilir (Camilli ve Shepard,1994). DMF, madde ile ölçülmek istenilen psikolojik yapının her bir yetenek düzeyinde, maddeyi doğru cevaplama olasılıklarının, alt gruplara göre farklılık gösterdiğini belirleyen bir fonksiyondur (Zumbo, 1999). Değişen madde fonksiyonu, madde yanlılığının bir indeksidir. Yanlı bir madde DMF gösterir; ancak bir maddenin DMF göstermesi o maddenin yanlı olduğunu belirtmek için yeterli değildir. DMF istatistiksel bir yöntem iken, madde yanlılığı kavramsaldır ve yoruma açıktır (Camilli ve Shepard, 1994). Yanlılık çalışmaları genel olarak DMF sonuçlarına dayalı olarak değerlendirilmektedir (Zumbo, 1999:6).

Değişen Madde Fonksiyonu (DMF) analizleri, bu çerçevede, gözlenen grup farklarının bireylerin yetenek düzeylerindeki farktan mı yoksa madde yanlılığından mı kaynaklandığını belirleme yönündeki inceleme sürecini ifade eder. Madde yanlılığı araştırılması süreci, hem istatistiksel analizlerin yapılmasını (DMF analizleri) hem de DMF gösterdiği belirlenen maddelerin olası DMF kaynaklarının tespit edilerek, ölçülen yetenekteki gerçek farklılıktan mı yoksa istenmeyen varyans kaynaklarından mı geldiğinin belirlenmesi aşamalarını içermektedir. DMF gösteren bir maddenin yanlı olup olmadığına yapı ve kapsam geçerliği bağlamında incelenmesinin ardından karar verilmektedir (Holland ve Wainer, 1993).

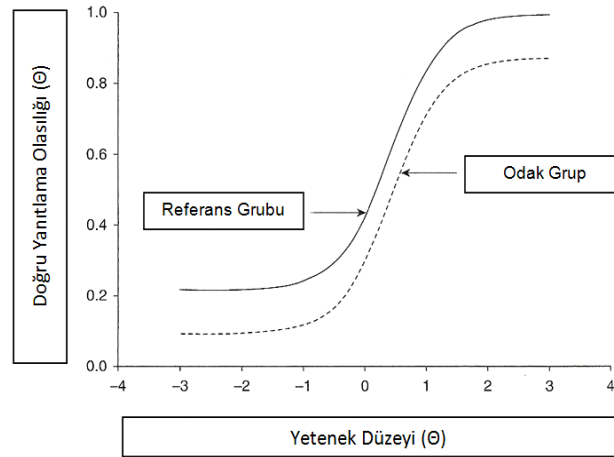
DMF belirleme yöntemlerinde genel olarak, “odak” ve “referans” olarak alınan iki gruptaki bireyler yetenek seviyelerine göre eşlenirler. Dezavantajlı olduğu düşünülen ve azınlığı (minority) teşkil eden grup odak (focal) grup olarak adlandırılırken, bu grubun performansı ile karşılaştırılan çoğunluk (majority) grubu ise referans (reference) grubu olarak adlandırılır (Camilli ve Shepard, 1994; McNamara ve Roeber, 2006). Referans ve odak gruplarının uygun şekilde belirlenmesinin ardından, söz konusu iki grubun bir değişkene göre, gösterdikleri performans anlamında eşitlenmesi gerekmektedir; bu süreç yetenek düzeylerinin eşitlenmesi anlamına gelmektedir. Yetenek düzeylerini eşitlemek için *içsel eşitleme (internal matching)* ve *dışsal eşitleme(external matching)* olmak üzere iki farklı yol bulunmaktadır. İçsel eşitlemede eşitleme değişkeni, incelenen maddeleri içeren testten, dışsal eşitlemede ise, incelenen maddeleri içeren test dışında farklı bir testten alınan toplam puandır (Karami ve Nodoushan, 2011). Yetenek düzeyleri içsel ya da dışsal eşitleme türüne göre eşitlendikten sonra maddeler üzerindeki grup farklılıklarını belirlemek için istatistiksel süreçler uygulanır.

Aynı yetenek düzeyine sahip fakat farklı alt gruplarda yer alan bireylerin ilgili maddeyi doğru cevaplama olasılıklarındaki farklılık olarak ortaya çıkan DMF; Tek Biçimli Değişen Madde Fonksiyonu (TB DMF) ve Tek Biçimli Olmayan Değişen Madde Fonksiyonu (TBO DMF) olmak üzere iki şekilde ortaya çıkar (Osterlind ve Everson, 2009)

### I. 1. 1. Tek Biçimli Değişen Madde Fonksiyonu

Aynı yetenek düzeyindeki iki grubun bir maddeyi doğru cevaplama olasılığındaki fark sabit ise, bu durum tek biçimli DMF olarak adlandırılır. Yani tek biçimli DMF, bir maddeyi doğru cevaplama olasılığındaki farklılığın bütün yetenek düzeylerinde tutarlılık göstermesi durumunu belirtir (Camilli ve Shepard, 1994; Zumbo, 1999). Bireylerin yetenek düzeyi değiştikçe maddeyi doğru cevaplamaları arasındaki farklılık da aynı düzeyde değişmektedir. Tek biçimli DMF gösteren maddelerin iki grup için belirlenen madde karakteristik eğrileri kesişim göstermez (Camilli ve Shepard, 1994).

Şekil 1’de referans ve odak gruplarının, tek biçimli DMF gösteren bir maddeyi doğru cevaplama olasılıklarının gösterildiği Madde Karakteristik Eğrisi (MKE) verilmiştir.

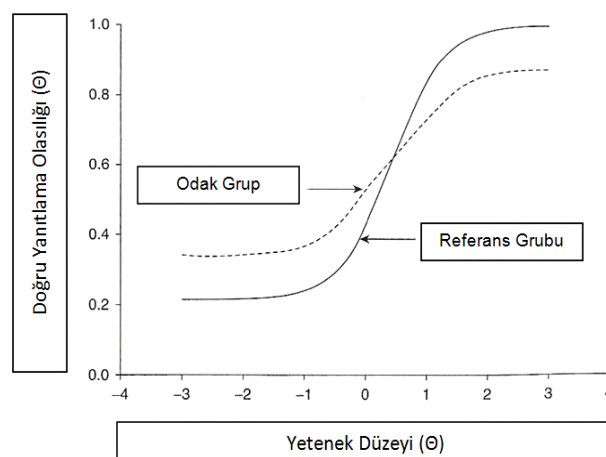


Şekil 1. Tek Biçimli DMF (Osterlind ve Everson, 2009: 12)

Şekil 1’de verilen grafik incelendiğinde, tüm  $\theta$  yetenek düzeyleri boyunca referans gruptaki bireylerin ilgili maddeyi doğru cevaplama olasılıkları, odak gruptaki bireylerden daha yüksek çıkmaktadır.

### I. 1. 2. Tek Biçimli Olmayan Değişen Madde Fonksiyonu

Tek biçimli olmayan DMF, maddeyi doğru cevaplama olasılığındaki farklılığın bütün yetenek düzeyleri için farklı olmasını ifade eder. Yani, yetenek ranjı boyunca bir grubun diğer gruba sağladığı avantaj görece olarak tek biçimli bir yapı göstermiyorsa “Tek Biçimli Olmayan DMF” söz konusu olur (Camilli ve Shepard, 1994). Aynı yetenek düzeyindeki iki grubun maddeyi doğru cevaplama olasılığındaki farklılık miktar ve yön bakımından sabit değildir. Grupların yetenek düzeyi değiştiğinde bireylerin maddeyi doğru cevaplama olasılıkları da değişmektedir. Bir madde  $\theta$ 'nın bir seviyesinde referans gruba küçük bir avantaj sağlarken  $\theta$ 'nın daha yüksek bir seviyesinde daha büyük bir avantaj sağlayabilir. Ya da bir madde odak grupta  $\theta$ 'nın bir ucunda göreceli avantaja sahip iken, diğer ucunda da referans gruba bir avantaj sağlayabilir. Sonuç olarak Tek Biçimli Olmayan DMF grup üyeleri ve yetenek seviyesi ( $\theta$ ) arasında bir etkileşim olduğu zaman ortaya çıkar (Osterlind ve Everson, 2009). Şekil 2’de referans ve odak gruplarının, Tek Biçimli Olmayan DMF gösteren bir maddeyi doğru cevaplama olasılıklarının gösterildiği Madde Karakteristik Eğrisi (MKE) verilmiştir.



Şekil 2. Tek Biçimli Olmayan DMF (Osterlind ve Everson, 2009: 12)

Tek biçimli DMF, DMF'nin belirlenmesinde büyük bir öneme sahiptir. Tek biçimli DMF'nin ortaya çıkması, maddenin belli bir alt gruba sistematik bir avantaj sağladığının göstergesi olarak değerlendirilebilir ve bu durumun meydana gelmesine ilişkin nedenler daha net ortaya konabilir. Fakat, Tek Biçimli Olmayan DMF bulgularının yorumlanması ve sonuçların ortaya çıkma nedenlerinin belirlenmesi oldukça güçtür. Tek Biçimli Olmayan DMF kuramsal olarak mümkün olsa da gerçek durumda nadiren ortaya çıkmaktadır (Camilli ve Shepard, 1994; Mazor, Clauser ve Hambleton, 1994).



## I.2.Değişen Madde Fonksiyonu Belirleme Yöntemleri

Alanyazında DMF yöntemlerini sınıflamada farklı bakış açıları ve buna bağlı oluşan farklı sınıflama biçimleri mevcuttur. DMF belirleme yöntemlerini, Klasik Test Kuramına dayalı yöntemler ve Madde Tepki Kuramı temelli yöntemler olarak iki grupta incelemek mümkündür. Camilli (2006) ve Zumbo (2009) sınıflandırmayı; olasılık tablosuna dayalı yöntemler, Madde Tepki Kuramına dayalı yöntemler ve çok boyutlu yöntemler biçiminde yaparken; Camilli ve Shepard (1994) ise; varyans analizine dayalı yöntemler, MTK'ya dayalı yöntemler, olasılık tablosuna dayalı yöntemler olmak üzere üçe ayırmıştır. Ayrıca DMF yöntemleri, verilere (parametrik olup olmaması), eşleme değişkenine (gözlenen ya da örtük), madde puanlarına (iki kategorili ya da çok kategorili) DMF'yi test etmesi ve etki büyüklüğü ölçümü vermesi ve Tek biçimli ve Tek biçimli olmayan DMF'yi tespit edebilme özelliklerine bağlı olarak da sınıflandırılabilir (Holland ve Wainer, 1993; Wiberg, 2007). Test maddelerindeki yanlılığı bulmaya yönelik ilk çalışmalarda madde güçlük (p) indeksleri arasındaki farklılıkların, grup-madde etkileşimleri ve korelasyon farklılıklarının incelendiği Klasik Test Kuramı ve varyans analizi yöntemleri kullanılmıştır. Klasik Test Kuramı ve varyans analizi kullanılarak geliştirilen DMF belirleme yöntemleri; dönüştürülmüş madde güçlük indeksi (delta plot metodu), altın kural prosedürü, varyans analizi ve korelasyona dayalı yöntemler(maddelerin nokta çift serili korelasyonları arasındaki fark analizi) olarak sıralanabilir. Klasik Test Kuramı ve varyans analizine dayanan bu yöntemlerin en önemli eksiklikleri, örnekleme bağımlı olması ve iki alt gruba ait indeksler birbirinden farklı bulunduğunda, farklılığın grup performansındaki farklılıktan mı yoksa maddeden mi kaynaklandığını söylemenin mümkün olmamasıdır (Camilli ve Shepard, 1994: 22). Madde Tepki Kuramı (MTK) kapsamında ele alınan DMF belirleme yöntemleri ise; alan indeksleri tekniği, madde parametreleri tekniği, olabilirlik oran analizi şeklinde sıralanabilir. (Hambleton, Swaminathan ve Rogers, 1991; Camilli ve Shepard, 1994; Zumbo, 1999). Klasik Test Kuramı'na (KTK) dayalı DMF belirleme yöntemlerinden Mantel-Haenszel ve Lojistik Regresyon yöntemleri iki kategorili maddelerde özellikle Tek biçimli DMF'yi belirlemede kullanılan güçlü yaklaşımlardır (Camilli ve Shepard, 1994).

Ayrıca DMF içeren maddelerin belirlenmesinde istatistiksel testlerin yanı sıra, DMF büyüklüğünün bir ölçüsünü vermeleri açısından bu yöntemler önemli görülmektedir. Bu çalışma kapsamında Klasik Test Kuramı'na dayalı olarak geliştirilen  $\chi^2$  Analizi ile Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) kullanılacağı için bu yöntemler aşağıda kısaca açıklanmıştır.

### I. 2. 1. $\chi^2$ Analizi ile DMF Belirleme Yöntemleri

$\chi^2$  analizi, verilerin kategorik olduğu durumlarda kullanılan parametrik olmayan bir manidarlık testidir.  $\chi^2$  analizinde gözlenen değerler oranı ile beklenen değerler oranı karşılaştırılarak, aradaki farklılığın anlamlı olup olmadığı test edilir.  $\chi^2$  analizine dayanan DMF belirleme yöntemlerinde farklı gruplardan yer alan bireylerin bir maddeyi doğru cevaplama oranları toplam test puanı kategorilerine göre analiz edilir. Kullanılan tekniklerin hesaplanması oldukça kolay olup, Madde Tepki Kuramı'ndaki gibi büyük örneklemelere ulaşılması gerekmemektedir.  $\chi^2$  analizi ile DMF belirleme yöntemleri; tam  $\chi^2$  yöntemi, mantel-haenszel yöntemi, loglinear modeller ve lojistik regresyon yöntemi olmak üzere dörde ayrılır (Benito ve Ara, 2000).

#### I. 2. 1. 1. Mantel-Haenszel Yöntemi (MH)

Mantel-Haenszel yöntemi, iki medikal araştırmacı Nathan Mantel ve William Haenszel tarafından 1950'li yıllarda eşleştirilmiş gruplarda uygulanmak üzere bir Ki-kare tekniği olarak geliştirilmiş ve 1985 yılında Holland tarafından güncellenerek Değişen Madde Fonksiyonu incelenmesi sürecinde kullanılmak üzere önerilmiş, sonrasında Holland ve Thayer (1988) tarafından değişen madde fonksiyonu değerlendirilmesine uyarlanmıştır.

Ki-kare metodu,  $2 \times 2$ 'lik matrislerde satır ve sütunlarda yer alan değişkenler arasındaki lineer ilişki yardımıyla gruplar arası karşılaştırmalar yapmak amacıyla kullanılmaktadır. Günümüzde en yaygın kullanılan DMF belirleme yöntemlerinden biridir (Osterlind ve Everson, 2009). Burada DMF etkisi, hipotez testleri yardımıyla test edilmekte olup, ne kadarlık bir farka DMF denileceği de böylece istatistiksel olarak karara bağlanmaktadır. Yöntemde odak ve referans grubunda yer alan bireyler testten aldıkları toplam puanlar gözetilerek eşleştirme yapmak amacıyla yetenek gruplarına ayrılır. Böylece aynı yetenek puanına sahip bireyler arasında karşılaştırma yapılarak; "aynı yetenek düzeyindeki kişilerin maddeyi doğru cevaplama olasılıkları arasında fark yoktur" hipotezi test edilir. Bu teknik, madde etkisini incelemek için kullanılan olasılık tablosuna dayalı teknikler arasındadır. Olasılık tablosu tekniklerinde parametrik olmayan istatistikler kullanılmaktadır. Bu yöntemlerde gözlenemeyen örtük özelliklerden çok, ölçekle gözlenen değerler sunulmaktadır ve MTK modellerinin gerektirdiği yüksek düzeyde matematik ön bilgisi gerektirmemektedir. Bu durum DMF analizi sonuçlarının kitlelere duyurulması noktasında bir avantaj olarak görülmektedir (Camilli ve Shepard, 1994:102).

Olasılık tablosu oluşturmaya dayanan yöntemlerde DMF analizinin ilk adımında test verisi puanlanmalı ve kodlanmalıdır. Bu veri daha sonra her bir madde için olasılık tablosu hazırlanmasında kullanılacaktır. Bunun için her birey ile ilgili;

- a. Grup üyeliğini gösteren kod ya da etiket
- b. Her bir madde için gerçek cevapları (doğru ya da yanlış şeklinde)
- c. Testten aldığı toplam puan

bilgilerinin bilinmesi gerekmektedir.

Toplam puan her bir birey için doğru cevap sayısı olarak hesaplanmaktadır. Bu durumda K tane 1-0 şeklinde puanlanan maddeden oluşan bir testte toplam puan 0 ve K değerleri arasında değer alacaktır. Ardından karşılaştırma sağlayabilmek için aynı toplam puana sahip bireyler birlikte gruplandırılarak her bir grup için olasılık tablosu oluşturulmaktadır. Bu tablo her bir bireyin odak ya da referans grup üyesi olmasına ve maddeyi doğru ya da yanlış cevaplama göre verilerin çapraz sınıflandırılmasıyla oluşturulur (Camilli ve Shepard, 1994:104). Tüm bireylerin aynı toplam puana sahip olmasından dolayı birbirleri ile kıyaslanabileceği düşünülmektedir. Toplam puan şartı kontrol altına alındıktan sonra odak ve referans grubundaki bireylerin maddeyi doğru cevaplama oranlarının eşit olup olmadığı, Tablo 1'deki veriler gözönüne alınarak referans grup için;  $Prj=Aj/Nrj$  referans grubunun maddeyi doğru cevaplama oranı;  $qrj=Bj/Nrj$  referans grubunun maddeyi yanlış cevaplama oranı hesaplanır ve bu oranlar odak grubu içinde aynı şekilde hesaplanır, hesaplanan bu değerler Tablo 2' de gösterilen şekilde yerleştirilir.

**Tablo 1:** Mantel-Haenszel Yönteminde Kullanılan Veri Düzeni

	1	0	Toplam
R	$A_j$	$B_j$	$N_{rj}$
F	$C_j$	$D_j$	$N_{fj}$
Toplam	$M_{1j}$	$M_{0j}$	$T_j$

R: Referans grup

F: Odak grup

$A_j$ : Referans grupta maddeyi doğru cevaplayanların sayısı

$C_j$ : Odak grupta maddeyi doğru cevaplayanların sayısı

$B_j$ : Referans grupta maddeyi yanlış cevaplayanların sayısı

$D_j$ : Odak grupta maddeyi yanlış cevaplayanların sayısı

$M_{1j}$ : j yetenek düzeyindeki tüm grupta maddeyi doğru cevaplayanların sayısı

$M_{0j}$ : j yetenek düzeyindeki tüm grupta maddeyi yanlış cevaplayanların sayısı

$N_{rj}$ : j yetenek düzeyindeki referans grupta bulunan kişi sayısı

$N_{fj}$ : j yetenek düzeyindeki odak grupta bulunan kişi sayısı

$T_j$ : j yetenek düzeyindeki kişilerin sayısı

**Tablo 2:** j Yetenek Seviyesindeki Bireyler İçin i Maddesinin Olasılık Tablosu

	1	0	Toplam
R	$p_{rj}$	$q_{rj}$	1
F	$p_{fj}$	$q_{fj}$	1

$p_{rj} = A_j / nR$  referans grubunun maddeyi doğru cevaplama oranı,

$q_{rj} = B_j / nR_j$ , referans grubunun maddeyi yanlış cevaplama oranı,

$p_{rj} + q_{rj} = 1$ .

$p_{fj} = C_j / nF$ , odak grubunun maddeyi doğru cevaplama oranı,

$q_{fj} = D_j / nF_j$ , odak grubunun maddeyi yanlış cevaplama oranı,

$p_{fj} + q_{fj} = 1$ .

Olasılık tablosuna dayalı metotlar kategorisine giren tüm DMF belirleme yöntemlerinde bu tablonun kullanımı faydalı olmaktadır (Camilli ve Shepard, 1994:106-107). Eşitlik 1'deki gibi hesaplanan olasılık oranı, bir gruptaki bireylerin bir maddeye doğru cevap verme olasılığının diğer gruptaki bireylerden yüksek olup olmadığını belirler. Denkleştirilmiş iki grubun performansı daha sonra olasılık oranı değerleri dikkate alınarak karşılaştırılır (Camilli ve Shepard, 1994).

$$\alpha_j = \frac{p_{rj}/q_{rj}}{p_{fj}/q_{fj}} \quad (1)$$

$$\alpha_{MH} = \frac{\sum_{j=1}^s A_j D_j / T_j}{\sum_{j=1}^s B_j C_j / T_j} \quad (2)$$

Olasılık oranı  $(0, \infty)$  arasında değerler alabilir. Olasılık oranının hesaplanmasından sonra Eşitlik 2'de belirtilen  $\alpha_{MH}$  hesaplanır.  $\alpha_{MH}$  değeri, test sonuçlarının odak grubundaki kişiler için o kadar kat daha zor ( $\alpha_{MH}=3$  ise 3 kat daha zor) olduğu anlamına gelir.  $\alpha_{MH}$  değerinin 1 olması maddede DMF olmadığı; 1'den küçük olması odak grup lehine DMF olduğu; 1'den büyük olması da referans grup lehine DMF olduğu şeklinde yorumlanır.

$$\beta_{MH} = \log_e(\alpha_{MH}) \quad (3)$$

$\alpha$  MH değerinin yorumlamasının zor olması nedeniyle bu değer, yorumlamaların daha kolay ve pratik olduğu, eşitlik 3'te belirtildiği gibi logaritmik ölçeğine dönüştürülür.  $\beta_{MH}$  değeri ise bir indeks değeridir.  $\beta_{MH} = 0$  olması DMF'nin yokluğunu yani gruplar arasında maddeyi cevaplama olasılıkları için bir fark olmamasını;  $\beta_{MH} < 0$  olması maddenin odak grup lehine DMF gösterdiğini ve  $\beta_{MH} > 0$  olması ise maddenin referans grup lehine DMF gösterdiğini belirtir (Angoff, 1993).

$$\Delta_{MH} = -2,35 \beta_{MH} \quad (4)$$

Dönüşüm sonrasında indeks; eşitlik 4'te belirtilen odds oranının doğal logaritması olarak  $\Delta_{MH}$  şeklinde ifade edilmektedir (Holland ve Thayer, 1993; Wiberg, 2007).  $\Delta_{MH} = 0$  olması DMF'nin yokluğunu yani gruplar arasında maddeyi cevaplama olasılıkları için bir fark olmamasını,  $\Delta_{MH} < 0$  olması maddenin referans grup lehine DMF gösterdiğini ve  $\Delta_{MH} > 0$  olması ise maddenin odak grup lehine DMF gösterdiğini belirtir (Holland ve Wainer, 1993). Etki büyüklüğü istatistiği olarak  $\Delta_{MH}$  istatistiği Eğitimsel Test Hizmetleri (Educational Testing Service-ETS) tarafından DMF miktarının ne derece büyük olduğunu yorumlanmasında da kullanılır (Wiberg, 2007).  $\beta_{MH}$  DMF'nin büyüklüğü hakkında bilgi verirken;  $MH\chi^2$  değeri istatistiksel olarak DMF'nin test edilmesinde kullanılır.  $MH\chi^2$ , 1 serbestlik derecesinde, yaklaşık normal dağılıma sahip bir  $\chi^2$  istatistiğidir ve aşağıdaki gibi hesaplanır (Camilli ve Shepard, 1994:120).

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

$$VAR(A_j) = \frac{n_{Rj}n_{Fj}m_{1j}m_{0j}}{T_j^2(T_j-1)} \quad (5)$$

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

$A_j$ : Madde için referans grubun  $j$  yetenek seviyesindeki gözlenen doğru cevap sayısı,

$E(A_j)$ : Madde için referans grubun  $j$  yetenek seviyesindeki beklenen doğru cevap sayısı

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

$m_{1j}$ :  $j$  yetenek seviyesi için maddeye 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 .

Madde için gözlenen doğru cevap sayısı, madde için beklenen doğru cevap sayısından fazla ise ( $A_j > E(A_j)$ ) referans grup lehine; tersine durumda ise referans grup aleyhine DMF gözleendiği yorumu yapılabilir (Camilli ve Shepard, 1994:120). Madde etkisinin istatistiksel olarak test edilmesi önerilmektedir (Holland ve Thayer, 1988; Angoff, 1993; akt: Holland ve Wainer, 1993).  $MH\chi^2$  istatistiği yardımıyla madde performansı ile grup üyelikleri arasında bir ilişki olup olmadığı belirlenebilir. Fakat yapılan kıkare testi DMF'nin büyüklüğü yani ilişkinin gücü hakkında bilgi vermez.

$\Delta_{MH}$  kullanılarak test maddelerinin farklı derecelerde gösterdiği DMF'yi belirlemede Zieky (1993) tarafından üç kategori düzeyi önerilmiştir. Bu üç kategori önemsiz (A), orta derece (B) ve büyük (C) olarak belirlenmiştir. Bu kategoriler ve karar verme ölçütleri Tablo 3'te gösterilmiştir.

**Tablo 3:** Mantel Haeszel DMF İstatistiğinin Yorumlanması

Kategoriler	D Değeri	Yorum
A	$ \Delta_{MH}  < 1$	DMF yoktur veya göz ardı edilebilecek düzeydedir.
B	$1 \leq  \Delta_{MH}  \leq 1,5$	Orta düzeyde DMF vardır
C	$ \Delta_{MH}  > 1,5$	Yüksek düzeyde DMF vardır.

Zieky (1993), elde edilen  $\Delta_{MH}$  değeri 1'den küçük ise maddede DMF'nin olmadığını veya ihmal edilebileceğini; 1 ile 1,5 arasında ise maddede B düzeyinde DMF bulunduğunu ve 1,5'den büyük ise C düzeyinde DMF bulunduğunu belirtmiştir.

### 1.2.1.2. Lojistik Regresyon Yöntemi (LR)

Lojistik Regresyon bağımlı değişkenin sürekli ya da nicel bir değişken olmadığı, kategorik ya da sınıflamalı olduğu durumlar için uygun bir analiz türüdür. Lojistik regresyon genellikle bir test veya alt test puanı ölçüt alınarak, aynı yetenek düzeyindeki farklı grup üyelerinin bir maddeye doğru cevap verme olasılıklarının istatistiksel olarak modellenmesine dayanır. Lojistik regresyon analizinde amaç, kategorik bağımlı değişkenin değerini tahmin etmek olduğundan burada yapılmaya çalışılan iki ya da daha fazla gruba ilişkin 'üyelik' tahminidir. Bu yöntemde, 1 ve 0 biçiminde puanlanan madde cevapları bağımlı değişken; grup değişkeni, bireylerin ölçek puanı ve grupla ölçek puanı etkileşimi ise bağımsız değişken olarak kullanılır.

Bu yöntemde iki kategorili bir bağımlı değişken (maddeden alınan 0 – 1 puan), bir gruplama değişkeni (1=referans grubu, 2= odak grubu), testten alınan toplam puan ve gruplama değişkeni ile toplam puanın ortak etkisinin incelendiği bir eşitlik kurulur. Lojistik regresyon yardımıyla, bağımsız değişkenlerin bağımlı değişkenin varyansının yüzde kaçını açıklayabildiği, bağımsız değişkenlerin bağımlı değişkenler üzerindeki etki büyüklüğü ve bağımsız değişkenler arasında etkileşim olup olmadığı belirlenebilir.

LR yöntemiyle DMF durumunu belirleme çalışmaları ilk olarak Swaminathan ve Rogers(1990) tarafından uygulanmıştır. Bir maddenin doğru cevaplanma olasılığını tahmin eden lojistik regresyon modeli; maksimum olabilirlik yönteminin kullanıldığı formüller eşitlik 6 ve eşitlik 7’de verilmiştir (Swaminathan ve Rogers, 1990; Camilli ve Shepard, 1994:126). Olabilirlik oranının doğal logaritması olarak bağımlı değişken,

$$P(uij = 1 | \theta) = \frac{\exp \psi_i}{1 + \exp \psi_i} \quad (6)$$

6 nolu eşitlikte,  $P(ui=1 | \theta)$ ,  $\theta$  yetenek düzeyindeki j. gruptaki i. bireyin maddeyi doğru cevaplama olasılığını ve  $\Psi_i$ , bir maddeden i bireyi için beklenen log odds değerini yani olabilirlik oranının doğal logaritmasını göstermektedir (Osterlind ve Everson, 2009). DMF, eşlenmiş yetenek düzeyindeki bireylerin maddeyi cevaplama farklılık göstermesi durumunda ortaya çıkar. Yani LR eğrileri her iki grupta (referans ve odak gruplarda) aynı olduğunda DMF yoktur (Osterlind ve Everson, 2009).

LR modeli;

$$\Psi_i = \delta + \zeta_1 X_i + \zeta_2 G_i + \zeta_3(G_i X_i) \quad (7)$$

i: Birey,

$X_i$ : Bireyin toplam puanı,

$G_i$ : Bireyin grubu,

$\zeta_1$ : Odds oranı logaritmalarının birleştirilmiş hali,

$\zeta_2$ : Birey yeteneklerinin eşit olduğunu gösteren parametre,

$\zeta_3$ : Tek biçimli Değişen Madde Fonksiyonu parametresi.

şeklinde verilmektedir. LR modelindeki katsayılar maksimum olabilirlik yöntemiyle kestirilmekte ve modellerin birbirlerine göre anlamlı fark oluşturup oluşturmadığı olabilirlik oran testiyle incelenmektedir (Osterlind ve Everson, 2009).

Tek biçimli olmayan DMF incelemesi için regresyon denkleminde bir de etkileşim terimi eklenmektedir. Eşitlik 7 göz önünde bulundurularak 3 modelde bulunan katsayılar maksimum olabilirlik yöntemiyle kestirilmekte ve modellerin birbirlerine göre anlamlı fark oluşturup oluşturmadığı incelenmektedir ( Camilli ve Shepard, 1994:125).

$$\Psi_i = \delta + \zeta_1 X_i \quad (8) \text{ (Model I)}$$

$$\Psi_i = \delta + \zeta_1 X_i + \zeta_2 G_i \quad (9) \text{ (Model II)}$$

$$\Psi_i = \delta + \zeta_1 X_i + \zeta_2 G_i + \zeta_3 (G_i X_i) \quad (10) \text{ (Model III)}$$

$\zeta_1$ : Maddeyi doğru cevaplama eğilimindeki bireyler arasındaki yetenek farklılıkları

$\zeta_2$ : bir maddeyi doğru cevaplama olasılığının iki grup için farklılığı

$\zeta_3$ : Etkileşim terimi(TBO DMF varlığı göstergesi)

DMF analizinde değişkenler modele hiyerarşik olarak sırayla dahil edilirler. Model I’de sınırlanacak değişken, yani toplam puan “X” modelde yerini alır. Model II’de, grup değişkeni (referans, odak) eklenir ve üçüncü modelde ise önceki değişkenlere ilave olarak etkileşim değişkeni (toplam puan ve grup) modele alınır. Kurulan bu modeller sayesinde incelenen maddenin tek biçimli ya da tek biçimli olmayan DMF gösterip göstermediği belirlenebilir. Model III’ten elde edilen  $\chi^2$  değeri ile Model I’deki  $\chi^2$  değerinin karşılaştırılmasıyla DMF’nin varlığı tespit edilir. Elde edilen  $\chi^2$  değerinin anlamlı olması DMF’nin varlığına işaret eder. Aşamaların her birinde ki kare değeri elde edildiği gibi, aynı zamanda  $R^2$  değeri de hesaplanır. Model I ve Model III’den elde edilen  $\Delta R^2 = R_3^2 - R_1^2$  değeri ile var olan DMF ’nin büyüklüğü belirlenir. Model III’den elde edilen  $R_3^2$  ile Model II’den elde edilen  $R_2^2$  farkı ( $R_3^2 - R_2^2$ ) Tek Biçimli Olmayan DMF’nin, Model II’den elde edilen  $R_2^2$  ile Model I’den elde edilen  $R_1^2$  farkı ( $R_2^2 - R_1^2$ ) ise Tek Biçimli DMF’nin varlığına işaret eder. Model III ve Model I arasındaki R değerlerinin farkı  $\Delta R^2$  değeri (Nagelkerke  $R^2$  değer farkları) DMF yorumlanmasında kullanılan etki büyüklüğünü belirtmektedir (Zumbo, 1999). DMF’nin düzeyine karar vermede,  $\Delta R^2$  değerlerinin yorumlanması için



Zumbo ve Thomas (1996) ile Jodoin ve Gierl (2001) iki ayrı ölçüt sınıflaması önermişlerdir. Bu önerilen sınıflama kategorileri Tablo 4'te gösterilmektedir.

**Tablo 4:** Lojistik Regresyon Tekniğinde Değişen Madde Fonksiyonu Düzeyi İçin  $R^2$  Ölçütleri

Gierl ve Arkadaşlarının Ölçütleri	Zumbo ve Thomas'ın Ölçütleri	KARAR
$R^2 < 0,035$	$R^2 < 0,13$	Önemsiz düzeyde DMF vardır (A).
$0,035 \leq R^2 < 0,070$	$0,13 \leq R^2 < 0,2$	Orta düzeyde DMF vardır (B).
$R^2 \geq 0,070$	$R^2 \geq 0,26$	Yüksek düzeyde DMF vardır (C).

Zumbo (1999), LR yönteminde ölçüt değişkenin kategori edilmesine ihtiyaç olmaması, Tek biçimli ve Tek biçimli Olmayan DMF'yi belirleyebilmesi, ikili puanlanan maddelerde kullanılan sıralı puanlanan maddeler için de yaygınlaştırılabilir olmasını LR yönteminin avantajları olarak sıralamaktadır. Bu avantajlarının yanında, lojistik regresyon analizinin, tutarlı sonuçlar sağlayabilmesi için Mantel-Haenszel yöntemine göre daha büyük örneklere ihtiyaç duyması en önemli sınırlılığıdır (Holland ve Wainer, 1993; Zumbo, 1999).

### I.3. İlgili Araştırmalar

Bu bölümde, maddelerin test içerisinde madde güçlük indekslerine göre kolaydan zora, zordan kolay ve seçkisiz olarak sıralanmasının, test ve madde istatistiklerine etkisini, diğer değişkenlerle olan etkileşimlerini ve testte yer alan maddelerin farklı sıralanmasının Değişen Madde Fonksiyonu'yla (DMF) incelenmesini ve DMF analizlerinde bu araştırma kapsamında yöntemleri ele alan çalışmalardan bahsedilecektir. Yurt içinde ve yurt dışında yapılan çalışmalar incelendiğinde, maddelerin çeşitli şekillere göre sıralanmasının test ve madde istatistikleri üzerinde etkisini ya da diğer değişkenlerle etkileşimini inceleyen çalışmaların bulunduğu, fakat madde sıralanmasının DMF'yle incelenmesini araştıran çok az çalışmanın bulunduğu görülmektedir.

#### I.3. 1. Yurt İçinde Yapılan Çalışmalar

Yergin (2007), çoktan seçmeli testlerde maddeleri madde güçlük düzeylerine göre kolaydan zora, zordan kolay ve seçkisiz olarak sıralamanın bazı madde ve test istatistiklerine etkisi olup olmadığını araştırmıştır. Çalışmaya Mersin ili merkez ilköğretim okullarında 7. sınıfta okuyan 513 öğrenci katılmıştır. Çalışmada çoktan seçmeli 25 maddeden oluşan bir okuduğunu anlama testi geliştirilmiş ve

test içindeki maddeler, madde güçlük düzeylerine göre kolaydan zora, zordan kolay ve seçkisiz olarak sıralanarak farklı üç test formu oluşturulmuştur. Testin bu üç formu ile test içindeki maddeleri, madde güçlük düzeylerine göre kolaydan zora, zordan kolay ve seçkisiz olarak sıralamanın test puanlarına, maddelerin güçlük ve ayıricılık düzeyi indekslerine ve testin güvenilirliğine etkisi olup olmadığı araştırılmıştır. Elde edilen verilere yapılan analizler sonucunda testin maddelerini güçlük düzeylerine göre kolaydan zora, zordan kolay ve seçkisiz olarak sıralamanın, bazı maddelerin madde güçlük indeksleri ve madde ayıricılık güçlükleri arasında farklılaşmaya neden olduğu gözlenmiştir. Ayrıca öğrencilerin testin üç formundan aldığı puanların ortalamaları arasında anlamlı bir farklılık gözlenmemiştir. Öğrencilerin test maddelerinin güçlüğüne ilişkin algı puanlarında da anlamlı bir farklılık gözlenmemiştir.

Tan (2009), çoktan seçmeli testlerde aynı maddelerin farklı sıralarda öğrencilere sorulmasının maddelerin güçlükleri ve madde ayırt edicilik indekslerine etkisini araştırmıştır. Çalışmaya 2007-2008 öğretim yılında, Uludağ Üniversitesi, Eğitim Fakültesi, birinci sınıf öğrencilerinden “Eğitim Bilimine Giriş” dersinin dönem sonu sınavına giren 996 kişi katılmıştır. Dönem sonu sınavı 50 çoktan seçmeli maddeden oluşan maddelerin sıralarının ilgili oldukları konuya göre değiştirildiği A ve B formu olmak üzere iki farklı formdan oluşmaktadır. A ve B formları öğrencilere random olarak dağıtılmıştır. Yapılan analiz sonucunda 50 maddenin 11 tanesinde madde güçlük düzeyleri arasındaki farklar istatistiksel olarak anlamlı bulunmuştur. Yani maddelerin sorulma sırasının madde güçlüklerini anlamlı olarak etkilediği görülmüştür. Aynı maddelerin farklı sıralarda sorulması sonucu her madde için A ve B formundan olmak üzere iki tane madde ayırt edicilik gücü elde edilmiştir 50 maddenin 3 tanesinde madde ayırt edicilik güçlükleri arasındaki farklar istatistiksel olarak anlamlı bulunmuştur. Yani maddelerin testte soruluş sırasının madde ayırt edicilik güçlüklerini az da olsa anlamlı olarak etkilediği gözlenmiştir.

Yılmaz (2014), testteki madde ve seçenek sırasında yapılan değişikliklere bağlı olarak öğrencilerin cevapladığı soru kitapçık türüne göre bir maddeyi doğru cevaplandırma olasılığının farklılaşıp farklılaşmadığını araştırmıştır. Çalışmada 2010 yılında uygulanan Uzman Jandarma Okulu Giriş Sınavı (JANU) kullanılmıştır. Sınavda Türkçe, Matematik, Tarih, Coğrafya ve Temel Yurttaşlık Bilgisi alanlarını içeren 100 çoktan seçmeli maddenin yer aldığı, maddelerin ve seçeneklerinin yerleri değiştirilerek oluşturulan dört test kitapçığı kullanılmıştır. Bazı, köklerinde öncüller içeren maddelerin öncüllerinin yerlerinin değiştirilmesinin, değişen madde fonksiyonuna neden olup olmadığı belirlenmeye çalışılmıştır. Öncülleri değiştirilen maddelerin seçeneklerinin yerleri değiştirilmemiştir. Verilerin Değişen Madde Fonksiyonu analizinde Mantel Haenszel yöntemi kullanılmıştır. Yapılan analizler sonucunda alt testler göz önünde bulundurulmadan yapılan madde sıralamalarındaki değişikliklerin, Eğitimsel Test Hizmetleri’ne (ETS) göre B ve C kategorisinde DMF’ye sahip madde sayılarında artışa

neden olduğu gözlenmiştir. Ayrıca maddelerin seçeneklerinin yerlerinin değiştirilmesinin C düzeyinde DMF'ye neden olduğu gözlenmiştir.

### **I.3. 2. Yurt Dışında Yapılan Çalışmalar**

MacNichol (1960), çalışmasında çoktan seçmeli test maddelerinin madde güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz olarak sıralanmasının öğrencilerin test puanları üzerine etkisini araştırmıştır. Çalışmada 50 çoktan seçmeli maddeden oluşan sözel analogi testinin farklı üç formu 1500 lise öğrencisine uygulanmıştır. Verilerin analizinde t-testi kullanılarak, maddelerin güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz sıralandığı formlardaki test ortalamaları karşılaştırılmıştır. Yapılan analiz sonucunda, öğrencilerin maddelerin kolaydan zora ve seçkisiz olarak düzenlendiği formlarda, maddelerin zordan kolaya doğru düzenlendiği forma göre daha yüksek test puanına sahip olduğu ve puanlar arasındaki farkın anlamlı olduğu gözlenmiştir.

Brenner (1964), test içindeki maddeleri madde güçlük düzeyine göre sıralamanın, madde ayırıcılığına, madde güçlüğüne ve testin güvenilirliğine etkisini araştırmıştır. Çalışmaya Ohio State Üniversitesi'nde Eğitim Psikolojisi dersi alan 407 öğrenci katılmıştır. Araştırmada, dersin konularıyla ilgili, her biri dört seçenekli ve çoktan seçmeli ve madde güçlük indeksleri 0,17 ile 0,94 arasında değişen 40 maddelik bir test dört ara sınavda uygulanmıştır. Birinci sınavda kullanılan maddeler madde güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz olarak sıralanarak üç test formu oluşturulmuştur. İkinci sınav için hazırlanan test maddeleri için iki test formu oluşturulmuştur. İkinci sınav, 80 madde içinden 10 en kolay, 10 en zor ve 20'si diğer güçlük düzeylerini temsil edecek şekilde seçilen 40 maddeden oluşmaktadır. İkinci sınavdaki birinci form, en kolay 10 madde en kolaydan en zora sıralanacak biçimde testin başına yerleştirilerek ve diğer 30 madde seçkisiz olarak sıralanarak hazırlanmıştır. İkinci form ise, 10 en zor madde en zordan en kolaya sıralanacak biçimde testin başına yerleştirilerek ve geriye kalan 30 madde seçkisiz sıralanarak oluşturulmuştur. Üçüncü sınavda ise, maddeler güçlük düzeylerine göre kolaydan zora ve zordan kolaya doğru sıralanarak iki form şeklinde uygulanmıştır. Dördüncü sınavda da ikinci sınavda olduğu gibi seçilen maddeler üzerinden artan ya da azalan güçlük düzeyine göre maddeler sıralanarak iki farklı form oluşturulmuştur. Dört sınav için de, madde güçlükleri arasındaki farkların manidar olmadığı; testlerin güvenilirlikleri arasında fark olmadığı ve yalnızca ikinci sınavda maddelerin ayırıcılıklarında farklılaşma olduğu sonucuna ulaşılmıştır.

Sax ve Cromax (1966), farklı madde düzenlemelerinden oluşan çeşitli test formlarının test performansı üzerine etkisini araştırmışlardır. Araştırma; Sax ve Carr'ın (1962) çalışmasında kullandığı, madde güçlük düzeylerine göre farklılık gösteren 70 maddeden oluşan Henmon-Nelson Zihinsel Yetenek Testi' nin dört farklı formu (Form E: maddelerin kolaydan zora doğru sıralandığı; Form H: maddelerin zordan kolaya doğru sıralandığı; Form M: kolay maddelerden oluşan her altı maddenin

arasına ortalama zorluk düzeyinde maddelerin serpiştirildiği; Form R; maddelerin seçkisiz olarak sıralandığı form) Hawaii Üniversitesindeki Eğitim Bilimine Giriş dersi alan 276 öğrenciye uygulanmıştır. Formlar öğrencilere random olarak dağıtılıp test için öğrencilere 30 dakika süre verilmiştir. Çalışmanın diğer paralel kısmında farklı dört form; Psikolojiye giriş dersi alan 191 öğrenciye 48 dakika süre verilerek uygulanmıştır. Bu uygulamanın amacı test süresi sınırlamasının madde sıralamasını önemli kılıp kılmadığı dolayısıyla test performansı üzerine etkisini araştırmaktır. Çalışmanın sonucunda maddelerin ortalamasının, maddelerin güçlük düzeylerine göre kolaydan zora doğru sıralandığı formda en yüksek olduğu ve zaman sınırlaması olan testlerde madde sıralamalarının önem teşkil ettiği ve böylece öğrencilerin test performansını etkilediği, ayrıca madde sıralamaları zordan kolayca doğru olan testin ortalama ve ortalamasının üzerindeki öğrenciler arasındaki farkı en iyi ortaya koyduğu sonucuna ulaşılmıştır.

Smouse ve Munz (1968), test içindeki maddeleri kolaydan zora, zordan kolayca ve seçkisiz olarak sıralamanın sınav kaygısı ve test başarısı üzerindeki etkisini araştırmıştır. Araştırmaya Oklahoma Üniversitesi'nde Psikoloji dersi alan 113 öğrenci katılmıştır. Araştırmaya katılan öğrenciler, yılsonu sınavı için iki gruba ayrılmışlar ve gruplardan birine yönerge verilerek sınav kaygı düzeyleri artırılmıştır. Sınavda, madde güçlük düzeyleri daha önceki dönemlerden belirlenmiş olan maddeler arasından seçilen, 100 çoktan seçmeli maddeden oluşan bir test formu kullanılmıştır. Bu test içindeki maddeler, madde güçlük indekslerine göre kolaydan zora, zordan kolayca ve seçkisiz sıralanarak üç test formu şeklinde hazırlanıp öğrenciler bu formlara seçkisiz olarak atanmışlardır. Sınav kaygı düzeyi ölçümü için, sınavın sonuna eklenen Çoklu Etki Belirleme Kontrol Listesi (Multiple Affect Adjective Check List) kullanılmıştır. Yapılan varyans analizi sonucunda, iki grupta yer alan ve farklı test formlarını cevaplayan öğrencilerin test puanı ortalamaları arasında farklılaşma olmadığı gözlenmiştir.

Marso (1970), çalışmasında testteki maddelerin düzenlenmesi ile güç testlerindeki öğrenci performansları ve test için ayrılan süre arasında bir ilişki bulunup bulunmadığını incelemek için iki deney düzeneği tasarlamıştır. Bu çalışma için; birinci hipotez madde güçlüğü gözetilerek yapılan madde sıralamalarının güç testlerindeki öğrenci performanslarını ve test için ayrılan süreyi etkilemediğidir. İkinci hipotez; maddelerin içerik olarak benzer konulara göre ya da derste işleme sıraları gözetilerek sıralanmasının, güç testlerindeki öğrenci performansı ve test için ayrılan süre üzerinde etkili olmadığıdır. Üçüncü hipotez; madde sıralama faktörünün test kaygısı faktörü ile etkileşim halinde olmadığıdır. Bunu araştırmak için madde güçlükleri 0 ile 1 aralığında değişen 139 maddeden oluşan Hızlı Kelime Testi'nin (Word Quick Test-QWT) maddeleri kolaydan zora, zordan kolayca ve seçkisiz olarak sıralandığı üç farklı formu düzenlenmiş ve Bowling Green State Üniversitesi'nde Eğitim Psikolojisine Giriş dersi alan 122 öğrenciye uygulanmıştır. Her öğrencinin testi bitirme süresi kaydedilmiştir. Sınav kaygısı ölçekleri birkaç gün önce her öğrenciye uygulanmıştır. Öğrencilerin başarı puanları ve bireysel testi çözme süreleri kullanılarak yapılan varyans analizi sonuçlarına göre; başarı ya da güç testlerinde, madde güçlük

düzeyine göre madde sıralamanın, öğrenci performansları ve test için gerekli süreye az etki ettiği ya da hiç etki etmediği sonucuna varılmıştır. Ayrıca güçlük düzeylerine göre madde sıralamanın öğrenci kaygı düzeyiyle ilişkili olmadığı görülmüştür. 2. deney düzeneğinde ise; Bowling Green State Üniversitesinde Eğitim Psikolojisine Giriş dersi alan 156 öğrenciye maddelerin seçkisiz olarak sıralandığı, benzer kavram ya da materyallerin birlikte sunulduğu ve derste işlenen konu başlıklarına göre sıralandığı, konulara göre sıralandığı üç farklı formdan oluşan 103 maddeden oluşan final testi uygulanmıştır. Final sınavı puanları ve her öğrencinin bu sınav için harcadığı zaman göz önünde bulundurularak yapılan varyans analizi sonucunda madde sıralamalarının öğrenci performansı üzerinde etkili olmadığı ve test kaygı faktörü ile etkileşim halinde olmadığı belirlenmiştir.

Kestenbaum ve Weiner (1970), testteki maddelerin, farklı sıralanmasının öğrencilerin başarı performanslarına, öğrencilerin başarı motivasyonlarına ve test kaygılarına etkisini araştırmıştır. Çalışmada 7. ve 8. sınıf öğrencilerinin okuma becerisini ölçen okuduğunu anlama testinin maddelerinin madde güçlük düzeylerine göre kolaydan zora doğru ve seçkisiz olarak sıralandığı iki farklı test formu kullanılmıştır. Çalışmaya 43 kız 36 erkek olmak üzere toplamda 79 öğrenci katılmıştır. Ayrıca öğrencilere Çocuklar İçin Test Kaygısı Ölçeği (The Test anxiety Scale For Children) ve Çocuklar İçin Başarı Motivasyonu Ölçeği (The Children's Achievement Motivation Scale) uygulanmıştır. Araştırmanın sonucunda, öğrencilerin başarı performanslarının madde sıralamalarıyla ilişkili olmadığı sonucuna varılmıştır. Ayrıca hem kız hem de erkek öğrencilerin okuma performanslarının başarı motivasyonlarıyla pozitif yönde ilişkili, test kaygılarıyla negatif yönde ilişkili olduğu sonucuna varılmıştır.

Huck ve Bowers (1972), çoktan seçmeli bir testteki bir maddenin güçlük düzeyinin, ardından gelen maddenin güçlük düzeyinden etkilenip etkilenmediğini, yani çoktan seçmeli başarı testlerinde maddelerin güçlük düzeylerini ve madde sıra etkilerini incelemişlerdir. Çalışmayı gerçekleştirmek amacıyla iki farklı çalışma yürütülmüş ve her iki çalışmada da madde sıra etkisinin maddelerin gerçek güçlük düzeylerini değiştirip değiştirmediğini incelemek amacıyla dengelenmiş Latin kareleri kullanılmıştır. 1. çalışmada, Northwestern Üniversitesinde Psikolojiye Giriş dersi alan 120 öğrenciye 60 çoktan seçmeli maddenin bulunduğu final sınavının on farklı formu uygulanmıştır. Yapılan analizler sonucunda, madde sıra etkisinin maddenin gerçek güçlük düzeyini değiştirmediği sonucuna varılmıştır. İkinci çalışmada ise, aynı şekilde madde sıra etkisinin maddelerin gerçek güçlük düzeyine etkisini incelemek amacıyla Northwestern Üniversitesinde Psikolojiye Giriş dersi alan 162 öğrenciye 50 çoktan seçmeli maddenin bulunduğu ara sınavın madde sıra etkisini daha az grupta kontrol altına almak için altı farklı formu uygulanmıştır. Yapılan analiz sonucunda, ilk çalışmada olduğu gibi madde sıra etkisinin maddelerin gerçek güçlük düzeyini değiştirmediği görülmüştür.

Hambleton ve Traub (1974), test maddelerini kolaydan zora ve zordan kolayca sıralamanın, sınav stresi, sınav kaygısı ve test performansına etkisini incelemişlerdir. Çalışma, Toronto ve Ontario'daki iki okulda, matematik başarısı düşük olduğu için yaz okuluna giden 11. sınıf öğrencileri arasından seçilen 106 öğrenciyle gerçekleştirilmiştir. Çalışmada sınav kaygısını ölçmek için Başarı Kaygı Test'i (The Achievement Anxiety Test); stresi ölçmek için kalp atım hızını kaydeden bir cihaz; matematik başarısını ölçmek için de Eğitimsel Test Hizmetleri (ETS) tarafından hazırlanan Yardımcı Matematik Testi Cebir II (Cooperative Mathematics Test Algebra II) kullanılmıştır. Başarı kaygı testi matematik sınavından iki hafta önce uygulanmıştır. Matematik testinin maddeleri madde güçlük indeksine göre kolaydan zora ve zordan kolayca sıralanarak iki form şeklinde öğrencilere sunulmuştur. Stres ölçümü için, katılımcılara testin başında, ortasında ve sonunda ölçüm alınacağı söylenmiş, test başladıktan sonra 10, 20, 30 ve 40. dakikalarda ve test bittikten sonra ölçüm yapılmıştır. Araştırmanın sonucunda, maddelerin güçlük düzeylerine göre kolaydan zora doğru sıralandığı testin ortalamasının, zordan kolayca doğru sıralanan testin ortalamasından daha yüksek olduğu; fakat cinsiyete ilişkin bir farklılaşma olmadığı belirlenmiştir. Ayrıca, test sırasındaki stres puanlarının farklı formları alan öğrenciler arasında değişmediği gözlenmiştir.

Barcikovski ve Olsen (1974); test maddelerinin güçlük düzeylerine göre kolaydan zora ve zordan kolayca doğru sıralanması durumunda test puanının etkilenip etkilenmediğini, testin uygulandığı kişilerin madde türüne göre (çoktan seçmeli, doğru- yanlış) testlerin güçlük düzeyini algılamasında fark olup olmadığını ve test puanının doğru-yanlış ve çoktan seçmeli madde türlerine göre değişip değişmediğini araştırmıştır. Araştırmaya, Ohio State Üniversitesinde Eğitim Bilimine Giriş dersini alan 85 öğrenci katılmıştır. Araştırmada ilkokullarda okuma öğretiminin temel ilkeleri ile ilgili olan bir test kullanılmıştır. Bu testin maddeleri daha önceki dönemlerdeki ara sınavlardan madde güçlük ve ayırıcılık indeksleri elde edilmiş olan çoktan seçmeli ve doğru yanlış maddelerinden oluşmaktadır. Testte 140 maddenin bulunduğu madde havuzundan madde güçlük düzeylerine göre kolay, orta ve zor olarak ayrılan 60 madde alınmıştır. 10 tane çoktan seçmeli zor madde (Z-ÇS), 10 tane doğru-yanlış zor madde (Z-DY), 10 tane çoktan seçmeli orta güçlükte madde (O-ÇS), 10 tane doğru yanlış orta güçlükte madde (O-DY), 10 tane çoktan seçmeli kolay madde (K-ÇS), 10 tane doğru yanlış kolay madde olmak üzere maddeler zordan kolayca ve kolaydan zora sıralanarak iki farklı form oluşturulmuştur. Testi cevaplarırken katılımcıların her maddeyi "1 çok kolay, 2 kolay, 3 orta, 4 zor, 5 çok zor" şeklinde değerlendirmeleri ve testteki maddeleri verilen sırada cevaplamaları istenmiştir. Yapılan çalışma sonucunda çoktan seçmeli 30 maddenin (10 kolay, 10 orta güçlük, 10 zor), maddelerin zordan kolayca düzenlenmiş formda kolaydan zora düzenlenmiş forma göre daha kolay algılandığı sonucuna varılmıştır. Doğru-yanlış türündeki 30 (10 kolay, 10 orta güçlük, 10 zor) maddede ise sadece 10 tane zor maddenin, maddelerin zordan kolayca doğru düzenlenmiş formunda kolaydan zora doğru düzenlenmiş formuna göre daha kolay algılandığı ve kolay ve orta güçlükteki maddelerde anlamlı bir farklılaşma olmadığı sonucuna

varılmıştır. Maddelerin ortalamaları incelendiğinde ise çoktan seçmeli maddelerde maddelerin puanının sadece kolay maddelerde, maddelerin zordan kolaya doğru sıralandığı formda kolaydan zora doğru sıralandığı forma göre daha yüksek olduğu görülmüştür. Doğru yanlış türündeki maddelerde ise kolay, orta ve zor maddelerden alınan puanlar açısından maddelerin zordan kolaya ve kolaydan zora doğru sıralandığı formlar arasında manidar bir farklılaşma olmadığı gözlenmiştir. Test puanları açısından çoktan seçmeli ve doğru-yanlış türündeki maddelerde maddelerin kolaydan zora ve zordan kolaya doğru sıralandığı form arasında manidar bir farklılaşma olmadığı gözlenmiştir.

Towle ve Merrill (1975), maddelerin güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz olarak sıralanmasının öğrencilerin test performansına etkisini ve madde güçlük indeksine göre maddelerin sıralanmasının Munz ve Smouse'un (1968) çalışmasında tanımlanan kaygı tipleriyle etkileşimini belirlemeye çalışmışlardır. Çalışmaya Tallahassee Community Kolejinde okuyan Matematik dersini alan ve Florida State Üniversitesinde Eğitim Psikolojisi dersini alan toplamda 82 gönüllü öğrenci katılmıştır. Çalışmada, madde güçlük indeksleri geniş aralıkta olan ve madde güçlük indeksleri daha önceki uygulamalardan elde edilen, çoktan seçmeli 48 maddeden oluşan Matematik testinin maddelerin güçlük düzeylerine göre sıralandığı üç farklı formu kullanılmıştır. Öğrencilerin test kaygı düzeylerini ölçmek için Uzun Süreli ve Anlık Kaygı Ölçeği ve Başarı Kaygısı Testi kullanılmıştır. Maddelerin seçkisiz olarak sıralandığı form 28, maddelerin güçlük düzeyine göre kolaydan zora doğru sıralandığı form 27, maddelerin güçlük düzeylerine göre zordan kolaya doğru sıralandığı form 27 öğrenciye uygulanmıştır. Kaygı ölçekleri ise hem matematik testi uygulanmadan önce hem de uygulandıktan sonra öğrencilere uygulanmıştır. Kaygı tipinin öğrencilerin Matematik başarıları üzerine etkilerini incelemek amacıyla öğrenciler kolaylaştırıcı, zorlaştırıcı, yüksek etki altında kalanlar, düşük etki altında kalanlar olarak gruplara ayrılmıştır. Çalışmanın analizleri iki yönlü varyans analizi ile yapılmıştır. Yapılan analizler sonucuna göre; öğrencilerin test ortalamalarının, maddelerin kolaydan zora doğru sıralandığı test formunda maddelerin zordan kolaya doğru sıralandığı test formuna göre anlamlı olarak daha yüksek olduğu sonucuna varılmıştır. Ayrıca öğrencilerin kaygı düzeyleri ile madde sıralamalarının etkileşimleri arasında anlamlı bir ortak etki gözlenmemiştir. Çalışmanın bir diğer sonucu da kaygı ölçeğinin matematik testi uygulanmadan önce alınan puanlarının, matematik testi uygulandıktan sonra alınan puanlarından daha düşük olduğudur.

Gerow (1980), test formlarında zor maddelerin ardında yer alan maddelerin, kolay maddelerin ardında yer alan maddelere göre daha sık yanlış yapıp yapılmadığını araştırmıştır. Bu çalışmayı gerçekleştirmek için Indiana Üniversitesi'nde Psikolojiye Giriş dersini alan 71 öğrenciye, 60 çoktan seçmeli maddeden oluşan final sınavının iki farklı formu uygulanmıştır. Testte bulunan maddelerin madde güçlük indeksleri, 5'i zor (bozucu etki yapan), 5'i kolay (motive ediciler) ve 50'si de orta güçlüktedir. Test maddelerinin sadece sıraları değiştirilerek iki farklı formu düzenlenmiştir. Birinci formda, 7., 19., 31., 43. ve 55. maddeler kolay (motive ediciler) ve 1., 13., 25., 37. ve 49. maddeler zor

(bozucu etki yapan), diğer maddeler ise orta güçlükte madde olarak yerleştirilmiştir. İkinci formda ise, kolay ve zor maddelerin yerleri değiştirilip, orta güçlükteki maddelerin yerleri ise her iki formda da aynı kalmıştır. Çalışmanın sonucunda, iki formu alan öğrencilerin test puanları açısından bir farklılık oluşmadığı ve kolay ve zor maddeleri takip eden maddelerde doğru cevapların ortalaması arasında anlamlı bir farklılık olmadığı bulunmuştur.

Plake (1980), testteki maddelerin güçlük düzeylerine göre kolaydan zora, zordan kolayla ve seçkisiz olarak sıralanmasının, öğrencilerin testteki madde sırasına ilişkin bilgi sahibi olup olmasının ve testin anlaşılabilirliği ve güçlüğünü algılayışlarının, öğrencilerin test performansına etkisini incelemiştir. Çalışmaya Nebraska Üniversitesindeki 104 öğrenci katılmıştır. Bu öğrencilere çoktan seçmeli 96 maddeden oluşan ve madde güçlük düzeyleri 0,20 ile 0,96 arasında değişen bir ara sınav uygulanmıştır. Test formları oluşturulurken maddelerin kolaydan zora, zordan kolayla ve seçkisiz olarak sıralandığı formlarda her beş maddeden sonra kolay bir madde eklenmiştir. Ayrıca her test formunda, her maddeden sonra testin güçlüğü için 1-5 arasında puanlanan ‘‘1-çok kolay, 2-kolay, 3-orta, 4-zor, 5-çok zor’’ tepki kategorileri eklenmiştir. Testin anlaşılabilirliği için ise her maddenin sonuna ‘‘oldukça anlaşılır’’, ‘‘anlaşılır’’, ‘‘orta’’, ‘‘anlaşılmıyor’’ ve hiç anlaşılmıyor’’ şeklinde ifadeler eklenmiştir. Araştırmacı, üç farklı formu alan öğrencilerin yarısına, formlardaki maddelerin nasıl sıralandığına ilişkin bilgi verirken diğer yarısına bilgi vermemiştir. Yapılan analiz sonucunda maddelerin kolaydan zora, zordan kolayla ve seçkisiz olarak sıralandığı formlarda madde sıralamaları hakkında bilgilendirilip bilgilendirilmemeye ilişkin test puanlarında, testin algılanan güçlüğünde ve testin anlaşılabilirliğinde anlamlı bir farklılık olmadığı gözlenmiştir.

Plake, Ansorge, Parker ve Lowry (1982), test içindeki maddelerin düzenlenmesi, maddelerin test içerisindeki düzen bilgisi, test kaygısı ve cinsiyetin etkileşimli olarak test performansı üzerindeki etkisini araştırmışlardır. Araştırmada 48 çoktan seçmeli maddeden oluşan Matematik testi kullanılmıştır. Bu testin madde güçlük düzeyleri 0,15 ile 0,84 arasında değişmekte ve maddelerin ortalama güçlük değeri 0,55, madde ayırt edicilik indeksleri ise; 0,15 ile 0,82 arasında değişmekte ve ortalama madde ayırt edicilik indeksi 0,53’tür. 48 maddelik testin maddelerinin kolaydan zora, periyodik olarak ve seçkisiz olarak sıralandığı üç farklı formu oluşturulmuştur. Oluşturulan formların yarısına madde düzenlenmesine ilişkin yönergede bilgiler verilirken diğer yarısına verilmemiştir. Böylece altı form, Nebraska-Lincoln Üniversitesinde İstatistiğe Giriş dersini alan toplam 170 öğrenciye uygulanmıştır. Öğrencilerin test kaygısını ölçmek amacıyla dört farklı kaygı ölçeğinden kaygı ölçümü alınmıştır. Test formlarının sonuna öğrencilerin testin zorluğunu algılama düzeyleri ve onların bekledikleri performansı değerlendirmeleri istenmiştir. Yapılan ANCOVA analizi sonucunda, matematik başarıları ile cinsiyete ilişkilendirildiğinde, erkek öğrencilerin test puanlarının, maddelerin kolaydan zora doğru sıralandığı test formunda daha yüksek olduğu, maddelerin algılanan güçlük düzeyi ile öğrencinin performans



değerlendirmesinin, madde sıralamalarından etkilendiği sonucuna varılmıştır. Madde düzenlemesi hakkında öğrencinin bilgi sahibi olması ise, test performansı ile ilişkili çıkmamıştır.

Plake, Melican ve Shaughnessy (1983), testteki madde sıralamalarının cinsiyet açısından öğrencilerin maddeleri cevaplama olasılıklarında farklılığa neden olup olmadığını araştırmışlardır. Çalışmaya 1982 yılında Midwestern Üniversitesinde Gelişimsel Eğitim Psikolojisi dersini alan, 39 erkek ve 128 kız öğrencilerden oluşan 167 öğrenci katılmıştır. Çalışmada daha önceki dönemlerde pilot uygulaması yapılmış olan sınavlardan seçilen maddelerden oluşan formlar kullanılmıştır. Birinci formda maddelerin önce güçlük düzeylerine göre kolaydan zora, sonra dairesel spiral ve en sonda da seçkisiz; ikinci formda, maddelerin önce seçkisiz, sonra güçlük düzeylerine göre kolaydan zora doğru, ardından dairesel spiral; üçüncü formunda ise, önce dairesel spiral, sonra seçkisiz, en sonda da güçlük düzeylerine göre kolaydan zora doğru sıralandığı farklı üç form kullanılmıştır. Her form çoktan seçmeli 20 maddeden oluşmaktadır. Yapılan varyans analizleri sonucunda cinsiyet ile madde sıralamaları arasında .05 düzeyinde anlamlı bir ilişki bulunamamıştır; fakat .01 düzeyinde anlamlı bir ilişki gözlenmiştir. Üç formda da kız öğrencilerin erkek öğrencilerden daha yüksek test puanına sahip oldukları gözlenmiştir. Diğer etkileşimler arasında anlamlı bir fark gözlenmemiştir. Ayrıca kız öğrencilerin, maddelerin önce kolaydan zora sonra dairesel spiral en sonda da seçkisiz olarak sıralandığı formda en yüksek test puanına sahip oldukları görülmüştür. Erkek öğrencilerin formlardan alınan test puanları arasında anlamlı bir farklılık gözlenmemiştir.

Allison (1984), testteki maddeleri güçlük düzeylerine göre sıralamanın öğrencilerin test performansına etkisini yetenek düzeyi ve cinsiyet ile etkileşimli olarak incelemiştir. Çalışmada 64 çoktan seçmeli maddeden oluşan madde güçlük düzeyleri 0,178 ile 0,981 arasında değişen ve ortalama madde güçlük düzeyi 0,673 olan fen bilimleri testi kullanılmıştır. Teste yer alan maddeler güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz olarak sıralanarak farklı üç form oluşturulmuştur. Öğrencilerin yetenek düzeyleri okul kayıtlarında bulunan IQ puanlarıyla belirlenmiştir. Çalışmaya 160 erkek, 167 kız olmak üzere toplamda 327 öğrenci katılmıştır. Çalışmanın sonucunda üç farklı test formunun ortalamaları arasında anlamlı bir farklılık gözlenmemiştir. Fakat yetenek düzeyi ve cinsiyet etkileşiminin anlamlı olduğu sonucuna varılmıştır. Yüksek yetenek düzeyindeki erkek öğrencilerin test puanlarının daha yüksek olduğu gözlenmiştir. Ayrıca cinsiyet, yetenek düzeyi ve madde sıralamaları arasında anlamlı bir etkileşimin olmadığı görülmüştür.

Laffittee (1984), çalışmasında maddelerin güçlük düzeylerine ve konularına göre sıralanmasının öğrencilerin testin güçlüğünü algılaması ve öğrencilerin test performansı üzerine etkisini araştırmışlardır. Bu amaçla, Winthrop Kolejinde Psikolojiye Giriş dersini alan 82 öğrenciye 50 maddeden oluşan çoktan seçmeli bir testin dört farklı formu uygulanmıştır. Öğrenciler dört farklı gruba ayrılmışlar ve bu gruplara seçkisiz olarak atanmışlardır. Dört gruptan üçü 21 öğrenciden, biri ise, 19 öğrenciden oluşmaktadır. Öğrencilerin toplamda 24'ü erkek, 58'i kız öğrencilerden oluşmaktadır. Bu

öğrencilere üç textbook ünitesi ve dersle ilgili 50 çoktan seçmeli maddeden oluşan bir test uygulanmıştır. Araştırmada kullanılan test dört form şeklinde düzenlenmiştir (Form 1: Her konuyu kapsayan maddelerin kendi arasında kolaydan zora doğru sıralandığı, Form 2: konuların kendi arasında zordan kolaya doğru sıralandığı, Form 3: Her konuyu kapsayan maddelerin kendi arasında seçkisiz olarak sıralandığı, Form 4: konuların kendi arasında seçkisiz olarak sıralandığı form). Dört farklı form her gruba farklı sıralarda uygulanmıştır. Birinci grup sırasıyla 1-2-3-4 formlarını; ikinci grup sırasıyla 2-3-4-1; üçüncü grup sırasıyla 3-4-1-2, dördüncü grup sırasıyla 4-1-2-3 numaralı formları almışlardır. Birinci grupta testin sonuna 51. madde olarak öğrencilerin testin zorluk düzeylerini nasıl algıladıkları (çok kolay, kısmen kolay, ortalama zorluk düzeyinde, kısmen zor, çok zor) eklenmiştir. Çalışmada madde sıralamalarının bir fonksiyonu olan test başarı puanlarının analizleri için; tekrarlı ölçümler için varyans analizi kullanılmıştır. Testin dört farklı formunun test ortalamaları arasında anlamlı bir farklılık gözlenmemiştir. Öğrencilerin testin zorluk düzeyini nasıl algıladıklarına ilişkin cevapları ise Ki kare kullanılarak analiz edilmiştir. Bu analize göre öğrencilerin dört farklı testin zorluk düzeylerini algıları arasında bir farklılık gözlenmemiştir. Bu çalışmaya göre maddelerin zorluk düzeyine göre veya maddelerin konuya göre sıralanması öğrencilerin toplam test puanlarında bir farklılığa yol açmamıştır.

Klimko (1984), çeşitli bilişsel düzeylere ve kaygı düzeylerine sahip Eğitim Psikolojisi dersini alan kız ve erkek öğrencilerin testteki maddelere verdiği tepkilerin, testteki madde sıralanmalarından nasıl etkilendiğini ortaya çıkarmaya yönelik bir araştırma yapmıştır. Öğrencilerin bilişsel giriş düzeyleri Eğitim Psikolojisi dersini almadan önce araştırmacı tarafından geliştirilen 45 çoktan seçmeli maddeden oluşan bir test uygulanarak belirlenmiştir. Çalışmaya 93 kız ve 18 erkek öğrenciden oluşan seçkisiz olarak seçilen toplamda 111 lise öğrencisi katılmıştır. 50 çoktan seçmeli madde, maddelerin madde güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz olarak sıralanarak, üç farklı form şeklinde hazırlanmıştır. Çalışmanın analizleri iki aşamalı çoklu regresyon analizi kullanılarak yapılmıştır. Yapılan analiz sonuçlarına göre sadece öğrencilerin bilişsel giriş düzeyleri ile öğrencilerin test puanları arasında anlamlı bir ilişki bulunmuştur ( $p < .0001$ ). Öğrencilerin test puanları ile maddeleri güçlük düzeylerine göre sıralamanın, cinsiyet ve öğrencilerin kaygı düzeyleri arasında anlamlı bir ilişki bulunmamıştır.

Miller (1989), madde sıralamalarının cinsiyete göre değişen madde fonksiyonu içerip içermediğini araştırmıştır. Çalışmaya ABD'nin 16 eyaletinden Genel Eğitimsel Gelişim (General Educational Development) Deneysel Ünite Testini 1985 yılı Güz döneminde alan 404 kız, 375 erkek öğrenci olmak üzere 779 öğrenci katılmıştır. Çalışmada madde güçlük indeksleri 0,22 ile 0,80 arasında madde ayırt edicilik indeksleri 0,23 ile 0,43 arasında değişen; hesaplama, ölçme, cebir, geometri ve istatistik konularını kapsayan 20 çoktan seçmeli maddeden oluşan Matematik testinin aynı maddelerden oluşan üç farklı formu kullanılmıştır. Form 1: maddelerin güçlük düzeylerine göre kolaydan zora doğru sıralandığı form; Form 2: maddelerin derste işlenen konuya göre sıralandıktan sonra her konu içindeki

maddelerin güçlük düzeylerine göre kolaydan zora doğru sıralandığı form; Form 3: madde güçlük düzeylerine göre kolaydan zora doğru sıralanmış olan her beş maddeden oluşan spiral-dairesel (spiral-cyclical) düzende sıralandığı ve her maddenin kendinden önce gelen maddeden daha zor olduğu form (1-6-11-16-2-7-12-10...).Form 1, 113 kız ve 143 erkek öğrenciye, Form 2, 131 erkek 131 kız öğrenciye, Form 3 ise 130 kız 131 erkek öğrenciye uygulanmıştır. Form 1, Form 2 ve Form 3'ten alınan puanların ortalama, standart sapma ve KR-20 değerleri cinsiyete göre incelendiğinde her bir alt grupta kendi içinde bütün formlarda farklılık gözlenmezken; erkek öğrencilerde bütün madde düzenlemeleri için kız öğrencilerden daha yüksek değerlere sahip olduğu gözlenmiştir. Çalışmanın analizleri iki farklı şekilde yapılmıştır. Testteki kız ve erkek öğrencilerin değişen performanslarını incelemek için her bir formdaki maddelerin kız ve erkek öğrenciler için madde güçlük indeksleri hesaplanmıştır. Bu maddelerin madde güçlük indeksleri sıraya dizildikten sonra her form için cinsiyet ve cinsiyet grupları arasında mümkün olan dokuz tane spearman sıra farkları korelasyon katsayısı hesaplanmıştır. Bunlardan dört tane korelasyon katsayısı beklenen değerinden daha düşük çıkmıştır. Analizin bir diğer kısmı ise bireyin üç formdaki aynı maddeye verdiği cevabın farklılaşıp farklılaşmadığıdır. Kız ve erkek öğrencilerin madde düzeyinde değişen performansı olup olmadığını belirlemek için madde düzenlemenin ve yetenek seviyesinin bağımsız değişken olarak alındığı ki-kare analizi yapılmıştır. Bu analiz sonuçlarına göre erkek öğrencilerde hiçbir maddede değişen madde performansı gözlenmezken kız öğrencilerde üç maddede değişen madde performansı gözlenmiştir. Analiz sonuçları form bazında incelenecek olursa kız ve erkek öğrencilerin performanslarının maddelerin kolaydan zora doğru sıralandığı ve maddelerin spiral bir şekilde düzenlendiği formlardaki hiç bir maddede değişiklik göstermediği belirlenmiştir. Fakat maddelerin önce konuya göre sıralanıp ardından her konu içinde maddelerin kolaydan zora doğru sıralandığı formda beş maddede erkek ve kız öğrenciler arasında değişen madde performansı olduğu gözlenmiştir.

Tippets ve Benson (1989), maddelerin madde güçlüğüne göre kolaydan zora, zordan kolaya ve seçkisiz olarak sıralamanın sınav kaygısı üzerine etkisini araştırmışlardır. Araştırmaya Maryland Üniversitesinde birinci ve ikinci düzey Educational Measurement ve İstatistik dersi alan öğrencilerden birinci düzey dersi alan 28 erkek 50 kız olmak üzere 78 öğrenci, ikinci düzey dersi alan 8 erkek 40 kız olmak üzere 48 öğrenci, toplamda araştırmaya 136 öğrenci katılmıştır. Araştırmacılar birinci düzey ve ikinci düzeyde ders alan öğrenciler için güçlük düzeyleri önceden hazırlanmış çoktan seçmeli test hazırlamışlardır. Madde güçlükleri 0,39 ve 0,97 arasında değişen 44 maddeden oluşan çoktan seçmeli test birinci düzey öğrencileri için hazırlanmıştır. Test maddeleri kolaydan zora, zordan kolaya ve seçkisiz olmak üzere üç form şeklinde hazırlanıp birinci düzey öğrencilerine uygulanmıştır. Maddelerin daha önce uygulanmasından elde edilen madde güçlük indeksleri ile bu uygulamadan elde edilen madde güçlük indeksleri arasındaki korelasyon 0,88 olarak bulunmuştur. Madde güçlükleri 0,20 ile 1,00 arasında değişen çoktan seçmeli 40 maddeden oluşan bir test ikinci düzey öğrencileri için

hazırlanmıştır. Bu testin de maddelerin kolaydan zora, zordan kolay ve seçkisiz olarak sıralandığı üç farklı formu ikinci düzey öğrencilerine uygulanmıştır. Maddelerin daha önce uygulanmasından elde edilen madde güçlük indeksleri ile bu uygulamadan elde edilen madde güçlük indeksleri arasındaki korelasyon 0,83 bulunmuştur. Araştırmacılar sınav kaygısını ölçmek için 20 maddelik Likert maddelerden oluşan Cronbach alfa katsayısı 0,92 olan Sınav Kaygısı Envanteri (Test Anxiety Inventory)'ni test formlarının sonuna ekleyerek uygulamışlardır. İki ayrı düzey için, madde sırasına göre sınav kaygısı puanları arasındaki farklılaşma test kaygı puanlarının bağımlı değişken ve madde sıralamalarının bağımsız değişken olarak ele alındığı tek yönlü ANOVA ile incelenmiştir.

Analiz sonucunda birinci düzey ders alan öğrencilerin maddelerin seçkisiz olarak sıralandığı formda, kolaydan zora doğru sıralandığı forma göre sınav kaygı puanlarının daha yüksek olduğu, maddelerin zordan kolay doğru sıralandığı formdaki sınav kaygı puanları arasında fark olmadığı gözlenmiştir. İkinci düzey dersi alan öğrencilerin sınav kaygı puanlarının ise zordan kolay doğru maddelerin sıralandığı formda, kolaydan zora doğru maddelerin sıralı olduğu forma göre daha yüksek olduğu, seçkisiz sıralamada ise bir fark olmadığı gözlenmiştir. Sonuç olarak, testlerin maddelerinin güçlüklerinin artan düzeyde sıralanması öğrencilerin kaygı puanını azaltırken test puanını artırmıştır.

Aamondt ve Teige (1992), çoktan seçmeli maddelerde seçenek sayısının, test puanlarına ve testi cevaplama süresine; madde güçlüğüne ve konuların derste işleniş sırasına göre madde sıralamanın sınav kaygısına ve test puanına etkisini incelemişlerdir. Bu amaçla, seçenek sayısı ile ilgili önceden yapılmış 8 çalışmadan elde edilen 14 örneklem, madde güçlük düzeyi sırası ile ilgili daha önce yapılan 20 çalışmadan elde edilen 26 örneklem, maddelerin dersin işleniş sırası ile ilgili daha önceden yapılan 4 çalışmadan elde edilen 16 örneklemden oluşan meta analiz çalışması yapılmıştır. Madde güçlük düzeyine göre yapılan analizin sonucunda; öğrencilerin en yüksek test puanını maddelerin güçlük düzeylerine göre kolaydan zora doğru sıralandığı test formundan elde ettiği; maddelerin kolaydan zora doğru sıralandığı 100 maddelik bir testten elde edilen puanın maddelerin seçkisiz olarak sıralandığı formdan elde edilen puanlardan 1,49 puan daha fazla olduğu ve zordan kolay doğru sıralanmış formdan alınan puanlardan 2,97 puan daha fazla olduğu gözlenmiştir. Ayrıca maddelerin zordan kolay doğru sıralanmış test formundan alınan puan ile maddelerin seçkisiz olarak sıralandığı formdan alınan test puanları arasında seçkisiz sıralama lehinde 1,08'lik puan fark oluştuğu gözlenmiştir. Maddelerin güçlük düzeylerine göre sıralanması, maddelerin algılanan güçlük düzeyleri ve sınav kaygısı üzerinde önemli bir farklılaşmaya yol açmamıştır. Maddelerin konuya göre sıralandığı formda testten alınan puan maddelerin seçkisiz olarak sıralandığı testten alınan puanlardan 0,54 puan fazla olduğu gözlenmiştir.

Carlson ve Ostrosky (1992), çoktan seçmeli test maddelerinin test formunda derste işlenen konuya göre ya da karışık olarak sıralanmasının öğrencilerin test performansını etkileyip etkilemediğini araştırmışlardır. Çalışmaya Illinois State Üniversitesinde Ekonominin Temelleri dersi alan 400 öğrenci katılmıştır. Araştırmacılar öğrencilere her sınavın iki test formunun düzenlendiği dört sınav uygulamışlardır. Öğrenciler dört sınavda da farklı iki test formuna seçkisiz olarak atanmışlardır. Sınavlarda kullanılan test formlarının birincisinde (Form A) maddeler derste işlenen konulara göre ve her konu ile ilgili maddelerin güçlük düzeyine göre, ikincisinde (Form B) ise maddeler konuya ve madde güçlük düzeyine göre karışık şekilde sıralanmışlardır. Elde edilen veriler; madde düzenleme etkisi, öğrenci performanslarının ortalama düzeyi, sınavlarda alınan puanların varyansı ve puanların dağılımları, sınavın güvenilirliği açısından analiz edilmiştir. Yapılan analizler sonucunda dört sınav için iki ayrı test formunun ortalamaları arasında manidar bir fark olmadığı, varyansların eşit olduğu gözlenmiştir. Dört sınavın üç'ünde, maddelerin derste işlenen konuya ve madde güçlük düzeyine göre karışık sıralanmış formların, maddelerin derste işleniş sıraya göre sıralandığı formlara göre daha çok sola çarpık olduğu ve daha yüksek basıklık katsayısına sahip olduğu gözlenmiştir. Güvenirlik incelemesinde de testlerin KR-20 katsayıları arasında manidar bir farklılık gözlenmemiştir.

Pang, vd. (1994), DMF 'yi belirlemede kullanılan LR ve MH yöntemlerinin örneklem büyüklüğü ve ölçüt değişkenlerden ne oranda etkilendiğini araştırmışlardır. Çalışma sonucunda Tek biçimli DMF'yi belirleme bakımından iki yöntemin uyumlu olduğu gözlenirken, Tek biçimli olmayan DMF 'yi belirlemede LR yönteminin MH yönteminden daha güçlü olduğu sonucu bulunmuştur. Ayrıca sonuçların her iki yöntemde de örneklem büyüklüğünden etkilendiği ve örneklem arttıkça yöntemlerin daha güvenilir sonuçlar ürettiği sonucuna ulaşılmıştır.

Gierl, Khaliq ve Boughton (1999), Matematik ve Fen başarı testindeki maddelerin DMF içerip içermediğini araştırmıştır. Çalışmaya 12000 öğrenci katılmıştır. Test maddelerinin DMF içerip içermediği MH, LR ve SIBTEST yöntemi ile analiz edilmiştir. Yapılan analizler sonucunda 469 maddeden MH ile 34 maddede, SIBTEST ile 40 maddede, LR yöntemi ile 49 maddede, SIBTEST ile 63 maddede, LR ile 74 maddede DMF gözlenmiştir. Çalışmanın sonucunda yöntemler arasında tam bir uyum olmadığı gözlenmiştir.

Jodoin ve Gierl (1999), lojistik regresyon tekniğinin tek biçimli ve tek biçimli olmayan DMF belirleme performansını istatistiksel güç ve I. Tip hata ölçütlerine göre değerlendirmeyi amaçlamışlardır. Araştırmacılar, örneklem büyüklüğü, yetenek dağılımı ve testteki DMF içeren maddelerin yüzdesini manipüle edilen değişkenler olarak belirlemişlerdir. Araştırma sonucunda, referans grup örneklem büyüklüğü 1000, odak grup örneklem büyüklüğü 250 olduğunda hem tek biçimli hem de tek biçimli olmayan DMF'in belirlenmesinde LR tekniğinin istatistiksel gücünün düşük olduğu gözlenmiştir.

Gomez-benito ve Navas-ara (2000), DMF belirleme yöntemlerinden madde tepki kuramı (IRT) çerçevesinde, MH, LR, logit modeli, restricted faktör analiz metodlarını birbirlerini karşılaştırmışlardır. Çalışmanın sonuçları incelendiğinde, tüm tekniklerin, DMF olmayan maddelerde DMF bulmaya eğilimli olduğu gözlenmiştir. Bu eğilimin en açık LR yönteminde olduğu görülmüştür.

Ryan ve Chiu (2001) tarafından yapılan çalışmanın amacı, madde içerik etkisi (item context effect) ile Değişen Madde Fonksiyonu (DMF) arasındaki ilişkiyi araştırmaktır. Cinsiyete göre DMF bulgularının maddelerin test içindeki yerinden etkilenip etkilenmediği de incelenmiştir. Çalışmada Üniversite birinci sınıf düzeyinde Matematik yerleştirme testi üzerinde cinsiyete göre DMGF analizleri yapılmıştır. Araştırmada Matematik testi için iki farklı test formu oluşturulmuştur. Bu test formlarının her biri 18 cebir, 12 trigonometri, 5 geometri ve 5 analitik geometri sorusu olmak üzere 40 maddeden oluşmaktadır. Birinci test formunda maddeler seçkisiz olarak sıralanırken ikinci formda ise maddeler derste işlenen konuların sırasına ve maddelerin güçlük düzeylerine göre sıralanmıştır. Yapılan SIBTEST analizlerine göre; maddelerin farklı sıralandığı her iki formda da kelime problemlerinin belirgin ölçüde erkek öğrencilerin lehine işlediği görülmüştür. Maddelerin seçkisiz olarak sıralandığı formda geometri maddelerinin, maddelerin derste işlenen sırada ve madde güçlük düzeylerine göre sıralandığı formda ise şekil veya grafik içeren maddelerin, grup olarak erkek öğrencilerin lehine işledikleri görülmüştür.

Ziaee (2007), çalışmasında, lojistik regresyon ve Mantel-Haenszel tekniklerini karşılaştırmayı amaçlamıştır. MH tekniğinin değişen madde fonksiyonu biçimine duyarlı olduğu ve yetenek düzeyi boyunca tek-biçimli olmayan DMF'yi belirlemede etkili olmadığını belirtmiştir. Buna karşın, LR yönteminin, bağımsız değişkenlerin eş zamanlı analiz edilebilmesine izin veren bir yöntem olduğu belirtilmiştir.

Ollenu ve Etsey (2009), tarafından yapılan çalışmanın amacı, madde sıralamalarının öğrencilerin başarı performansını etkileyip etkilemediğini ortaya çıkarmaktır. Çalışmaya on iki farklı okuldan 810 öğrenci katılmıştır. Araştırmada İngilizce, Matematik ve Fen alanlarını kapsayan 40 maddelik çoktan seçmeli test kullanılmıştır. Testin maddelerinin seçkisiz olarak sıralandığı form 82 öğrenciye uygulandıktan sonra maddelerin güçlük düzeyleri belirlenmiştir. Maddelerin güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz olarak sıralandığı üç farklı formu oluşturulmuştur. Böylece matematik, İngilizce ve fen alanlarından oluşan Temel Eğitim Sertifika sınavının her bir alanında madde sıralamalarının etkisi araştırılmaya çalışılmıştır. Çalışmanın analizleri bağımsız değişken olarak farklı üç formun, bağımlı değişken olarak da üç formdan alınan puanların alındığı ANOVA kullanılarak yapılmıştır. Yapılan analizlerin sonucuna göre; İngilizce dersinin konularını kapsayan maddelerde; maddelerin seçkisiz olarak sıralandığı formda öğrencilerin test puanlarının daha yüksek olduğu; maddelerin güçlük düzeylerine göre zordan kolaya ve kolaydan zora sıralandığı formlarda ise test puanlarında anlamlı bir farklılaşma olmadığı gözlenmiştir. Matematik dersinin konularını kapsayan maddelerde; maddelerin güçlük düzeylerine göre kolaydan zora doğru

sıralandığı formda öğrencilerin testten aldıkları puanın en yüksek olduğu; en düşük test puanlarının ise maddelerin güçlük düzeylerine göre zordan kolaya doğru sıralandığı formdan alındığı gözlenmiştir. Fen bilimleri dersinin konularını içeren maddelerde ise öğrencilerin maddelerin güçlük düzeylerine göre zordan kolaya doğru sıralandığı formda en yüksek test puanlarının alındığı; maddelerin seçkisiz ve güçlük düzeylerine göre kolaydan zora doğru sıralandığı formlarda ise anlamlı bir farklılık olmadığı gözlenmiştir.

Gök, Kelecioğlu ve Doğan (2010) çalışmalarında, Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırmak amacıyla, 2005 yılında yapılan Ortaöğretim Kurumları Sınavı (OKS)'nin matematik ve fen bilgisi alt test maddelerinin cinsiyet ve okul türüne göre DMF içerip içermediğini incelemiştir. Araştırmanın sonucunda, MH ve LR teknikleri arasında matematik ve fen bilgisi alt testlerinde cinsiyet ve okul türü değişkenlerine göre genel olarak düşük düzeyde bir uyum görüldüğü rapor edilmiştir.

Ayan (2011), çalışmasında, Uluslararası Öğrenci Değerlendirme Programı (PISA) fen okuryazarlığı alt testinde yer alan maddelerle, cinsiyet değişkenine göre, lojistik regresyon ve Mantel Haenszel tekniklerini kullanarak DMF analizi yapmıştır. DMF analizleri sonucunda LR tekniğinde 1 adet orta düzeyde TBO DMF belirlenirken, MH tekniğinde 4 adet orta düzeyde DMF gösteren madde belirlenmiş ve bu maddelerin 3'ünün kızlar 1'inin erkekler lehine çalıştığı tespit edilmiştir. DMF Analizlerinde kullanılan MH ve LR tekniklerinden elde edilen sonuçlar kıyaslandığında bu iki tekniğin DMF büyüklükleri sıralaması bakımından yakın sonuçlar ürettiği belirlenmiştir. Ayrıca yöntemlerle tespit edilen DMF düzeyleri bakımından düşük düzeyde uyum belirlenirken, belirledikleri maddeler bakımından ise tamamen farklı sonuçlar ürettikleri tespit edilmiştir. Ayrıca DMF elde edilen sonuçlara göre TB DMF belirleme de MH tekniğinin daha duyarlı sonuçlar ürettiği de tespit edilmiştir.

Chiu ve Irwin (2011), çalışmasında testteki maddeleri farklı şekillerde sıralamanın öğrenci performansı üzerine etkisini incelemiştir. Çalışmaya liselerde 9. ve 11. sınıflarda okuyan Tarih ve Devlet Değerlendirmesi (History and Government Assessment) dersini alan toplamda 19.479 öğrenci katılmıştır. Araştırmada tarih, coğrafya ve ekonomi konularını içeren 30 çoktan seçmeli maddeden oluşan test formunun, üç farklı formu kullanılmıştır (Form 1: maddelerin derste işlenen konuya göre standart sıralandığı form; Form 2: maddelerin kronolojik olarak geçmişten günümüze doğru sıralandığı form; Form 3: maddelerin kronolojik olarak günümüzden geçmişe doğru sıralandığı form). Maddelerin bazılarında tarih belirtilip bazılarında belirtilmemiştir. Test formlarına öğrenciler seçkisiz olarak atanmış ve madde sırasının kontrolü için öğrencilere formlar bilgisayar ortamında uygulanmıştır. 6502 öğrenci Form 1'i, 6489 öğrenci Form 2'yi, 6488 öğrenci de Form 3'ü almıştır. Öğrencilerin test formlarındaki her kategorideki (tarih belirtilen ve belirtilmeyen tarih, coğrafya ve ekonomi) maddelerdeki başarılarının madde sıralamasından nasıl etkilendiği üç test formundaki maddelerin güçlük indekslerine ve madde karakteristik fonksiyonlarına bakılarak test edilmiştir. Form 1 ile Form 2 ve Form 2 ile Form 3'ün

ortalama güçlükleri arasında anlamlı bir fark gözlenmemiştir. Form 3'den alınan puanların Form 1'den alınan puanlara göre daha yüksek olduğu belirlenmiştir. Öğrencilerin tarih belirtilen ve belirtilmeyen Tarih, coğrafya ve ekonomi konularını içeren maddelere verdiği doğru cevap sayısında fark olup olmadığı tek yönlü MANOVA kullanılarak karşılaştırılmıştır. Konulara göre ayrılmadan tarih belirtilen tüm (tarih, coğrafya, ekonomi) maddelere verilen doğru cevap sayısı, Form 2'de Form 1 ve Form 3'e göre daha yüksek olduğu ve Form 1 ile Form 3'de bir farklılık gözlenmediği, tarih belirtilen Tarih maddelerinde de aynı şekilde maddelere verilen doğru cevap sayısı, Form 2'de Form 1 ve Form 3'e göre daha yüksek olduğu ve Form 1 ile Form 3'de bir farklılık gözlenmediği sonucuna varılmıştır. Tarih belirtilen coğrafya maddelerinde ise Form 1 ve Form 3'te öğrencilerin Form 2'ye göre daha yüksek puana sahip oldukları, ekonomi maddelerinde ise Form 2'de Form 1 ve Form 3'e göre daha düşük puan aldıkları gözlenmiştir. Konulara göre ayrılmadan tarih belirtilmeyen tüm (tarih, coğrafya, ekonomi) maddelere verilen doğru cevap sayısı ,Form 3'de Form 1 ve Form 2'ye göre daha yüksek olduğu gözlenmiştir. Her madde için madde güçlük indeksi hesaplanıp farklı test formlarındaki değeri ile karşılaştırılmıştır. Formlarda madde karakteristik fonksiyonu, Değişen Madde Fonksiyonun (DMF) varlığı Mantel-Haenszel ki kare tekniği ve Breslow-Day kay kere tekniği (BD) (Breslow & Day, 1980) kullanılarak karşılaştırılmış ve belirlenmeye çalışılmıştır. Yapılan Mantel-Haenszel ki kare analiz sonucuna göre Form 1 ve Form 2 karşılaştırıldığı zaman 9, Form 1 ve Form 3 karşılaştırıldığında 16, Form 2 ve Form 3 karşılaştırıldığında ise 16 maddenin tek biçimli değişen madde fonksiyonu gösterdiği belirlenmiştir. Breslow-Day (BD) kay kere tekniğine göre yapılan analizlere göre Form 1 ve Form 3 karşılaştırıldığı zaman 8, Form 2 ve Form 3 karşılaştırıldığında ise 9 maddenin tek biçimli olmayan değişen madde fonksiyonu gösterdiği belirlenmiştir.

Chiu (2012), çalışmasında testteki madde sıralamalarının öğrencilerin test performansına etkisini incelemiştir. Çalışmaya Kansas Üniversitesinin 2009 bahar döneminde Matematik dersini alan 7.sınıftaki 12.247 öğrenci katılmıştır. Araştırmada farklı üç bölümden oluşan, her bölümdeki madde sayısının 27 ile 29 arasında değiştiği işlemler, cebirsel düşünme, işlemler ve sayılar, veri ve ölçme ve geometri konularını kapsayan 84 çoktan seçmeli maddelerden oluşan iki paralel test kullanılmıştır. Testlerin birinci kısmı; sırasıyla cebir, geometri, sayılar ve işlemler ünitelerini içeren 28 maddeden ve aynı maddelerden oluşan sadece maddelerin yerleri değiştirilerek oluşturulan paralel formu, ikinci kısmı; cebir, sayılar ve işlemler, geometri konularını içeren 29 maddeden ve aynı maddelerden oluşan sadece maddelerin yerleri değiştirilerek oluşturulan paralel formu, üçüncü kısmı; sayılar ve işlemler, geometri ve veri konularını içeren 27 maddeden oluşan ve aynı maddelerden oluşan sadece maddelerin yerleri değiştirilerek oluşturulan paralel formundan oluşmaktadır. Öğrencilerin testlere atanması seçkisiz bir şekilde gerçekleştirilmiştir. Öğrencilerin matematik başarılarının madde sıralamasından nasıl etkilendiği iki test formundaki maddelerin güçlük indekslerine ve madde karakteristik fonksiyonlarına bakılarak test edilmiştir. Her testin ortalama güçlük düzeyi ANOVA yapılarak karşılaştırılmıştır. Her madde için



madde güçlük indeksi hesaplanıp farklı test formlarındaki değeri ile karşılaştırılmıştır. Formlarda Değişen Madde Fonksiyonun varlığı Mantel-Haenszelyöntemi kullanılarak karşılaştırılmış ve belirlenmeye çalışılmıştır. Tek yönlü ANOVA ile yapılan analizler sonucunda, maddenin farklı test formlarındaki madde güçlük indeksleri kıyaslandığında, madde testin başında yer aldığımda, madde güçlük indeksinin de daha büyük olduğu görülmüştür. Yapılan Mantel-Haenszel ki-kare analiz sonucuna göre 34 maddenin tek biçimli değişen madde fonksiyonu gösterdiği belirlenmiştir. Breslow-Day (BD) kikare tekniğine göre yapılan analizlere göre tek biçimli olmayan değişen madde fonksiyonu gösteren 4 madde olduğu görülmüştür.

Louisa (2013), test maddelerinin madde güçlük düzeylerine göre artan zorluk düzeyinde ve seçkisiz olarak sıralanmasının testin ortalama puanlarına, güvenilirlik katsayılarından test-tekrar-test ve KR-20 katsayısı üzerine etkisini araştırmıştır. Çalışmaya Nijeryada Odi, Sagbama ve Yenagoa bölgesel hükümetlerinde bulunan liselerde öğrenim gören random seçilen 264 erkek ve 216 kız öğrenciden oluşan toplamda 480 öğrenci katılmıştır. Çalışmada geçerlik çalışması iki psikometri uzmanı ve dört Matematik öğretmeni tarafından yapılan, çoktan seçmeli beş seçenekli 40 maddeden oluşan, Matematik Başarı Testi kullanılmıştır. Testin maddeleri maddelerinin artan zorluk düzeyinde sıralandığı Form A ve maddelerin seçkisiz olarak sıralandığı Form B olmak üzere iki farklı form şeklinde düzenlenmiştir. Test-tekrar test güvenilirliğini belirleyebilmek için her form için test uygulaması tekrarlanmıştır. Her formun uygulanması 5'er gün aralıklarla olurken, testlerin tekrarları 15'er gün aralıklarla uygulanmıştır. Çalışmanın sonucuna göre maddelerin artan zorluk düzeyine göre sıralandığı Form A ve maddelerin karışık olarak sıralandığı Form B'nin test ortalamaları arasında anlamlı ve maddelerin seçkisiz sıralandığı Form B lehine bir farklılık gözlenmiştir. Form A ve Form B'nin test-tekrar test ve KR-20 güvenilirlik katsayıları arasında anlamlı bir farklılık gözlenmemiştir.

Bulut (2015), maddelerin farklı şekillerde düzenlenmesinin cinsiyet açısından değişen madde fonksiyonu içerip içermediğini araştırmıştır. Bu amaçla, çalışmada ÖSYM tarafından yılda iki kez uygulanan 2010 yılında yapılmış olan ALES sınavının verileri kullanılmıştır. Çalışmada Sözel Alt Yetenek testinin maddelerinin farklı sıralarda sunulduğu A, B, C, D olmak üzere dört farklı soru formu kullanılmıştır. Bu sınava A formuna 2524 erkek, 2476 kız; B formuna 2517 kız, 2483 erkek; C formuna 2516 kız, 2484 erkek; D formuna 2521 kız, 2479 erkek toplamda her formda 5000 öğrenci olmak üzere 20.000 öğrenci katılmıştır. Çalışmanın Değişen Madde Fonksiyonu analizleri için Mantel-Haenszel ve Breslow-Day yöntem kullanılmıştır. Bu yöntemlerle analizler yapılırken üç farklı yol izlenmiştir. Birincisinde, ALES sınavının A, B, C ve D formları birleştirilip tek bir form halinde maddeler sıralanıp bu analizler yapılmıştır. İkincisinde, her bir soru formu için bu analizler ayrı ayrı yapılmıştır. Üçüncüsünde ise, A soru formu referans grup olarak diğer soru formları ise odak grup olarak seçilmiş ve analizler gerçekleştirilmiştir. ALES sınavının A, B, C ve D formlarının birleştirilip tek bir form haline getirilip yapılan DMF analizleri sonucunda 3 tane maddenin ( madde 1 ,madde 4, madde 14) Eğitimsel

Test Hizmetleri ( ETS) kategorisine göre B düzeyinde değişen madde fonksiyonu içerdiği tespit edilmiştir. Madde 1 ve Madde 4'ün kız öğrenciler, madde 14'ün ise erkek öğrenciler lehine işlediği görülmüştür. Mantel-Haenszel ve Breslow-Day metodu kıyaslandığında BD metodu ile 13 maddenin yani testin yaklaşık %16'sının tek biçimli olmayan Değişen Madde Fonksiyonuna sahip olduğu gözlenmiştir. Her iki analizde de ortak olarak 1.maddede Değişen Madde Fonksiyonu olduğu gözlenmiştir. Her bir soru formu için analizlerin ayrı yapıldığı ikinci analiz kısmında ise, yapılan analiz sonucunda A soru formunda 1. ve 3. maddede B düzeyinde erkek öğrenciler lehine, 4. maddede ise C düzeyinde erkek öğrenciler lehine, B soru formunda 2. ve 4.maddelerde C düzeyinde erkek öğrenciler lehine, C soru formunda ise 2. ve 5. maddelerde B düzeyinde erkek öğrenciler lehine; 14. ve 24. maddelerde B düzeyinde kız öğrenciler lehine; D soru formunda ise 3.maddede B düzeyinde erkek öğrenciler lehine, 13.maddede ise B düzeyinde kız öğrenciler lehine DMF gözlenmiştir. En fazla DMF'nin gözlendiği Form ise C formu olduğu görülmüştür. BD'nin tek biçimli olmayan ve MH'nin tek biçimli değişen madde fonksiyonunu belirlemesinden kaynaklı olarak bu metotlarla yapılan DMF analizlerinin sonucu farklı çıkmıştır. A soru formunun referans grup olarak diğer soru formlarının ise odak grup olarak seçildiği analizin üçüncü kısmında; A formunun referans, B formununun odak grup olarak alındığı analiz sonuçlarına göre kız öğrenciler lehine 2 maddenin (madde 1, madde 9) C düzeyinde; 3 maddenin (madde 6, madde 38, madde 51) B düzeyinde, erkek öğrenciler lehine ise 1 maddenin (madde 16) B düzeyinde DMF gösterdiği görülmüştür. A formunun referans, C formununun odak grup olarak alındığı analiz sonuçlarına göre, erkek öğrenciler lehine 2 maddenin (madde 4, madde 10) B düzeyinde, kız öğrenciler lehine 2 maddenin (madde 5, madde 10) B düzeyinde, 1 maddenin (madde 5) C düzeyinde DMF gösterdiği görülmüştür. A formunun referans, D formununun odak grup olarak alındığı analiz sonuçlarına göre, 23.maddenin hem erkek hem kız öğrencilerin lehine B düzeyinde; 51.maddenin hem erkek hem de kız öğrencilerin aleyhine B düzeyinde; 16.maddenin ise kız öğrenciler aleyhine B düzeyinde değişen madde fonksiyonu gösterdiği gözlenmiştir. Sonuç olarak 9. maddenin A formunda B formuna göre erkek öğrenciler için daha kolay olduğu; 51.maddenin hem B formunda hem D formunda A formuna göre erkek öğrenciler için daha kolay olduğu;16.maddenin hem B hem de D formunda kız öğrencilere A formuna göre daha kolay olduğu gözlenmiştir.

#### **I.4.Problem**

Maddelerin güçlük düzeylerine göre kolaydan zora ve zordan kolaya doğru sıralandığı test formunun alınması durumuna göre, maddeler değişen madde fonksiyonu göstermekte midir?

#### I.4. 1.Alt Problemler

1. Mantel-Haenszel yöntemiyle yapılan analizde
  - a. odak grubunun maddelerin güçlük düzeylerine göre kolaydan zora, referans grubunun zordan kolayca doğru sıralandığı,
  - b. odak grubunun maddelerin güçlük düzeylerine göre zordan kolayca, referans grubunun kolaydan zora doğru sıralandığı,

test formunu alması durumuna göre, maddeler, değişen madde fonksiyonu göstermekte midir?

  - c. İki duruma göre Mantel-Haenszel yöntemiyle yapılan analizlerde değişen madde fonksiyonu gösteren maddeler farklılık göstermekte midir?
  
2. Lojistik Regresyon yöntemiyle yapılan analizde
  - a. odak grubunun maddelerin güçlük düzeylerine göre kolaydan zora, referans grubunun zordan kolayca doğru sıralandığı,
  - b. odak grubunun maddelerin güçlük düzeylerine göre zordan kolayca, referans grubunun kolaydan zora doğru sıralandığı,

test formunu alması durumuna göre, maddeler, değişen madde fonksiyonu göstermekte midir?

  - c. İki duruma göre Lojistik Regresyon yöntemiyle yapılan analizlerde, değişen madde fonksiyonu gösteren maddeler farklılık göstermekte midir?
  
3. Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle yapılan analizlerde,
  - a. odak grubunun maddelerin güçlük düzeylerine göre kolaydan zora, referans grubunun zordan kolayca doğru sıralandığı test formunu alması durumuna göre,
  - b. odak grubunun maddelerin güçlük düzeylerine göre zordan kolayca, referans grubunun kolaydan zora doğru sıralandığı test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddeler uyum göstermekte midir?

#### I.5. Amaç

Bu araştırmanın genel amacı, maddelerin güçlük düzeylerine göre test içerisinde farklı sıralarda (kolaydan zora ve zordan kolayca) yerleştirildiği farklı test formlarının verilmesinin, testte yer alan maddelerde DMF oluşturup oluşturmadığının ve kullanılan DMF belirleme yöntemleri arasındaki uyumun belirlenmesidir.

## 1.6. Önem

Ülkemizde geniş çapta yapılan KPSS, ALES, YGS gibi sınavlarda ya da sınıf içi uygulamalarında kullanılan sınavlarda öğrencilerin kopya çekmesinin önüne geçilmesi için maddelerin çeşitli şekillerde sıralandığı testin farklı formları oluşturulmaktadır. Maddelerin test içindeki sıralarının değiştirilmesiyle farklı formların oluşturulması, bu farklı formları cevaplayan bireylerin maddelere verecekleri tepkileri ve bununla birlikte bireylerin test puanlarını farklı biçimde etkileyebilmesi söz konusudur. Özellikle seçme ve yerleştirme gerektiren bireyler hakkında ciddi kararların verildiği böylesine önemli olan sınavlarda, bu durumun bireylerin test puanlarını etkilememesi gerekmektedir. Ülkemizde ve yurtdışında çoktan seçmeli testlerdeki madde sıralamalarının, madde ve test istatistikleri, sınav stresi, sınav kaygısı ve test performansı üzerine etkisi üzerine yapılan çalışmalara rastlanmaktadır. Ancak çoktan seçmeli bir testte yer alan maddelerin güçlük düzeylerine göre sıralanmasının maddelerde Değişen Madde Fonksiyonu (DMF) yaratıp yaratmadığına ilişkin doğrudan bir çalışmaya rastlanmamıştır. Başka bir ifadeyle, DMF kaynağı olarak maddelerin güçlük düzeylerine göre sıralandığı formların ele alındığı bir çalışmaya rastlanmamıştır. Türkiye'deki araştırmalar genel olarak değerlendirildiğinde DMF'yle ilgili çalışmaların ulusal geniş ölçekte yapılan testler üzerinde odaklandığı ve cinsiyet, yaşanan bölge, okul türü ve bölüm değişkenlerine göre incelenmiş olduğu görülmektedir. Bu açıdan, bu çalışmanın alan yazına katkı sunacağı ve bu tarz çalışmalara ve geniş çapta yapılan sınavlara da ışık tutacağı düşünülmektedir.

Madde sıralamalarıyla ilgili yapılan literatür incelemelerinde, bazı çalışmalarda kullanılan testlerde, maddelerin test formlarında düzenlenen sırada çözülmesi gerektiğini belirten yönerge yardımıyla madde sırasının kontrolünün yapıldığı; bazı çalışmalarda ise, her madde farklı sayfaya basılarak madde sırasının kontrolü sağlanmaya çalışıldığı görülmektedir.

Bu çalışmalarda, maddelerin cevaplama sırasına uyulup uyulmadığını tam olarak belirlemek mümkün olmadığından bu durum bir sınırlılık olarak ortaya çıkmaktadır. Bu çalışmada ise, madde sırasının kontrolünün bilgisayarla yapılması, bu sınırlılığın ortadan kaldırılması açısından önemli olduğu düşünülmektedir.

## 1.7. Sayıtlar

- Araştırmaya katılan öğrenciler, test maddelerini okuyup, bilgi düzeylerini yansıtacak ciddiyette ve duyarlılıkta cevaplamışlardır.
- Araştırmaya katılan öğrenciler, testleri cevaplarırken kopya çekme davranışında bulunmamışlardır.

-Araştırma sürecine katkıda bulunan uzmanlar değerlendirmelerini birbirlerinden bağımsız şekilde, içtenlikle ve ciddiyetle yapmışlardır.

-Araştırmaya katılan öğrencilere uygulanan tüm sınavlar aynı sınav koşullarında gerçekleştirilmiştir.

-Araştırmaya katılan öğrencilere uygulanan paralel testler arasında geçen zaman süresince öğrencilerde yeni öğrenmeler gerçekleşmemiştir.

### **1.8. Sınırlılıklar**

1. Araştırmada kullanılan testler, Temel Matematik dersindeki Sayılar Ünitesine ait ‘‘Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler’’ alt öğrenme alanı ile sınırlıdır.

2. Araştırma, uygulamanın yapıldığı Mersin Üniversitesi Erdemli Meslek Yüksek Okulu ve Sosyal Bilimler Meslek Yüksek Okulunda öğrenim görmekte olan 300 öğrenci ile sınırlıdır.

3. Araştırma, çalışma grubunda bulunan öğrencilerin yetenek düzeyleriyle sınırlıdır.

4. Araştırma, Klasik Test Kuramına Dayalı DMF belirleme yöntemlerinden Mantel-Haenszel ve Lojistik Regresyon yöntemleriyle ve hesaplamalarda kullanılan R programıyla sınırlıdır.

## BÖLÜM II: YÖNTEM

Bu bölümde araştırmanın türü, verilerin toplandığı grup, veri toplama araçlarının geliştirilmesi ve uygulanması ve verilerin analizi konusunda açıklamalar yer almaktadır.

### II.1. Araştırmanın Türü

Bu araştırmada; madde güçlük düzeyleri kolaydan zora ve zordan kolayla sıralanarak oluşturulan farklı test formlarına göre maddelerin, çeşitli yöntemlere göre Değişen Madde Fonksiyonu (DMF) içerip içermediğinin belirlenmesi söz konusu olduğundan araştırma temel araştırma olarak değerlendirilebilir. Ayrıca araştırma kapsamında kullanılan yöntemlerin benzerlik ve farklılıkları yönünden bilgi vermesi sebebiyle, araştırma kuramsal bir çalışma özelliği de göstermektedir.

### II.2. Araştırma Grubu

Bu araştırmanın çalışma grubunu, 2015-2016 eğitim-öğretim yılında, Mersin Üniversitesi Erdemli Meslek Yüksekokulu ve Sosyal Bilimler Meslek Yüksekokulu'nda öğrenim gören öğrenciler arasından amaçlı örnekleme yöntemiyle seçilen toplam 300 öğrenci oluşturmaktadır. Araştırma tekrarlı ölçümlere dayanıp, öğrencinin tüm uygulamalara katılmış olması gerektiğinden, yapılan eşleştirme ve veri ön izleme süreçlerinin ardından belirlenen 300 öğrenciyle çalışma tamamlanmıştır. Maddelerin güçlük düzeylerine göre kolaydan zora ve zordan kolayla doğru sıralandığı test formlarına göre DMF analizi yapılırken, matematik dersindeki bilgi ve becerileri açısından yetenek düzeyleri eşitlenmesi amacıyla uygulanan testten alınan puanlara göre yetenek düzeyleri eşitlenen öğrenciler, odak grup (150 öğrenci, %50) ve referans grup (150 öğrenci, %50) olmak üzere ikiye ayrılmış ve çalışma bu iki grubun maddelerin güçlük düzeylerine göre kolaydan zora ve zordan kolayla doğru sıralandığı test formlarına verdikleri cevaplar dikkate alınarak gerçekleştirilmiştir.

### II.3. Veri Toplama Araçları

Araştırmada, Kareköklü Sayılar ve Kareköklü sayılarla işlemler konusunda hazırlanan ikisi paralel toplam üç tane test kullanılmıştır. Paralel testler, maddelerin güçlük düzeylerine göre kolaydan-zora ve zordan-kolayla sıralanarak verilmesi durumunun DMF yaratıp yaratmadığının tespit edilmesinde, diğer test ise öğrencilerin Temel Matematik dersindeki bilgi ve becerileri açısından yetenek düzeylerinin belirlenmesinde kullanılmıştır. Kullanılan testlerin geliştirilme süreciyle ilgili bilgiler aşağıda yer almaktadır.

### II.3. 1. Deneme Uygulaması İçin Testlerin Hazırlanması

Araştırmada ilk olarak Mersin Üniversitesi Meslek Yüksek Okullarında ortak olarak okutulan Temel Matematik Dersinin içeriği, her bir okulda dersi veren öğretim görevlileriyle incelenmiştir. Testin uygulanacağı zaman dilimi de göz önünde bulundurularak, Temel Matematik dersindeki Sayılar Ünitesine ait alt öğrenme alanı olan “Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler” konusuna ait bir tane, öğrencilerin Temel Matematik dersindeki bilgi ve becerileri açısından yetenek düzeylerinin belirlenmesinde kullanılacak testin (Test 1) ve iki tane paralel (Test 2 ve Test 3) ve geliştirilmesine karar verilmiştir. Test geliştirme sürecinde konunun ve kazanımların yer aldığı belirtke tablosu oluşturulmuştur. Belirtke tablosu oluşturulurken, Mersin Üniversitesi web sayfasında yer alan ‘Bilgi Paketi Ders Kataloğu’ sekmesinde Ön lisans Programında yer alan Temel Matematik dersi öğrenim çıktıları incelenmiş ve testlerin uygulanacağı Meslek Yüksek Okullarında Temel Matematik derslerini vermekle sorumlu olan öğretim görevlilerinin ve Matematik Eğitimi alanında çalışan uzmanların görüşüne başvurulmuştur. Sonuçta geliştirilecek test için 8 kazanım belirlenmiş ve her kazanım için madde güçlük düzeyleri kolay, orta ve zor olacak şekilde en az 3 madde yazma çalışması yapılmıştır. Bu şekilde her biri 40 maddeden oluşan üç adet kareköklü sayılar ve kareköklü sayılarla işlemler testi elde edilmiştir.

Paralel testler oluşturulurken paralel olduğu düşünülen maddelere aynı soru sıra numarası verilmiştir. 40 maddeden oluşan testlerin deneme uygulaması yapılmadan önce geliştirilen maddelerin bazı ölçütler bakımından incelenmesi için uzman görüşlerine başvurulmuştur. Uzman görüşleriyle ilgili verilerin alınmasında Ek 1’de belirtilen form kullanılmıştır. Ölçme ve değerlendirme, Matematik Eğitimi alanında çalışan uzmanlardan ve Meslek Yüksek Okullarında Temel Matematik derslerini vermekle sorumlu olan öğretim görevlilerinden oluşan 10 kişilik bir gruptan uzman görüşleri alınmıştır. Uzmanlar arası uyumun hesaplanmasında Fleiss’in Kappa uyum katsayısı hesaplanmıştır. Fleiss (1971), ikiden fazla puanlayıcı arasındaki uyumu, genellenmiş bir Kappa istatistiği ile ortaya koymuştur. Fleiss Kappa istatistiği, ikiden fazla değerlendiricinin uyumunu kategorik ya da sıralı yapıda olan tanı testi sonuçlarını ölçmek amacıyla kullanılır (aktaran: Kanık, Orekeci Temel ve Ersöz Kaya, 2010). Testlerde yer alması düşünülen 40’ar maddelik üç test için Fleiss Kappa uyum katsayıları; öğrencilerin matematik dersindeki bilgi ve becerileri açısından yetenek düzeylerinin belirlenmesi için kullanılan Test 1 için 0,931, paralel olarak tasarlanan Test 2 ve Test 3 için 0,930 olarak hesaplanmıştır.

Fleiss’in (1971) önerisi;  $kappa < 0,40$ : kötü uyum,  $0,40 \leq kapa \leq 0,59$ : zayıf uyum,  $0,60 \leq kapa \leq 0,74$ : iyi uyum ve  $kappa \geq 0,75$  mükemmel uyum şeklindedir (aktaran Yurdugül ve Alsancak, 2013). Bu sonuçlar doğrultusunda testler için uzmanlar arasında mükemmel bir uyumun olduğu

söylenbilir. Ayrıca Test 2 ve Test 3'deki maddelerin paralel olup olmadığıyla ilgili de uzman görüşü alınmıştır. Uzman görüşü ile ilgili verilerin alınmasında Ek 2' deki form kullanılmıştır.

Buna göre Test 2 ile Test 3 arasında 0,947 Fleiss Kappa uyum katsayısı elde edilmiştir. Bu katsayıya göre testlerin paralelliği konusunda uzmanlar arasında mükemmel bir uyumun olduğu söylenebilir.

### **II.3. 2. Testlerin Deneme Uygulamaları**

#### *Yetenek Düzeylerinin Belirlenmesi İçin Hazırlanan Testin Deneme Uygulaması*

Araştırma kapsamında kullanılacak testlerdeki madde sırasını kontrol altına almak amacıyla, nihai testlerin bilgisayar aracılığıyla, Moodle uzaktan eğitim sistemi kullanılarak uygulanması kararlaştırılmıştır. Nihai testlerin uygulanması sırasında bilgisayar laboratuvar ortamında ve test formlarının düzenlendiği Moodle uzaktan eğitim sisteminde yaşanabilecek aksaklıkları öngörebilmek ve gerekli tedbirleri alabilmek amacıyla deneme uygulamaları da bilgisayar kullanılarak gerçekleştirilmiştir. Deneme uygulamaları sürecinde, madde sırası kontrol altına alınmadan, öğrencilerin sınav esnasında istedikleri maddeye tekrardan dönüp cevaplayabilecekleri şekilde test formları düzenlenip öğrencilere bilgisayar ortamında sunulmuştur.

Araştırmanın amacına bağlı olarak odak ve referans gruplarının oluşturulması için, öğrencilerin matematik alanındaki bilgi ve becerileri açısından yetenek düzeylerinin belirleneceği kareköklü sayılar ve kareköklü sayılarla işlemler ünitesine ait kazanımlar göz önünde bulundurularak hazırlanan testin geliştirilmesi sürecinde, 40 maddeden oluşan denemelik test formu, toplamda 365 öğrenciye uygulanmıştır. Deneme uygulamasından elde edilen madde istatistikleri Ek 3' de yer almaktadır. Uygulamadan sonra maddelerin dağılımları, maddelerin güçlük düzeyleri ve madde ayırt edicilikleri incelenerek madde sayısı 20'ye düşürülmüştür. Bu madde eleme sürecinde, testteki her kazanım için en az bir madde olmasına dikkat edilerek ve madde güçlük indeksleri ve madde ayırt edicilik indeksleri göz önünde bulundurularak maddelerin kolay, orta ve zor olarak sınıflandırılmasında her kategoriye yeterli düzeyde temsil edecek şekilde madde seçimine dikkat edilmeye çalışılmıştır. Nihai teste alınan 20 maddenin madde istatistikleri Tablo 5'de belirtilmiştir.



**Tablo 5:** 20 Maddelik Test 1'in Madde İstatistikleri

Madde Numarası(40 Maddelik Testteki Madde Numarası)	Test 1		
	Madde Güçlüğü	Madde Ayırt ediciliği	Madde Standart Sapması
1(10)	0,73	0,71	0,44
2(1)	0,68	0,84	0,46
3(8)	0,66	0,79	0,47
4(23)	0,62	0,77	0,48
5(4)	0,61	0,78	0,49
6(31)	0,57	0,68	0,49
7(26)	0,55	0,75	0,5
8(38)	0,53	0,79	0,5
9(34)	0,52	0,74	0,5
10(37)	0,50	0,7	0,5
11(39)	0,49	0,76	0,5
12(36)	0,48	0,81	0,5
13(32)	0,47	0,79	0,5
14(9)	0,46	0,72	0,5
15(12)	0,44	0,77	0,5
16(5)	0,43	0,59	0,49
17(16)	0,42	0,84	0,49
18(28)	0,41	0,69	0,49
19(33)	0,38	0,76	0,48
20(14)	0,3	0,52	0,46

Uygulamalar ve analizler sonucunda belirlenen 20 maddelik nihai testten elde edilen test istatistiklerine Tablo 6'da yer verilmiştir. Test istatistikleri R-3.2.0 programı kullanılarak elde edilmiştir.

**Tablo 6:** 20 Maddelik Test 1'in Test İstatistikleri

	<b>Test 1</b>
Madde Sayısı	20
Katılımcı Sayısı	365
Ortalama	10,3
Varyans	45,024
Standart Sapma	6,71
Çarpıklık Katsayısı	0,218
Basıklık Katsayısı	1,457
Minimum	0,00
Maksimum	20,00
Medyan	8,00
KR 20	0,941

Tablo 6 incelendiğinde; 365 öğrenciye uygulanan 20 maddelik testten alınan en düşük puanın 0, en yüksek puanın 20 olduğu; ortalama ve ortancanın birbirine yakın değerler aldığı; basıklık ve çarpıklık katsayılarının pozitif değere sahip olmakla birlikte sifıra yakın değerler aldığı ve güvenirliliğin oldukça yüksek olduğu görülmektedir. Nihai 20 maddelik Test 1, Ek 4' de yer almaktadır.

#### *Paralel Testlerin Deneme Uygulaması*

Paralel testlerin geliştirilme sürecinde 40 maddelik iki test bir hafta arayla öğrencilere uygulanmıştır. İlk teste 287 öğrenci ve ikinci teste 175 öğrenci katılmıştır. Hem birinci hem de ikinci teste katılan öğrencilerin belirlenmesinde R 3.2.0 programı kullanılarak eşleştirme işlemi gerçekleştirilmiştir. Yapılan eşleştirme işlemi sonucunda 167 öğrencinin hem birinci hem de ikinci teste katıldıkları görülmüştür. Deneme uygulaması sonucunda, her iki testte yer alan maddelerin madde istatistikleri elde edilmiştir. Elde edilen madde istatistikleri Ek 5'de belirtilmiştir. Zaman sınırlılığı da göz önünde bulundurularak, asıl uygulamada kullanılacak olan paralel testlerde 20 maddenin yer almasının uygun olacağı düşünülmüştür. Testlerde yer alan madde sayısı 40'tan 20'ye düşürülürken testteki her kazanım için en az bir madde olacak şekilde maddelerin dağılımları, maddelerin standart sapmaları, madde ayırt edicilik ve madde güçlük indeksleri iki test için de karşılaştırılmıştır. Bu aşamada, bu madde istatistiklerinin uygun ve madde istatistikleri arasındaki farkın en az olduğu ve madde güçlük düzeylerini kolay, orta ve zor olarak en iyi temsil edeceği düşünülen maddelerin seçilmesine özen gösterilmiştir. Madde eleme işlemi sonucunda kareköklü sayılar ve kareköklü sayılarla işlemler konusunda iki adet paralel test elde edilmiştir. Madde istatistiklerinin belirlenmesinde Excel ve R-3.2.0 programı kullanılmıştır. Madde seçim işleminden sonra tekrar madde analizi işlemleri

gerçekleştirilmiştir. 20 maddeden oluşan nihai paralel testlerde yer alan maddelerin madde istatistikleri Tablo 7’ de görülmektedir.

**Tablo 7:** 20 Maddelik Test 2 ve Test 3’ün Madde İstatistikleri

Madde No (40 Maddelik Testteki Madde Numarası)	Test 1			Test 2			
	Madde Güçlüğü	Madde Ayırt Ediciliği	Madde Standart Sapması	Madde No(40 Maddelik Testteki Madde Numarası)	Madde Güçlüğü	Madde Ayırt Ediciliği	Madde Standart Sapması
<b>1(8)</b>	0,77	0,46	0,42	<b>20(8)</b>	0,75	0,77	0,43
<b>2(10)</b>	0,71	0,54	0,45	<b>19(10)</b>	0,71	0,86	0,45
<b>3(23)</b>	0,7	0,76	0,46	<b>18(23)</b>	0,7	0,86	0,46
<b>4(3)</b>	0,67	0,66	0,47	<b>17(3)</b>	0,65	0,66	0,48
<b>5(1)</b>	0,66	0,67	0,48	<b>16(1)</b>	0,6	0,79	0,49
<b>6(38)</b>	0,6	0,78	0,49	<b>15(38)</b>	0,56	0,79	0,5
<b>7(40)</b>	0,58	0,72	0,49	<b>14(40)</b>	0,55	0,78	0,5
<b>8(11)</b>	0,57	0,73	0,49	<b>13(11)</b>	0,56	0,81	0,5
<b>9(30)</b>	0,56	0,77	0,5	<b>12(30)</b>	0,55	0,78	0,5
<b>10(5)</b>	0,54	0,7	0,5	<b>11(5)</b>	0,55	0,71	0,5
<b>11(39)</b>	0,53	0,78	0,5	<b>10(39)</b>	0,51	0,8	0,5
<b>12(12)</b>	0,51	0,57	0,5	<b>9(12)</b>	0,5	0,63	0,5
<b>13(27)</b>	0,47	0,64	0,5	<b>8(27)</b>	0,49	0,77	0,5
<b>14(6)</b>	0,45	0,75	0,5	<b>7(6)</b>	0,48	0,75	0,5
<b>15(34)</b>	0,41	0,78	0,49	<b>6(34)</b>	0,47	0,84	0,5
<b>16(33)</b>	0,4	0,74	0,49	<b>5(33)</b>	0,45	0,73	0,5
<b>17(7)</b>	0,38	0,77	0,49	<b>4(7)</b>	0,39	0,72	0,49
<b>18(19)</b>	0,34	0,74	0,48	<b>3(19)</b>	0,37	0,71	0,48
<b>19(18)</b>	0,32	0,56	0,47	<b>2(18)</b>	0,34	0,58	0,47
<b>20(14)</b>	0,21	0,64	0,41	<b>1(14)</b>	0,29	0,53	0,46

Bu çalışmada, kolaydan zora tüm güçlük düzeylerini kapsayacak ve temsil edebilecek madde güçlük düzeyleri olması istendiğinden, madde güçlük düzeyleri 0,20 - 0,39 arasındaki maddeler zor, 0,40 - 0,59 arasındaki maddeler orta, 0,60 - 0,79 arasındaki maddeler kolay olarak adlandırılmış ve değerlendirilmiştir. Tablo 7 incelendiğinde, nihai test formlarında madde güçlük düzeyleri 0,60 – 0,77 aralığında değişen; Test 2’de 6, Test 3’de 5 tane kolay madde, 0,40 – 0,58 aralığında değişen; Test 2’de 10 tane, Test 3’de 11 tane orta güçlükte madde, 0,21 -0,39 aralığında değişen; Test 2’ de 4, Test 3’ de 4 tane zor madde bulunmaktadır. Uygulamalar ve analizler sonucunda belirlenen 20 maddelik nihai testlerden elde edilen test istatistiklerine Tablo 8’de yer verilmiştir. Test istatistikleri R-3.2.0 programı kullanılarak elde edilmiştir.

**Tablo 8:** 20 Maddelik Test 2 ve Test 3'ün Test İstatistikleri

	Test No	
	2	3
Madde Sayısı	20	20
Katılımcı Sayısı	167	167
Ortalama	10,54	10,73
Varyans	40,96	46,92
Standart Sapma	6,4	6,85
Çarpıklık Katsayısı	1,545	1,378
Basıklık Katsayısı	0,0349	0,020
Minimum	0,00	0,00
Maksimum	20,00	20,00
Medyan	11,00	10,00
KR 20	0,933	0,945

Tablo 8 incelendiğinde, 167 öğrenciye uygulanan 20 maddelik her iki testten alınan en düşük puanın 0, en yüksek puanın 20 olduğu; ortalama ve ortancanın birbirine yakın değerler aldığı; basıklık ve çarpıklık katsayılarının pozitif değere sahip olmakla birlikte sifıra yakın değerler aldığı ve güvenilirliklerinin oldukça yüksek olduğu görülmektedir. Ayrıca, testlerin birbirleriyle korelasyonu 0,941 olarak hesaplanmıştır. Madde ve test istatistikleri incelendiğinde testlerin paralel olduğu kabul edilebilir.

Maddelerin istatistiksel paralelliğinin yanında kapsam olarak da paralel olup olmadığını denetlemek için uzman görüşü alınmıştır. Ölçme ve değerlendirme uzmanları ve Matematik Eğitimi alanında çalışan uzmanlardan oluşan 8 kişilik bir gruptan maddelerin paralel olup olmadığı ile ilgili görüş alınmıştır. Nihai testte yer alması düşünülen 20 madde için uzman görüşü alınarak Fleiss Kappa uyum katsayısı hesaplanmıştır. Buna göre Test 2 ile Test 3 arasında 0,908 Fleiss Kappa uyum katsayısı elde edilmiştir. Sonuç olarak, testlerin paralelliği konusunda uzmanların mükemmel bir uyumunun olduğu söylenebilir.

Testlerin kapsam geçerliğini ortaya çıkarmak amacıyla nihai teste seçilen maddelerin bazı ölçütler bakımından incelenmesi için uzman görüşlerine başvurulmuştur. Her test için Fleiss Kappa Uyum katsayısı hesaplanmıştır. Test 2 ve Test 3 için 0,869 Fleiss Kappa uyum katsayısı elde edilmiştir. Böylece uzmanlar arasında elde edilen mükemmel uyum testlerin kapsam geçerliğinin sağlandığının göstergesi olarak kabul edilmiştir. Her iki teste ait KR -20 güvenilirlik katsayılarının yüksek olması da (0,933 ve 0,945) testlerin yapı geçerliğine kanıt olarak kabul edilebilir ( Erkuş, 2003; Kan, 2006).Nihai 20 maddelik Test 2 ve Test 3, Ek 6 ve Ek 7' de yer almaktadır.

### II.3. 3. Asıl Uygulama İçin Testlerin Hazırlanması ( Moodle Sınav Yönetimi)

Testlerin güvenilirlik ve geçerlik çalışmaları tamamlandıktan sonra test maddeleriyle üç ayrı test formu oluşturulmuştur. Araştırmanın amacına bağlı olarak; öğrencilerin Temel Matematik dersindeki bilgi ve becerileri açısından yetenek düzeylerinin belirleneceği Test 1’de yer alan maddeler, alanyazında önerilen şekilde, güçlük düzeyleri kolaydan zora doğru, Test 2’de yer alan maddeler, güçlük düzeylerine göre kolaydan zora doğru, Test 3’de yer alan maddeler ise; güçlük düzeylerine göre zordan kolaya doğru sıralanıp nihai test formları oluşturulmuştur. Hazırlanan Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler testlerinin uygulanmasında, araştırmanın amacına bağlı olarak öğrencilerin maddeleri, testte yer alan sıraya göre cevaplamaları gerekmektedir. Madde sıralamalarıyla ilgili bireylerin maddeleri testte yer alan sıraya göre cevapladıklarından emin olmak için testlerin bilgisayar aracılığıyla uygulanması kararlaştırılmıştır. Bu amaçla, Moodle açık kaynak kodlu uzaktan eğitim sistemi kullanılmıştır. Moodle uzaktan eğitim sistemi ile sınav yönetiminin nasıl yapıldığına dair işlem basamakları aşağıda kısaca belirtilmiştir:

Uzaktan eğitim sistemi ile sınav yönetimi için bilgisayara Moodle programı kurulmadan önce, IP adresi üzerinden internet yayını yapmak amacıyla sunucu ve bir paket kurulumu olan Wamp Server, sunucu olarak belirlenen bilgisayara kurulmuştur. Moodle kurulumu, kullanıcı ve veri tabanı oluşturulduktan sonra yönetici (admin) olarak sisteme girilip sınavların uygulanmasına yönelik genel ayarlamalar yapılmıştır. Sistemde üç test için üç farklı sınav oluşturulup sınav ayarları her test için ayrı olacak şekilde düzenlenmiştir. Her sınava, testlerdeki madde sıraları gözetilerek maddeler eklenmiş ve her maddenin yeni bir sayfada yer alması sağlanmıştır. Test 2 ve Test 3 düzenlenirken öğrencinin bir maddeye tekrardan dönmeye engellenmiş böylece araştırmanın amacına bağlı olarak öğrencinin testteki maddeleri istenen sırada çözmesi sağlanmıştır. Öğrencilerin matematik dersindeki bilgi ve becerileri açısından yetenek düzeylerinin belirleneceği Test 1 ise, sınav devam ederken öğrencilerin istediği maddeye tekrardan dönmelerine izin verilecek şekilde düzenlenmiştir. Ayrıca uygulamaya katılacak öğrencilere her bir test formu için bir kez uygulamaya katılma hakkı tanınmış, sınavların açılacağı ve kapanacağı tarih, saat ve sınav süreleri ayarlanmıştır. Sınavların genel ayarları yapıldıktan sonra uygulamaya katılacak öğrencilerin ad-soyad, öğrenci numaraları ve e-mail adres bilgileri öğrenci işlerinden temin edilmiştir. Bu bilgiler her öğrenciye bir kullanıcı adı ve şifre tanımlanacak şekilde düzenlenmiş ve tüm öğrenciler sisteme kullanıcı olarak eklenmiş ve kaydedilmiştir. Kullanıcı adı ve şifreyle sisteme giriş yapan öğrencinin karşısına düzenlenen sınav gelmekte ve hazırlanan sınav önce yönerge sayfasıyla başlamaktadır. Yönergede testin amacı özetlenmiş, testin süresi ve testteki madde sayısı belirtilerek testin nasıl uygulanacağı, yönerge sayfasından sonra hangi sayfaların geleceği bu sayfalardaki butonların işlevi ve nasıl kullanılacağı anlatılmıştır. Yönerge sayfasının ekran görüntüsü Ek 8’de verilmiştir. Yönerge sayfasındaki ‘‘Sınavı şimdi uygula’’ butonuna basıldığında uygulama

başlamakta ve testin maddeleri her sayfada bir madde olacak şekilde ekrana gelmektedir. Şekil 3’de 2.Test maddelerinden seçilmiş bir maddeye ilişkin ekran görüntüsü bulunmaktadır.

**Temel Matematik 1**

Benim sayfam ▶ TM-1

**SINAV GEZİNTİSİ**

1 2 3 4 5 6 7 8 9 10  
11 12 13 14 15 16 17 18 19 20

Uygulamayı bitir ...

Kalan Süre 0:20:51

**Soru 17**

Henüz cevaplanmadı

5,00 üzerinden işaretlenmiş

▼ Soruyu işaretle

$\sqrt{0,12}$  ifadesi aşağıdakilerden hangisine eşittir?

Lütfen birini seçin:

A.  $\frac{\sqrt{3}}{5}$

B.  $\frac{2\sqrt{3}}{5}$

C.  $\frac{3\sqrt{3}}{5}$

D.  $\frac{4\sqrt{3}}{5}$

E.  $\frac{6\sqrt{3}}{5}$

Sonraki

**Şekil 3.** Test 2 ’de yer alan 17. maddenin ekran görüntüsü

Şekil 3’de görüldüğü gibi maddeyi cevaplayan öğrenci ”Sonraki” butonuna basarak diğer maddeye geçebilmektedir.”Sınav Gezintisi” kısmında öğrencinin hangi maddeyi cevaplayıp hangi maddeyi boş bıraktığı ve sınav için kalan süresi hakkında bilgi edinilebilmektedir. Tüm maddeleri cevaplayan öğrenci “Uygulamayı bitir” butonuna basarak testi bitirmiş olacaktır.

### II.3. 4.Uygulama

Uygulama yapılmadan önce bilgisayar laboratuvar ortamında gerekli düzenlemeler yapılmıştır. Uygulama esnasında herhangi bir aksaklıkla karşılaşılmaması için bilgisayarlardaki internet bağlantısının mevcut olup olmadığı, internet tarayıcılarının sürümünün güncel olup olmadığı kontrol edilmiştir. Sunucu olarak kullanılacak bilgisayarda WampServer açılarak, moodle uzaktan eğitim sistemiyle sınav yönetimi hazır hale getirilmiştir. Uygulamanın yapılacağı laboratuvar ortamında bulunan bilgisayar sayısı kadar öğrenci, bilgisayar başına oturtulduktan sonra öğrencilere yapılacak uygulamayla ilgili kısa bir açıklama yapılmış ve uygulama başlayana kadar bilgisayarlarında herhangi bir işlem yapmamaları gerektiği konusunda uyarılmışlardır. Laboratuvar ortamındaki projektör yardımıyla öğrencilere sınav uygulamasının nasıl yapılacağı, beş tane farklı örnek maddeden oluşan sınav üzerinden uygulamalı olarak anlatılmıştır. Yönetici (admin) olarak giriş yapılan ana sayfada, kullanıcı adı ve şifre için, her öğrencinin kendisine ait olan öğrenci numaralarını kullanmaları ve daha sonra, sisteme giriş yapmaları söylenmiştir. Ardından perdeye yansıtılan yönerge, tüm öğrencilere okunmuştur.

Bu işlem tamamlandıktan sonra öğrencilere ekranda görülen ‘‘Sınavı şimdi uygula’’ butonuna basmaları söylenmiştir.

Çıkan sayfada örnek ilk madde perdeye yansıtılarak arařtırmacı tarafından okunmuş ve cevaplamamanın nasıl yapılacağı uygulamalı olarak öğrencilere gösterilmiştir. Yapılacak uygulamanın tüm öğrenciler tarafından anlaşılması sağlandıktan sonra, test uygulaması başlatılmıştır. Her test formunda yer alan 20 maddeyi cevaplamaları için öğrencilere 30 dakika süre verilmiş ve işlem becerisi gerektiren maddelerde kullanmaları için sınav başlamadan her öğrenciye kağıt ve kalem temin edilmiştir. Uygulamaya katılan öğrenciler kullanıcı adı ve şifresi olan öğrenci numaraları ile sisteme giriş yaptığında cevaplayacakları test formu ile karşılaşmıştır. Testlerin uygulanma şekline ait desen Tablo 9’ da gösterilmektedir.

**Tablo 9:** Testlerin Uygulanma Şekline Ait Desen

Grup	Test 1	Önce	Sonra
Odak	✓	Test 2	Test 3
Referans	✓	Test 3	Test 2

Odak Grubu: Önce Test 2’yi sonra Test 3’ü alan öğrenciler

Referans Grubu: Önce Test 3’ü sonra Test 2’yi alan öğrenciler

Test 1: Öğrencilerin yetenek düzeylerinin belirlendiği test formu

Test 2: Maddelerin güçlük düzeyine göre kolaydan zora doğru sıralandığı test formu, KZ test formu

Test 3: Maddelerin güçlük düzeyine göre zordan kolayca doğru sıralandığı test formu, ZK test formu

Sınava başlamadan önce arařtırmanın amacına baėlı olarak aynı anda uygulamaya katılacak öğrencilerin, Test 1’den aldıkları test puanlarına göre matematik dersindeki bilgi ve beceri açısından yetenek düzeyleri eşitlendikten sonra, odak ve referans grubu olmak üzere iki gruba ayrılmıştır. Odak grubunun; Test 2’den, referans grubunun ise; Test 3’den başlaması sağlanarak testler dengeli olarak uygulanmıştır. Bir hafta arayla, önceki uygulamada Test 2’yi alan odak grubundaki öğrencilere Test 3, Test 3’ü alan referans grubundaki öğrencilere Test 2 uygulanmıştır. Böylece uygulamaya katılan öğrencilerin üç test formunu da alması sağlanmıştır. Karşıt dengelenmiş desenin kullanılmasındaki amaç, testlerdeki sıra etkisini ortadan kaldırmaktır Tüm öğrenciler testlerini tamamlayıp laboratuardan çıktıktan sonra moodle uzaktan eğitim sistemine yönetici (admin) girişi yapılarak her öğrencinin her maddeye verdiği cevabının gösterildiği veriler Excel dosyası şeklinde elde edilmiştir.

## II.4.Verilerin Analizi

Uygulamadan elde edilen verilerin analizinde aşağıdaki çalışmalar yapılmıştır.

### I.Test maddelerine verilen cevapların puanlanması

Bütün test formlarında (Test 1,Test 2 ve Test 3) yer alan maddeler doğru cevaplar için '1' yanlış cevaplar için '0' şeklinde puanlanmış ve testler için her bir öğrencinin test puanı hesaplanmıştır.

### II. Testlerden elde edilen verilerin tek boyutluluğunun sınanması

Dorans ve Holland (1982), DMF analizleri öncesinde eşleme değişkeninin (toplam test puanı) geçerliliğine yönelik kanıtların sağlanması gerektiğini ve eşleştirme değişkenine ilişkin (toplam puan) tek boyutluluk analizinin, DMF belirleme sürecinin esası olduğunu belirtmektedir. Bu nedenle çalışma kapsamında DMF analizlerinin temel gereksinimi olarak tek boyutluluk sayılısı kontrol edilmiştir

Araştırmada, verilerin tek boyutluluk sayılısını karşılayıp karşılamadığını incelemek için Dimpack 1. 0 programı kullanılarak, DIMTEST T istatistiği hesaplanmıştır. Testlerin tek boyutluluğuna ilişkin yapılan analiz sonuçları Tablo 10' da gösterilmektedir.

**Tablo 10:** Testlere İlişkin Tek Boyutluluk Analiz Sonuçları

	DIMTEST			
	TL	TGbar	T	P
Test 1	3.3870	1.9888	1.3912	0.0821
Test 2	3.6996	2.3030	1.3897	0.0823
Test 3	3.2432	2.0061	1.2309	0.1092

Tablo 10 incelendiğinde, üç testin de p değerleri 0,05'ten büyük olduğu için testler tek boyutlu olarak kabul edilmiştir ( Test 1 ( $p=0,0821>0,05$ ), Test 2 (  $p=0,0823>0,05$ ), Test 3 ( $p=0,1092>0,05$ )).

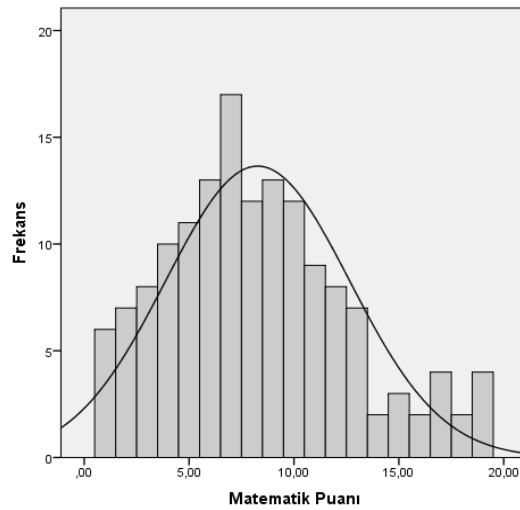


## II. Test puanları dağılımına ilişkin betimsel istatistiklerin hesaplanması

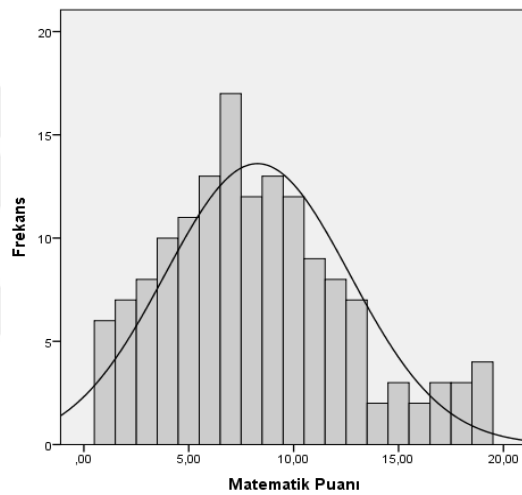
Alt problemlere ilişkin istatistiksel çalışmalara başlamadan önce, çalışmaya katılan bireylerin, Test 1'deki test puanları göz önünde bulundurularak matematik dersindeki bilgi ve becerileri açısından yetenek düzeyleri eşitlenerek odak ve referans gruplarında yer alacak bireyler belirlenmiştir. Odak ve referans gruplarında yer alan öğrencilere uygulanan Test 1'e ilişkin betimsel istatistikler hesaplanmıştır. Elde edilen sonuçlar Tablo 11'de gösterilmiştir. Ayrıca iki gruba ait dağılımların grafikleri çizilmiş, Şekil 4 ve Şekil 5'de gösterilmiştir.

**Tablo 11:** Odak ve Referans Grubuna Uygulanan Test 1'e İlişkin Betimsel İstatistikler

	Grup	
	Odak	Referans
Öğrenci Sayısı	150	150
Ortalama	8,266	8,273
Ortanca	8,000	8,000
Mod	7,00	7,00
Standart Sapma	4,38	4,39
Varyans	19,217	19,341
Basıklık Katsayısı	-0,146	-0,127
Çarpıklık Katsayısı	0,528	0,539
Minimum	1,00	1,00
Maksimum	19,00	19,00



**Şekil 4.** Odak Grubunda Yer Alan Öğrencilerin Matematik Puanı Dağılım Grafiği



**Şekil 5.** Referans Grubunda Yer Alan Öğrencilerin Matematik Puanı Dağılım Grafiği

Tablo 11 incelendiğinde, odak ve referans gruplarındaki öğrencilere Test 1 uygulandığında, her iki grup üzerinden elde edilen istatistiklerin birbirine yakın olduğu görülmektedir. Bu durumda, formların uygulandığı grupların benzer özellikler gösterdiği ve grup ortalamaları ve homojenlikleriyle ilgili benzer istatistiksel değerlere ulaşıldığı söylenebilir. Çarpıklık ve basıklık değerleri incelendiğinde normal dağılıma göre küçük sapmalar gösterdiği görülmektedir.

Odak ve referans grubundaki öğrencilere Test 1 uygulandığında, gruplarda yer alan öğrencilerin ortalamaları arasında anlamlı bir farklılık olup olmadığını belirlemek için yapılan Mann Whitney U testi sonuçları Tablo 12’de gösterilmiştir.

**Tablo 12:** Test 3'ün Uygulandığı Odak Ve Referans Gruplarına Yapılan Mann Whitney U Testi Sonuçları

Grup	n	Sıra Ortalaması	Sıra Toplamı	U	p	z
Odak	150	150.48	22572.00	11247.00	0.997	-0,004
Referans	150	150.42	22578.00			

Tablo 12 incelendiğinde, odak ve referans grubundaki öğrencilere Test 1 uygulandığında, bu grupların sıra farkı ortalamaları arasında 0,05 anlamlılık düzeyinde istatistiksel açıdan anlamlı bir farklılığın olmadığı görülmektedir ( $Z = -0,004$ ,  $p > 0,05$ ).

### III. Alt problemlere İlişkin Analizlerin Gerçekleştirilmesi

Birinci alt probleme ilişkin bulguların elde edilmesi aşamasında; odak grubunun KZ, referans grubunun ZK test formunu alması ve odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre iki MH analizi yapılmıştır. İlk MH analizinde veriler düzenlenirken, odak grubundaki 150 öğrencinin Test 2'deki maddelere verdiği cevaplar ile referans grubundaki 150 öğrencinin Test 3'deki maddelere verdiği cevaplar birleştirilmiştir. Verilerin birleştirilmesi sürecinde, Test 2 ve Test 3'deki paralel maddeler gözetilerek, paralel maddelere verilen cevaplar Excel dosyasında aynı sütunlarda yer alacak şekilde düzenlenmiştir. İkinci MH analizinde veriler düzenlenirken, odak grubundaki 150 öğrencinin Test 3'deki maddelere verdiği cevaplar ile referans grubundaki 150 öğrencinin Test 2'deki maddelere verdiği cevaplar birleştirilmiştir. Verilerin birleştirilme sürecinde, Test 2 ve Test 3 'deki paralel maddeler gözetilerek, paralel maddelere verilen cevaplar Excel dosyasında aynı sütunlarda yer alacak şekilde düzenlenmiştir. Her iki duruma göre yapılan analizler sonucunda, DMF gösteren maddeler, DMF düzeyleri ve bu maddelerin hangi grubun lehine işledikleri tespit edilmiş ve sonuçlar tablolastırılmıştır. MH istatistiği sonucunda test maddeleri içerdikleri DMF oranına göre A, B ve C düzeyi olmak üzere üç düzeyde ele alınır. Bu çalışmada, B ve C düzeyinde olan maddelerde DMF olduğu kabul edilmiştir (Camilli ve Shepard, 1994; Osterlind, 1983). Elde edilen sonuçlara göre iki analiz için DMF gösteren maddeler, avantajlı gruplar ve  $\Delta_{MH}$  değerleri ile bu değerlerin düzeyleri tespit edilmiş ve tablolar halinde düzenlenmiştir. MH yöntemi istatistiklerini elde etmek için R-3.2.0 programından faydalanılmıştır. İlk MH ve ikinci MH analizi sonucu değişen madde fonksiyonu gösteren maddeler, madde sayısı ve bunların düzeyleri açısından karşılaştırılmıştır. Birinci ve ikinci MH analiz sonuçları Ek 9 ve Ek 10 'da verilmiştir.

İkinci alt probleme ait bulguların elde edilmesi aşamasında, odak grubunun KZ, referans grubunun ZK test formunu alması ve odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre iki LR analizi yapılmıştır. İlk LR analizinde, ilk MH analizinde kullanılan; ikinci LR analizinde ise, ikinci MH analizinde kullanılan veriler analiz edilmiştir. LR analizi yönteminde, Tek Biçimli ve Tek Biçimli olmayan DMF'yi belirlemek üzere iki farklı analiz sonucu elde edilmiştir. Her iki duruma göre yapılan analizler sonucunda, Tek Biçimli ve Tek Biçimli Olmayan DMF'yi belirlemek için yapılan LR Analizlerine ait  $\beta$ ,  $p$  ve  $\Delta R^2$  değerleri elde edilmiştir. Her analiz için DMF gösterdiği belirlenen maddelerdeki DMF'lerin, etki büyüklüklerini ve önemlilik düzeylerini belirleyebilmek için, Jodoign ve Gierl'in (2001) önermiş oldukları sınıflama kategorilerine göre, DMF'lerin hangi düzeyde (A, B ya da C) olduğu belirlenip, B ve C düzeyinde olan maddelerde DMF olduğu kabul edilmiştir. DMF gösterdiği tespit edilen maddelerin hangi gruba avantaj sağladığı ilişkisiz örneklem için t testi kullanılarak belirlenmiştir. Elde edilen sonuçlara göre DMF gösteren maddeler, bu maddelerin DMF biçimleri, avantajlı gruplar ve  $\Delta R^2$  değerleri ile bu değerlerin düzeyleri tespit edilmiş ve tablolar halinde düzenlenmiştir. İlk LR ve ikinci LR analizi sonucu değişen madde fonksiyonu gösteren maddeler, madde sayısı ve bunların düzeyleri açısından karşılaştırılmıştır. Birinci LR analizinde, Tek biçimli ve Tek biçimli olmayan DMF'yi belirlemek için yapılan analizler Ek 11 ve Ek 13' de, ikinci LR analizinde, Tek biçimli ve tek biçimli olmayan DMF'yi belirlemek için yapılan analizler Ek 12 ve Ek 14'de verilmiştir.

Üçüncü alt probleme ait bulguların elde edilmesi aşamasında, odak grubunun maddelerin KZ, referans grubunun ZK test formunu alması ve odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, Mantel-Haenszel ve Lojistik Regresyon yöntemiyle yapılan analizlerde, değişen madde fonksiyonu ile ilgili elde edilen sonuçların uyum gösterip göstermediği belirlenmiştir. İlk Mantel-Haenszel ve Lojistik Regresyon yöntemiyle yapılan analizlerde, değişen madde fonksiyonu ile ilgili elde edilen sonuçların uyum gösterip göstermediği, her iki yöntemden elde edilen sonuçlara göre DMF gösteren maddeler, tespit edilen DMF düzeyleri ve toplam DMF gösteren madde sayıları kıyaslanarak belirlenmiştir. Yöntemlerin maddelerdeki DMF miktarlarının büyüklük sıralaması bakımından benzerliklerinin belirlenebilmesi için ise, her iki yöntem için elde edilen DMF büyüklüğü istatistikleri arasında Spearman Brown sıralama farkı korelasyon katsayısı hesaplanmıştır. İkinci analizde ise, ikinci MH ve LR yöntemleriyle yapılan analizlerde, değişen madde fonksiyonuyla ilgili elde edilen sonuçların uyum gösterip göstermediği, ilk kısımdaki bulguları elde etmek için izlenen işlem basamakları izlenerek gerçekleştirilmiştir.

## BÖLÜM III: BULGULAR

Bu bölümde, araştırmanın alt problemlerine ilişkin bulgulara yer verilmiştir.

### **III.1.a. Mantel-Haenszel yöntemiyle yapılan analizde, odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular**

Odak grubunun, KZ, referans grubunun ZK test formunu alması durumuna göre maddelerin DMF gösterip göstermediğini belirlemek üzere yapılan MH yöntemine ilişkin analiz sonuçları Tablo 13'de belirtilmiştir.



**Tablo 13:** Odak Grubunun Test 2' yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Mantel-Haenszel Yöntemine İlişkin Analiz Sonuçları

Madde No	$\alpha_{MH}$	$\chi^2$	p	$\Delta_{MH}$	DMF düzeyi
1	0,7034	1,6836	1,1944	0,8267	A
2	0,9312	0,0648	0,7991	0,1674	A
3	1,0253	0,0072	0,9325	-0,0586	A
4	1,0210	0,0061	0,9376	-0,0487	A
5	1,0705	0,0554	0,8140	-0,1600	A
6	0,8402	0,3718	0,5420	0,4091	A
7	2,0501	6,5341	0,0106*	-1,6871	C
8	1,4643	1,6782	0,1952	-0,8963	A
9	1,1046	0,1367	0,7115	-0,2338	A
10	0,8848	0,2306	0,6311	0,2877	A
11	0,8023	0,4433	0,5055	0,5177	A
12	0,8000	0,6063	0,4362	0,5245	A
13	0,8296	0,3501	0,5541	0,4389	A
14	0,7962	0,6634	0,4154	0,5356	A
15	0,6483	1,8249	0,1767	1,0184	B
16	0,9315	0,0493	0,8242	0,1667	A
17	0,6404	1,9241	0,1654	1,0472	B
18	1,6512	1,5771	0,2092	-1,1786	B
19	1,8604	4,6908	0,0303*	-1,4589	B
20	1,1811	0,1819	0,6697	-0,3912	A

\*  $p \leq .05$

Referans grup: Maddelerin güçlük düzeylerine göre zordan kolaya doğru sıralandığı test formunu (Test 3) alan öğrenciler ( n=150 )

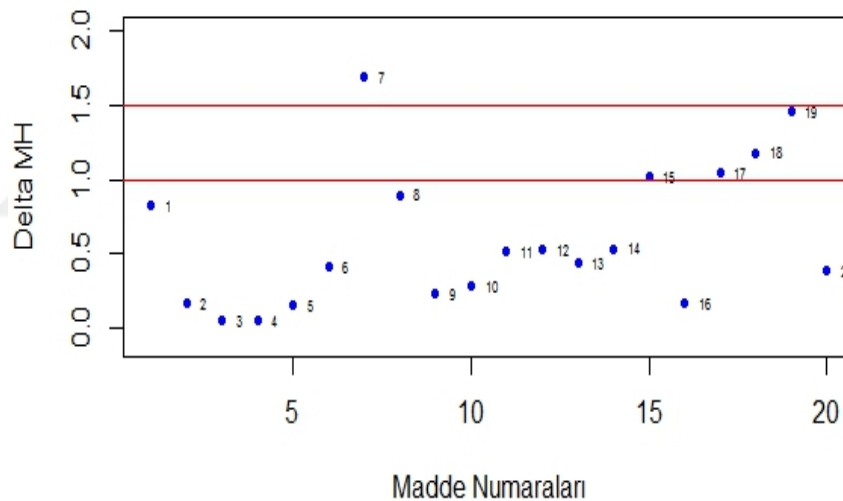
Odak grup: Maddelerin güçlük düzeylerine göre kolaydan zora doğru sıralandığı test formunu (Test 2) alan öğrenciler ( n=150 )

Tablo 13 'de yer alan maddelere ait  $\Delta_{MH}$  değerleri, MH yöntemi için kullanılan  $\Delta_{MH}$  eşik değerleriyle kıyaslanarak maddelerin hangi düzeyde DMF gösterdiği tespit edilmiştir.  $\Delta_{MH}$  değerinin negatif ya da pozitif olmasına göre, DMF gösteren ve güçlük düzeyleri kolay, orta ve zor olan maddelerin hangi gruba avantajlı davrandığı belirlenmiştir.

Yapılan değerlendirme sonuçları Tablo 14'de belirtilmiş olup, çeşitli düzeylerde DMF gösterdiği belirlenmiş test maddelerinin  $\Delta_{MH}$  değerlerine ilişkin grafik Şekil 6'da gösterilmiştir.

**Tablo 14:** Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Mantel-Haenszel Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar

Avantajlı Grup	DMF Düzeyi					
	B			C		
	Kolay	Orta	Zor	Kolay	Orta	Zor
Odak Grup	-	15	17	-	-	-
Referans Grup	-	-	18 19	-	7	-



**Şekil 6.** Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Mantel-Haenszel Analizine İlişkin Madde Dağılımları

Tablo 13 ve Tablo 14'den elde edilen sonuçlar incelendiğinde, kareköklü sayılar ve kareköklü sayılarla işlemler testinde yer alan, 4 maddenin orta düzeyde (B), 1 maddenin de yüksek düzeyde (C) DMF gösterdiği görülmüştür. KZ test formunu alan öğrencilerin (odak grup) lehine olan 2 maddenin orta düzeyde (B) DMF gösterdiği görülmektedir. B düzeyinde DMF gösteren bu maddelerden, madde 15'in orta güçlükte madde, 17'nin ise zor madde olduğu görülmektedir. ZK test formunu alan öğrencilerin (referans grup) lehine olan 2 maddenin orta düzeyde (B) ve 1 maddenin de yüksek düzeyde (C) DMF gösterdiği görülmektedir. B düzeyinde DMF içeren madde 18 ve madde 19'un, zor maddeler

olduğu görülmektedir. C düzeyinde DMF içeren 7. maddenin ise, orta güçlükte bir madde olduğu görülmektedir.

### III.1.b. Mantel-Haenszel yöntemiyle yapılan analizde, odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular

Odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre DMF içerip içermediğini belirlemek üzere yapılan MH yöntemine ilişkin analiz sonuçları Tablo 15'de belirtilmiştir.

**Tablo 15:** Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Mantel-Haenszel Yöntemine İlişkin Analiz Sonuçları

Madde No	$\alpha_{MH}$	$\chi^2$	p	$\Delta_{MH}$	DMF düzeyi
1	1,2154	0,4763	0,4901	-0,4585	A
2	0,9070	0,1265	0,7221	0,2294	A
3	0,6787	1,7910	0,1808	0,9110	A
4	1,2027	0,4526	0,5011	-0,4337	A
5	2,0171	5,2957	0,0214*	-1,6490	C
6	1,2501	0,6696	0,4132	-0,5246	A
7	0,4359	7,9903	0,0047**	1,9512	C
8	1,3927	1,3655	0,2426	-0,7785	A
9	0,7567	1,0466	0,3063	0,6551	A
10	1,3697	1,4356	0,2308	-0,7393	A
11	1,0349	0,0143	0,9048	-0,0806	A
12	1,3178	0,9437	0,3313	-0,6485	A
13	2,8806	13,2429	0,0003***	-2,4863	C
14	1,0909	0,0963	0,7564	-0,2045	A
15	1,0332	0,0125	0,9109	-0,0768	A
16	0,4663	6,2051	0,0127*	1,7931	C
17	1,0232	0,0062	0,9373	-0,0540	A
18	0,7497	0,5678	0,4511	0,6769	A
19	0,4385	7,0855	0,0078**	1,9373	C
20	0,5911	2,0238	0,1548	1,2354	B

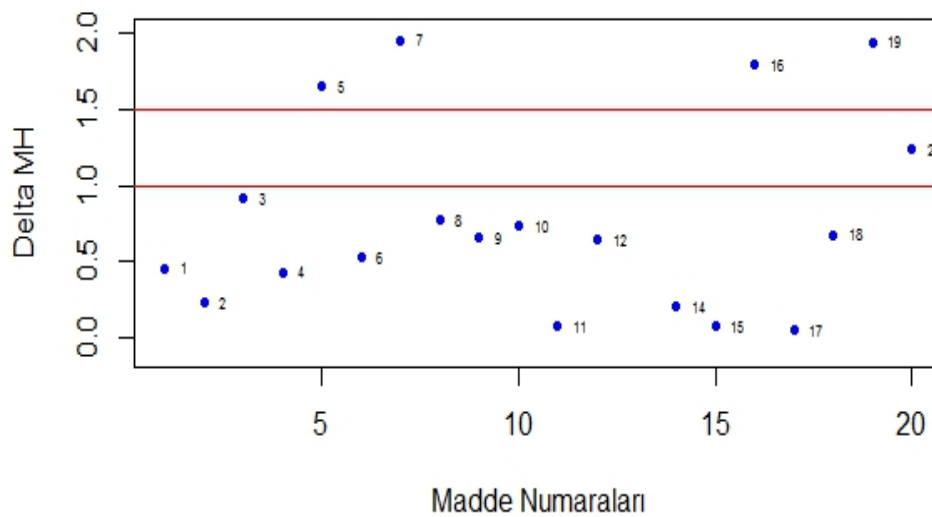
\*  $p \leq .05$  \*\*  $p \leq .01$  \*\*\*  $p \leq .001$



Tablo 15’de yer alan maddelere ait  $\Delta_{MH}$  değerleri, MH yöntemi için kullanılan  $\Delta_{MH}$  eşik değerleriyle kıyaslanarak maddelerin hangi düzeyde DMF gösterdiği tespit edilmiştir.  $\Delta_{MH}$  değerinin negatif ya da pozitif olmasına göre, DMF gösteren ve güçlük düzeylerine göre kolay, orta ve zor maddelerin hangi gruba avantajlı davrandığı belirlenmiştir. Yapılan değerlendirme sonuçları Tablo 16’da belirtilmiş olup çeşitli düzeylerde DMF gösterdiği belirlenmiş test maddelerinin  $\Delta_{MH}$  değerlerine ilişkin grafik Şekil 7’de gösterilmiştir.

**Tablo 16:** Odak Grubunun Test 3’ü, Referans Grubunun Test 2’yi Alması Durumuna Göre Mantel-Haenszel Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar

Avantajlı Grup	DMF Düzeyi					
	B			C		
	Kolay	Orta	Zor	Kolay	Orta	Zor
Odak Grup	-	-	5	13	-	-
Referans Grup	-	-	20	-	7,16	19



**Şekil 7.** Odak Grubunun Test 3’ü, Referans Grubunun Test 2’yi Alması Durumuna Göre Mantel-Haenszel Analizine İlişkin Madde Dağılımları

Tablo 15 ve Tablo 16' dan elde edilen sonuçlar incelendiğinde, kareköklü sayılar ve kareköklü sayılarla işlemler testinde yer alan, 1 maddenin orta düzeyde (B), 5 maddenin de yüksek düzeyde (C) DMF gösterdiği görülmektedir. ZK test formunu alan öğrencilerin (odak grup) lehine olan 2 maddenin yüksek düzeyde (C) DMF gösterdiği görülmektedir. C düzeyinde DMF gösteren bu maddelerden, madde 5'in kolay madde 13'ün ise orta güçlükte madde olduğu görülmektedir. KZ test formunu alan öğrencilerin (referans grup) lehine olan 1 maddenin orta düzeyde (B) ve 3 maddenin de yüksek düzeyde (C) DMF gösterdiği görülmüştür. B düzeyinde DMF içeren 20. maddenin zor bir madde olduğu görülmektedir. Ayrıca, B düzeyinde DMF içeren madde 7 ve madde 16'nın orta güçlükte ve madde 19'un zor madde olduğu görülmektedir.

### III.1.c. Mantel-Haenszel yöntemiyle yapılan analizlerde, değişen madde fonksiyonu gösteren maddelerin farklılık gösterip göstermediğine ilişkin bulgular

Mantel-Haenszel yöntemiyle yapılan analizlerde, değişen madde fonksiyonu gösteren maddeler Tablo 17'de belirtilmiştir.

**Tablo 17:** Mantel- Haenszel Yöntemleriyle Yapılan Analiz Sonucunda DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar

Yöntem	Avantajlı Grup	DMF Düzeyi					
		Kolay	Orta	Zor	Kolay	Orta	Zor
MH (Birinci)	Odak Grup	-	15	17	-	-	-
	Referans Grup	-	-	18,19	-	7	-
MH (İkinci)	Odak Grup	-	-	-	5	13	-
	Referans Grup	-	-	20	-	7,16	19

Tablo 17 incelendiğinde, birinci MH analizi sonucunda, 1 orta güçlükte (madde 15) ve 1 zor maddenin (madde 17) orta düzeyde (B) odak grubun lehine DMF gösterdiği görülmektedir. Ayrıca, 2 zor maddenin (madde 18 ve madde 19) orta düzeyde (B), 1 orta güçlükte maddenin (madde 7) yüksek düzeyde (C) referans grubunun lehine DMF gösterdiği görülmektedir. İkinci MH analizi sonucunda ise, 1 kolay (madde 5) ve 1 orta güçlükte maddenin (madde 13) yüksek düzeyde (C) odak grubunun lehine DMF gösterdiği görülmektedir. Ayrıca, ikinci MH analiz sonucunda ise, 1 zor maddenin (madde 20)

orta düzeyde (B), 2 orta güçlükte (madde 7 ve madde 16) ve 1 zor (madde 19) maddenin yüksek düzeyde (C) referans grubunun lehine DMF gösterdiği görülmektedir.

Birinci MH analizinde 5 maddede (7., 15., 17., 18. ve 19. madde), ikinci MH analizinde ise 6 maddede (5., 7., 13., 16., 19. ve 20. madde) DMF gözlenmiştir. Hem birinci hem de ikinci MH analizinde bulunan DMF gösteren maddeler incelendiğinde, iki maddenin ortak (7. ve 19. madde) ortak olduğu, ancak bu maddelerden birinin (19. madde) DMF düzeyinin farklı olduğu ve bu maddelerin avantaj sağladığı grupların farklı olduğu görülmektedir.

MH analizlerinde, KZ test formunu alan grubun lehine DMF içeren kolay maddenin (5. madde) bulunduğu, ancak ZK test formunu alan öğrencilerin lehine böyle kolay maddelerin olmadığı görülmektedir. Orta güçlükte yer alan maddelerin, her iki uygulamada da hem odak hem de referans grubunun lehine işlediği görülmektedir. Güçlük düzeyi zor olan maddelerin ise, hem KZ hem de ZK test formunu alan grupların lehine işlediği görülmektedir.

### **III.2.a. Lojistik Regresyon yöntemiyle yapılan analizde, odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular**

Odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, maddelerin, DMF gösterip göstermediğini belirlemek üzere yapılan LR yöntemine ilişkin analiz sonuçları Tablo 18'de belirtilmiştir. Tek biçimli DMF'yi belirlemek için yapılan analizde, grup değişkeni ile toplam puanın birlikte ele alındığı ikinci modelden elde edilen  $R^2$  ile, toplam puan değişkeninin ele alındığı birinci modelden elde edilen  $R^2$  arasındaki fark elde edilmiştir. Tek biçimli olmayan DMF'yi belirlemek için yapılan analizde ise, toplam puanın, gruplama değişkeninin ve etkileşim değişkeninin ele alındığı üçüncü modelden hesaplanan  $R^2$  ile toplam puan değişkeninin ele alındığı birinci modelden elde edilen  $R^2$  arasındaki fark elde edilmiştir. Elde edilen  $\Delta R^2$  değerleri, Jodoin ve Gierl (2001) tarafından belirlenen sınır değerler kullanılarak değerlendirilmiş ve maddelerin gösterdikleri DMF düzeyleri belirlenmiştir. Yapılan değerlendirme sonuçları Tablo 19'da gösterilmektedir. LR analizlerine göre elde edilen  $\Delta R^2$  değerleri Şekil 8 ve Şekil 9'da belirtilmiştir.

Tablo 18 ve Tablo 19'da yer alan bilgiler ışığında kareköklü sayılar ve kareköklü sayılarla işlemler testinde, hem Tek biçimli hem de Tek biçimli olmayan DMF'yi belirlemek için yapılan LR analizleri sonucunda, orta düzeyde (B) ve yüksek düzeyde (C) DMF gösteren maddenin bulunmadığı görülmektedir.

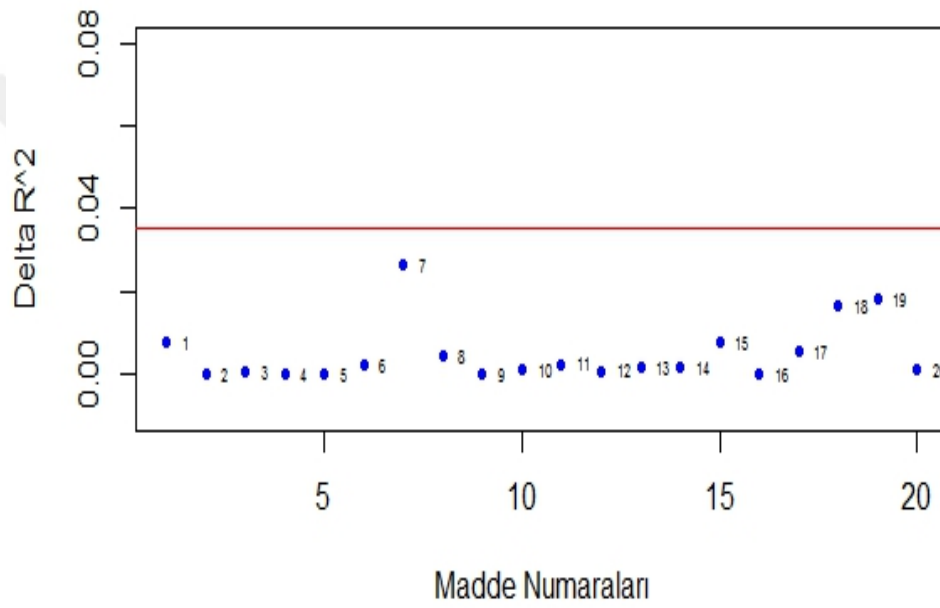
**Tablo 18:** Odak Grubunun Test 2'yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Lojistik Regresyon Analizine İlişkin Sonuçlar

Madde No	Tek Biçimli DMF Analizi				Tek Biçimli Olmayan DMF Analizi			
	$\beta$	p	$\Delta R^2$	DMF Düzeyi	$\beta$	p	$\Delta R^2$	DMF Düzeyi
1	1,8671	0,1718	0,0249	A	0,3793	0,5380	0,0015	A
2	0,0046	0,9462	0,0000	A	0,9301	0,3348	0,0036	A
3	0,1425	0,7058	0,0005	A	0,1047	0,7463	0,0009	A
4	0,0125	0,9112	0,0000	A	1,2872	0,2566	0,0045	A
5	0,0265	0,8706	0,0001	A	0,0950	0,7579	0,0003	A
6	0,6506	0,4199	0,0021	A	0,2731	0,6012	0,0009	A
7	7,4592	0,0063**	0,0262	A	0,8373	0,3602	0,0029	A
8	1,3626	0,2431	0,0044	A	0,0987	0,7534	0,0003	A
9	0,0023	0,9621	0,0000	A	1,4139	0,2344	0,0051	A
10	0,2429	0,6221	0,0011	A	0,2347	0,6280	0,0010	A
11	0,6050	0,4367	0,0024	A	1,2426	0,2650	0,0050	A
12	0,1646	0,6850	0,0007	A	0,2574	0,6119	0,0011	A
13	0,4898	0,4840	0,0017	A	1,5713	0,2100	0,0056	A
14	0,3064	0,5799	0,0014	A	2,9674	0,0850	0,0135	A
15	1,9838	0,1590	0,0075	A	3,9923	0,0457*	0,0149	A
16	0,0012	0,9728	0,0000	A	4,3166	0,0377*	0,0211	A
17	1,1475	0,2841	0,0054	A	0,8756	0,3494	0,0041	A
18	2,8021	0,0941	0,0164	A	0,8359	0,3606	0,0049	A
19	3,9403	0,0471*	0,0181	A	0,6384	0,4243	0,0029	A
20	0,1515	0,6971	0,0011	A	1,0268	0,3109	0,0071	A

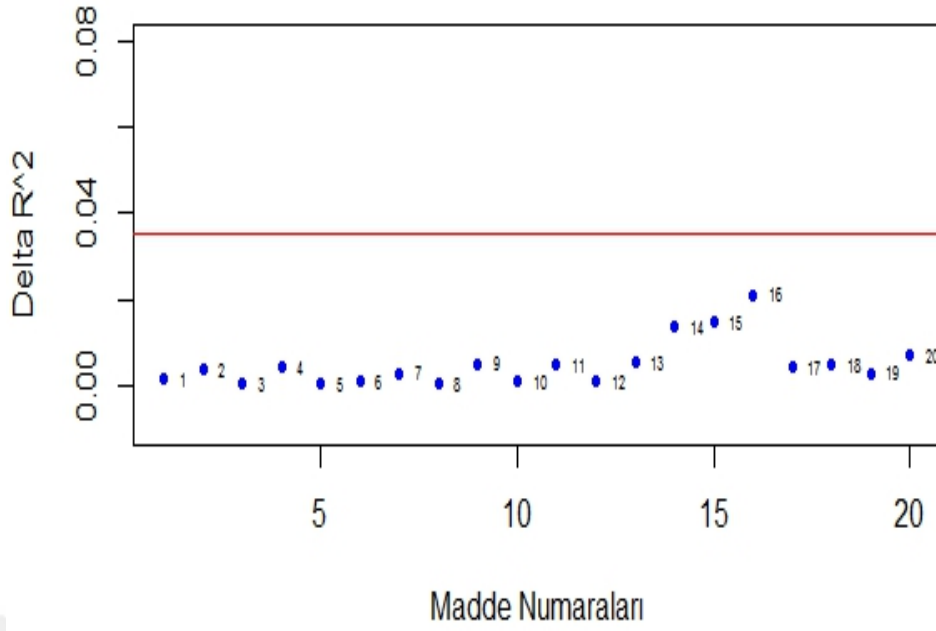
p ≤ .1 \* p ≤ .05 \*\* p ≤ .01

**Tablo 19:** Odak Grubunun Test 3'ü, Referans Grubunun Test 2' yi Alması Durumuna Göre Lojistik Regresyon Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar

DMF Biçimi	DMF Düzeyi	
	B	C
Tek Biçimli	-	-
Tek Biçimli	-	-
Olmayan		



**Şekil 8.** Odak Grubunun Test 2' yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı



**Şekil 9.** Odak Grubunun Test 2' yi, Referans Grubunun Test 3'ü Alması Durumuna Göre Tek Biçimli Olmayan DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı

### III.2.b. Lojistik Regresyon yöntemiyle yapılan analizde, odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelere ilişkin bulgular

Odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, maddelerin, DMF gösterip göstermediğini belirlemek üzere yapılan LR yöntemine ilişkin analiz sonuçları Tablo 20'de belirtilmiştir.

**Tablo 20:** Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Lojistik Regresyon Analizine İlişkin Sonuçlar

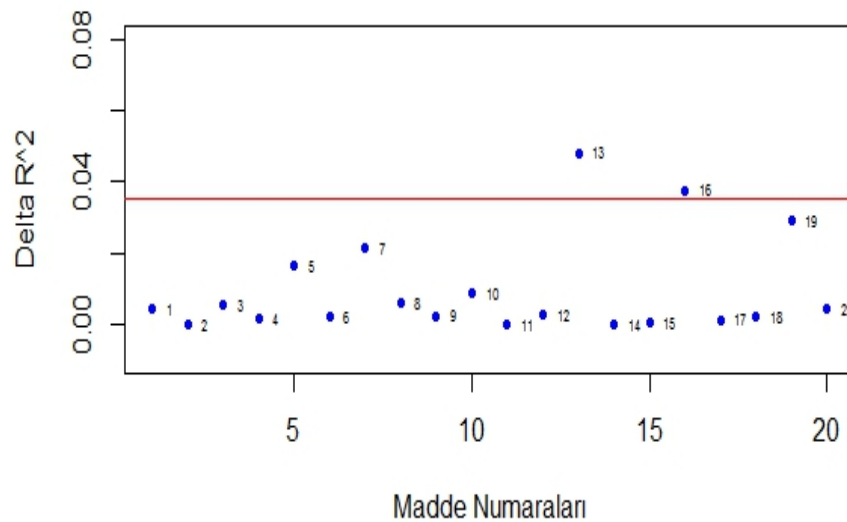
Madde No	Tek Biçimli DMF Analizi				Tek Biçimli Olmayan DMF Analizi			
	$\beta$	p	$\Delta R^2$	DMF Düzeyi	$\beta$	p	$\Delta R^2$	DMF Düzeyi
1	1,1469	0,2842	0,0043	A	0,0504	0,8223	0,0002	A
2	0,0313	0,8597	0,0001	A	0,2410	0,6235	0,0009	A
3	1,4358	0,2308	0,0052	A	2,4919	0,1144	0,0009	A
4	0,4039	0,5251	0,0014	A	0,1214	0,7275	0,0004	A
5	5,6473	0,0175*	0,0163	A	0,0750	0,7842	0,0002	A
6	0,5689	0,4507	0,0019	A	8,0753	0,0045**	0,0259	A
7	7,3505	0,0067**	0,0212	A	1,4429	0,2297	0,0041	A
8	1,7917	0,1807	0,0061	A	0,8552	0,3551	0,0029	A
9	0,7034	0,4016	0,0024	A	0,2647	0,6069	0,0009	A
10	2,3237	0,1274	0,0087	A	3,5740	0,0587	0,0133	A
11	0,0351	0,8514	0,0001	A	6,0795	0,0137*	0,0217	A
12	0,6506	0,4199	0,0028	A	1,7662	0,1839	0,0076	A
13	12,4331	0,0004***	0,0481	B	1,4313	0,2316	0,0054	A
14	0,0190	0,8904	0,0001	A	0,8620	0,3532	0,0035	A
15	0,1317	0,7167	0,0005	A	1,9959	0,1577	0,0071	A
16	8,3330	0,0039**	0,0376	B	0,0040	0,9498	0,0000	A
17	0,2932	0,5882	0,0013	A	4,9415	0,0262*	0,0209	A
18	0,4227	0,5156	0,0022	A	0,6751	0,4113	0,0036	A
19	6,9808	0,0082**	0,0290	A	0,7676	0,3810	0,0031	A
20	0,8308	0,3620	0,0043	A	0,2986	0,5848	0,0016	A

p ≤ .1 \* p ≤ .05 \*\* p ≤ .01

LR analizinde, Tek biçimli ve Tek biçimli olmayan DMF' yi belirlemek için yapılan analizlerden elde edilen  $\Delta R^2$  değerleri, Jodoin ve Gierl (2001) tarafından belirlenen sınır değerler kullanılarak değerlendirilmiş ve böylece maddelerin gösterdikleri DMF düzeyleri belirlenmiştir. Yapılan değerlendirme sonuçları Tablo 21'de gösterilmektedir. LR analizlerinden sonra elde edilen  $\Delta R^2$  değerleri Şekil 10 ve Şekil 11'de belirtilmiştir.

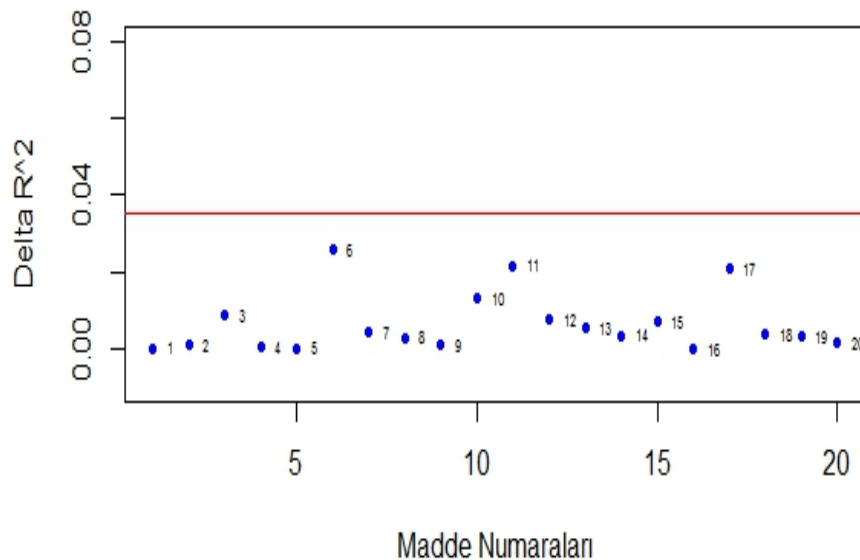
**Tablo 21:** Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre Lojistik Regresyon Yönteminde DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar

		DMF Düzeyi						
		B			C			
DMF Biçimi	Avantajlı Grup	Kolay	Orta	Zor	Kolay	Orta Güçlükte	Zor	
	Odak Grup	-	13	-	-	-	-	
Tek Biçimli	Referans Grup	-	16	-	-	-	-	
Tek Biçimli Olmayan	Odak Grup	-	-	-	-	-	-	
	Referans Grup	-	-	-	-	-	-	



**Şekil 10.** Odak Grubunun Test 3'ü, Referans Grubunun Test 2' yi Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı





**Şekil 11.** Odak Grubunun Test 3'ü, Referans Grubunun Test 2' yi Alması Durumuna Göre Tek Biçimli Olmayan DMF'yi Belirlemek İçin Yapılan Lojistik Regresyon Analizine İlişkin Maddelerin Dağılımı

Tablo 20 ve Tablo 21' de yer alan bilgiler ışığında, kareköklü sayılar ve kareköklü sayılarla işlemler testinde bulunan ve Tek biçimli DMF'yi belirlemek için yapılan LR analizine dahil edilen, 2 maddenin orta düzeyde (B) Tek biçimli DMF gösterdiği görülmektedir. Orta düzeyde (B) DMF gösterdiği belirlenen maddelerin hangi gruba avantaj sağladıklarını belirleyebilmek için maddelerin odak ve referans grubundaki puanları dikkate alınarak İlişkisiz Örneklemeler için t Testi uygulanmıştır. Orta düzeyde (B) DMF gösterdiği belirlenen 13. Maddenin ZK test formunu alan öğrencilerin (odak grup) lehine işlediği ve bu maddenin orta güçlükte bir madde olduğu görülmektedir. 16. maddenin ise, KZ test formunu alan öğrencilerin (referans grup) lehine işlediği ve orta güçlükte bir madde olduğu görülmektedir. Tek biçimli olmayan DMF'yi belirlemek için yapılan LR analizine dahil edilen 20 maddenin tümünün önemsiz düzeyde (A) DMF gösterdiği görülmektedir.

### **III.2.c. Lojistik Regresyon yöntemiyle yapılan analizlerde, değişen madde fonksiyonu gösteren maddelerin farklılık gösterip göstermediğine ilişkin bulgular**

Lojistik Regresyon yöntemiyle yapılan analizlerde, Tek biçimli ve Tek biçimli olmayan DMF'yi belirlemek için yapılan ilk LR analiz sonuçlarında, odak ve referans grubunun lehine, orta düzeyde (B) ve yüksek düzeyde (C) DMF gösteren maddenin bulunmadığı görülmektedir. İkinci LR analiz sonuçlarında ise, Tek biçimli DMF'yi belirlemek için yapılan analize göre, 1 orta güçlükte maddenin (madde 13) orta düzeyde (B) odak grup lehine, 1 orta güçlükte maddenin de (madde 16) orta düzeyde (B) referans grup lehine DMF gösterdiği görülmektedir.

Hem birinci hem de ikinci LR analizinde bulunan DMF'li maddelerin sayılarının, düzeylerinin ve bu maddelerin avantaj sağladığı grupların farklı olduğu görülmektedir.

### III.3.a. Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle yapılan analizlerde, odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelerin uyum gösterip göstermediğine ilişkin bulgular

MH ve Tek Biçimli DMF' yi belirleyebilmek için yapılan LR yöntemlerinden elde edilen sonuçların uyum derecesi, her iki yöntemden elde edilen sonuçlara göre DMF gösteren maddeler, bu maddelerin tespit edilen DMF düzeyleri ve toplam DMF gösteren madde sayıları karşılaştırılarak belirlenmiştir. Bu bilgiler Tablo 22' de belirtilmiştir.

**Tablo 22:** Odak Grubunun Test 2' yi, Referans Grubunun Test 3'ü Alması Durumuna Göre LR ve MH Yöntemleriyle DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar

	DMF Düzeyi	
	B	C
MH analizi	15, 17, 18, 19	7
LR analizi (Tek Biçimli)	-	-

Tablo 22'de yer alan bilgiler incelendiğinde, MH yönteminde DMF gösteren maddelerden, 4'ünün orta düzeyde (B), 1'inin de yüksek düzeyde (C) DMF gösterdiği belirlenmiştir. Tek Biçimli DMF'yi belirlemek için yapılan LR yöntemiyle yapılan analizde ise, orta düzeyde (B) ve yüksek düzeyde (C) DMF gösteren maddenin bulunmadığı görülmektedir.

Yöntemlerin maddelerdeki DMF miktarlarının büyüklük sıralaması bakımından benzerliklerinin belirlenebilmesi için, Tek Biçimli DMF'yi belirleyebilmek için yapılan LR ve MH yöntemlerine göre elde edilen ki-kare değerleri arasında Spearman sıra farkları korelasyon katsayısı hesaplanmıştır. Hesaplamalar sonucunda, iki yöntemin DMF büyüklük sıralamaları arasında manidar bir ilişkinin bulunduğu görülmektedir ( $r = 0.908$ ,  $p = 0.01$ ).

### III.3.b. Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle yapılan analizlerde, odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, değişen madde fonksiyonu gösteren maddelerin uyum gösterip göstermediğine ilişkin bulgular

MH ve Tek Biçimli DMF'yi belirleyebilmek için yapılan LR yöntemlerinden elde edilen sonuçların uyum derecesi, her iki yöntemden elde edilen sonuçlara göre DMF gösteren maddeler, bu maddelerin tespit edilen DMF düzeyleri ve toplam DMF gösteren madde sayıları karşılaştırılarak belirlenmiştir. Bu bilgiler Tablo 23'de belirtilmiştir.

**Tablo 23:** Odak Grubunun Test 3'ü, Referans Grubunun Test 2'yi Alması Durumuna Göre LR ve MH Yöntemleriyle DMF Gösteren Maddelerin DMF Düzeylerine İlişkin Sonuçlar

	DMF Düzeyi	
	B	C
MH analizi	20	5, 7, 13, 16, 19
LR analizi (Tek Biçimli)	13,16	-

Tablo 23'de yer alan bilgiler incelendiğinde, MH ile yapılan analiz sonuçlarına göre, 1 maddenin orta düzeyde (B), 5 maddenin de yüksek düzeyde (C) DMF gösterdiği belirlenmiştir. LR yöntemi ile Tek Biçimli DMF'yi belirlemek için yapılan analiz sonuçlarına göre 2 maddenin orta düzeyde (B) DMF gösterdiği belirlenmiştir.

Yöntemlerin maddelerdeki DMF miktarlarının büyüklük sıralaması bakımından benzerliklerinin belirlenebilmesi için Tek biçimli DMF'yi belirlemek için yapılan LR ve MH yöntemlerine göre elde edilen ki-kare değerleri arasındaki Spearman sıra farkları korelasyon katsayısı hesaplanmıştır. Hesaplamalar sonucunda, iki yöntemin DMF büyüklük sıralaması arasında manidar bir ilişki bulunduğu görülmektedir ( $r = 0,926$ ,  $p = 0,01$ ).

## BÖLÜM IV. TARTIŞMA VE YORUM

Bu bölümde elde edilen bulgular araştırma soruları çerçevesinde ele alınıp ilgili araştırmalarla tartışılıp yorumlanmıştır.

### IV.1. Mantel-Haenszel yöntemiyle yapılan analiz sonuçlarının değerlendirilmesi

Odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, maddelerin DMF içerip içermediğini belirlemek için yapılan MH yönteminden elde edilen analiz sonuçlarında, bir zor ve bir orta güçlükte maddenin, orta düzeyde (B), odak grubunun lehine DMF içerdiği görülmüştür. Referans grubu açısından bu durum incelendiğinde ise, iki zor maddenin orta düzeyde (B), bir orta güçlükte maddenin, yüksek düzeyde (C), bu grup lehine DMF içerdiği görülmüştür. Buradan hareketle, odak grubunda yer alan bireylerin iki maddede, referans grubunda yer alan bireylerin ise üç maddede daha yüksek performans gösterdiği ve farklı alt gruplarda yer alan bireylerin testteki toplam beş maddeyi doğru cevaplama olasılıklarının farklı olduğu söylenebilir.

Odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, maddelerin, DMF içerip içermediğini belirlemek için yapılan MH yönteminden elde edilen analiz sonuçlarında, bir kolay ve bir orta güçlükte maddenin, yüksek düzeyde (C), odak grubunun lehine DMF içerdiği görülmüştür. Bununla beraber, iki zor maddenin, birisi orta düzeyde (B), diğeri yüksek düzeyde (C), iki orta güçlükte maddenin yüksek düzeyde (C), referans grubunun lehine DMF içerdiği görülmüştür. Buradan hareketle, odak grubunda yer alan bireylerin iki maddede, referans grubunda yer alan bireylerin ise, dört maddede daha yüksek performans gösterdiği ve alt gruplarda yer alan bireylerin testteki toplam altı maddeyi doğru cevaplama olasılıklarının farklı olduğu söylenebilir.

Odak grubunda yer alan bireyler, KZ test formunu önce aldığımda, güçlük düzeyi zor olan maddelerde bu grubun lehine DMF'nin varlığı gözlenmiştir. Ancak, odak grubunda yer alan bireyler, ZK test formunu sonra aldıklarında, güçlük düzeyi kolay olan maddelerde bu grubun lehine DMF'nin varlığı görülmemektedir. Referans grubunda yer alan bireyler ise, KZ test formunu sonra aldıklarında, orta güçlük düzeyinde DMF gösteren madde sayılarında artmanın olduğu ve güçlük düzeyi zor olan maddelerdeki DMF düzeyinin daha yüksek olduğu gözlenmiştir. Bu bulgulardan hareketle, odak ya da referans grubunda yer alan bireylerin farklı test formlarını alma sıralarının, başka bir ifadeyle sıra etkisinin de maddelere doğru cevap verme olasılıklarını etkilediği söylenebilir.

Her iki analize dair bulgular, güçlük düzeyi kolay maddeler açısından incelendiğinde, KZ test formunu alan öğrenciler lehine, DMF içeren maddenin bulunmadığı, ZK test formunu alan öğrencilerin lehine, DMF içeren maddelerin bulunduğu görülmektedir. Bu durumda, kolay maddeleri sonra alan grubun, maddeleri doğru cevaplama olasılıklarında artışların olduğu söylenebilir. Orta güçlükte yer alan

maddelerin, her iki uygulamada da hem odak hem de referans grubunun lehine işlediği görülmektedir. Bu durumda, orta güçlükteki maddelerin kolay ya da zor maddeden sonra gelmesinin maddenin doğru cevaplama olasılığını etkilediği söylenebilir. Güçlük düzeyi zor maddeler açısından, her iki analize dair bulgular incelendiğinde, güçlük düzeyi zor olan maddelerin hem testin başında yer aldığı durumda hem de testin sonunda yer aldığı durumda, maddelerin doğru cevaplandırılma olasılığını etkilediği söylenebilir. Bu bulgular ışığında, özellikle zor maddelerin testin başında ya da sonunda yer alması ya da orta güçlükteki maddelerin kolay ya da zor maddelerden sonra gelmesinin, farklı gruplarda yer alan bireylerin maddelere doğru cevap verme olasılıklarını etkilediği sonucuna ulaşılabılır. Ayrıca, Mantel-Haenszel yöntemleriyle yapılan analiz sonuçlarında, DMF gösteren maddeler ve bu maddelerin sayılarının, DMF düzeylerinin ve avantaj sağladığı grupların farklı olduğu sonucu da çıkarılabilir.

Böylece bu çalışmada, maddelerin güçlük düzeylerine göre farklı şekilde (kolaydan zora ve zordan kolay) sıralanmasının farklı gruplarda yer alan bireylerin, maddelere, doğru cevap verme olasılıklarını etkilediği sonucuna ulaşılmıştır. Ayrıca, zor maddeler test formunun sonunda yer alması, maddelerin cevaplanma olasılığındaki farklılığın artmasına neden olmaktadır. Bu durum, maddelerin güçlük düzeylerine göre kolaydan zora doğru sıralandığı, yani madde güçlüğü açısından zor maddenin testin sonlarına doğru yer aldığı durumlarda pratik ya da öğrenme etkisinin varlığını destekler niteliktedir (Hohensinn, vd., 2011, Kubinger, 2008, Eignor ve Cook,1983; Davis ve Ferdous, 2005). Maddelerin güçlük düzeylerine göre farklı şekillerde sıralanmasının bireylerin test performanslarına etkisini inceleyen çalışmalarla ilgili literatür incelendiğinde anlamlı farklılıkların olduğu (Mollenkopf, 1950; Sax ve Cromack, 1966; Flaugher, Melton ve Meyers, 1968; Barciovski ve Olsen, 1974; Hambleton ve Traub, 1974; Plake, Ansorge,Parker ve Lowry, 1982; Tippets ve Benson) ve olmadığı (Brenner, 1964; Smouse ve Munz, 1968, Kestenbaum ve Weiner, 1970; Huck ve Bowers, 1972; Marso, 1970; Allison, 1984; Klimko, 1984; Zumbo, Perlini ve Lind, 1998; Ollenue ve Etsey, 2009) çalışmalara rastlanmıştır. Yapılan araştırmalarda sonuçların farklı çıkmasının nedenleri arasında, bazı çalışmalarda madde güçlüğü olarak uzman kanısına dayalı madde güçlük indekslerinin kullanılması (Zumbo, Perlini ve Lind, 1998), bazı çalışmalarda kullanılan ölçme aracının daha önceki uygulanmasından elde edilen ya da çalışmada geliştirilen ölçme aracının pilot uygulamasından elde edilen madde güçlük indekslerinin kullanılması (Hambleton ve Traub, 1974; Gerow, 1980; Carlson ve Ostrosky,1992), madde güçlük indeksine göre maddelerin kolay, orta ve zor olarak adlandırılmasında kesin bir ölçütün bulunmaması ve maddelerin sırasının kontrolünün yönerge ile yapılmasından kaynaklı olarak bireylerin maddeleri istenen sırada cevaplayıp cevaplamadığından emin olunmaması ve bu durumun karıştırıcı değişken olarak ortaya çıkması sayılabilir.

## IV.2. Lojistik Regresyon yöntemiyle yapılan analiz sonuçlarının değerlendirilmesi

Odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, hem Tek biçimli hem de Tek biçimli olmayan DMF'yi belirlemek için yapılan LR yönteminden elde edilen analiz sonuçlarında, orta düzeyde (B) ve yüksek düzeyde (C) DMF gösteren maddenin bulunmadığı görülmektedir. Buradan hareketle, odak ve referans grubunda yer alan bireylerin, maddelere verdikleri tepkilerin benzer olduğu, başka bir ifadeyle farklı alt gruplarda yer alan bireylerin testteki maddeleri doğru cevaplama olasılıklarının farklı olmadığı söylenebilir.

Odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, Tek biçimli olmayan DMF'yi belirlemek için yapılan LR yönteminden elde edilen analiz sonuçlarında, DMF gösteren maddenin bulunmadığı görülmektedir. Tek biçimli DMF'yi belirlemek için yapılan LR yönteminden elde edilen analiz sonuçlarında ise, orta güçlükte iki maddenin orta düzeyde (B) DMF gösterdiği, görülmektedir. Bu maddelerden bir tanesi, odak grubun lehine iken, diğeri referans grubunun lehine işlediği görülmektedir. Bu durumda, hem odak hem de referans grubunda yer alan bireylerin birer maddede, diğeri göre daha yüksek performans gösterdiği ve farklı alt gruplarda yer alan bireylerin testteki toplam iki maddeyi doğru cevaplama olasılıklarının farklı olduğu söylenebilir.

Yapılan iki Tek biçimli LR analizinden elde edilen bulgulardan hareketle, özellikle orta güçlükte yer alan maddelerin kolay ya da zor maddeden sonra gelmesinin bireylerin doğru cevap verme olasılığında farklılığa yol açtığı söylenebilir. Ayrıca, odak ya da referans gruplarında yer alan bireylerin test formlarını önce ya da sonra almasının onların doğru cevap verme olasılığını etkileyebileceği, başka bir ifadeyle sıra etkisinin etkili olabileceği sonucu çıkartılabilir.

## IV.3. Mantel-Haenszel ve Lojistik Regresyon yöntemiyle yapılan analiz sonuçlarının karşılaştırması

Odak grubunun KZ, referans grubunun ZK test formunu alması durumuna göre, yapılan MH analizi sonucunda, testte yer alan dört maddenin orta düzeyde (B), bir maddenin de yüksek düzeyde (C) DMF içerdiği görülmektedir. Tek Biçimli DMF'yi belirlemek için yapılan LR analiz sonucuna göre ise, orta düzeyde (B) ve yüksek düzeyde (C) DMF içeren maddelerin olmadığı görülmektedir. Böylece, MH ve Tek Biçimli DMF'yi belirlemek için yapılan LR yöntemleri, B ve C düzeyinde DMF gösteren maddelerin belirlenmesinde uyum gösterememiştir. Yöntemlerin, maddelerdeki DMF miktarlarının büyüklük sıralaması bakımından benzerliklerinin belirlenebilmesi için hesaplanan Spearman sıra farkları korelasyon katsayısı ise manidar bulunmuştur. Bu sonuca göre, iki yöntemin DMF büyüklüğü sıralamasına göre uyum gösterdiği söylenebilir.

Odak grubunun ZK, referans grubunun KZ test formunu alması durumuna göre, yapılan MH analizi sonucunda, testte yer alan bir maddenin orta düzeyde (B), beş maddenin de yüksek düzeyde (C) DMF içerdiği görülmektedir. Tek Biçimli DMF'yi belirlemek için yapılan LR analiz sonucuna göre ise, iki maddenin orta düzeyde (B) DMF içerdiği görülmektedir. Her iki yönteme göre DMF içerdiği belirlenen maddelerden sadece bir tanesinin ortak olduğu, fakat bu maddenin DMF düzeyi bakımından farklı olduğu görülmektedir. İki yöntemin DMF miktarlarındaki büyüklük sıralamasının benzerliği açısından aralarında manidar bir ilişkinin bulunduğu görülmektedir. Bu sonuca göre, iki yöntemin DMF büyüklüğü sıralamasına göre uyum gösterdiği söylenebilir.

Elde edilen bulgular doğrultusunda, yapılan her iki analizde kullanılan LR ve MH yöntemlerinin, DMF miktarlarındaki büyüklük sıralamalarında benzer belirledikleri DMF'li maddeler bakımından farklı sonuçlar ürettiği sonucuna ulaşılmıştır. Bu açıdan bakıldığında, DMF içeren madde sayısı bakımından, MH yönteminin LR yönteminden, daha duyarlı olduğu söylenebilir. Bunun nedenini MH yönteminde, odak ve referans grupların madde parametre tahminlerinin aynı anda yapılması ve toplam örneklem büyüklüğünün LR'ye göre daha fazla olabilmesi söylenebilir. Buradan hareketle, LR'nin daha az sayıda DMF'li madde bulmasının sebebi örneklem büyüklüğüne bağlı olabilir (Penfield ve Camilli, 2007). Daha büyük örneklerde LR yönteminin daha duyarlı sonuçlar verebileceği bildirilmektedir (Pang ve vd.,1994; Jodoin ve Gierl, 2001). LR ve MH yöntemlerinin her ikisi için de tüm düzeydeki DMF'leri hassasiyetle belirleyebilmek için grupların her birinin 500 ya da 500 'den büyük olması gerektiği belirtilmiştir (Pang ve vd.,1994). Yapılan başka araştırmalarda da bu iki DMF belirleme yöntemi arasında tam bir uyum bulunamamıştır (Gondal, 2001; Gierl, Khaliq ve Boughton, 1999; Gomez, Benito ve Navas Ara, 2000; Bertrand, Bouteau, 2003; Gök, Kelecioğlu ve Doğan, 2010). Alanyazın incelendiğinde MH yönteminin, DMF belirlemede güçlü bir yöntem olduğu ve daha tutarlı sonuçlar verdiğini belirten çalışmalar (Hambleton ve Rogers, 1989; Clauser, Mazor ve Hambleton, 1993; Narayanan ve Swaminathan, 1994; Narayanan ve Swaminathan, 1994; Bertrand ve Bouteau,2003) olduğu gibi, LR yönteminin en etkili ve tavsiye edilen yöntemlerden biri olduğunun ifade edildiği çalışmalara da (Clauser ve Mazor, 1998; Swaminathan ve Rogers, 1990; Wiberg, 2007) rastlanmaktadır.

DMF miktarlarındaki büyüklük sıralamalarındaki paralelliğe rağmen, her iki yöntemde DMF'li maddeleri belirlemede kullanılan ölçütlerdeki farklılıkların, DMF düzeylerinde ve DMF'li madde sayılarındaki farklılıklara neden olduğu düşünülmektedir. Bu açıdan, DMF için Ki-kare değeri üreten ve kullanan bu tekniklerin hepsinde kullanılacak ortak ölçütler geliştirmenin faydalı olacağını belirtilmektedir (Doğan ve Öğretmen, 2008). Ayrıca literatür incelendiğinde, LR ve MH yöntemlerinden, farklı sayıda DMF gösteren madde elde edilebilmesinin nedenleri arasında; maddelerin güçlük indeksi ve ayırt edicilik indekslerinin, grup ortalamalarının ve örneklem büyüklüklerinin farklı olması gösterilmektedir (Narayahan ve Swaminathan, 1996; Hidalgo ve Pina, 2004; Penfield ve Camilli, 2007).

## ÖNERİLER

1. Bu araştırma kapsamında, DMF'nin belirlenmesinde, Klasik Test Kuramı'na dayalı yöntemlerden MH ve LR yöntemleri kullanılmıştır. Daha sonraki yapılacak olan çalışmalarda, KTK'ya dayalı diğer yöntemler ve IRT'ye dayalı yöntemlerle DMF belirlenebilir. Farklı yöntemlerden elde edilecek sonuçlar karşılaştırılabilir.
2. Test geliştirme çalışmalarında, maddelerin konulara ve bilişsel güçlük düzeylerine göre farklı sıralarda yerleştirildiği test formlarının verilmesinin, maddelerde DMF yaratıp yaratmadığı belirlenmeye yönelik çalışmalar yapılabilir.
3. Testte güçlük düzeylerine göre çok kolay ve çok zor maddeler bulunduğu durumda, maddelerin güçlük düzeylerine göre farklı şekillerde (kolaydan, zora ve zordan kolaya) sıralandığı test formlarının verilmesinin, maddelerde DMF yaratıp yaratmadığı belirlenmeye çalışılabilir.
4. Maddelerin farklı şekillerde sıralanmasının çeşitli değişkenler (cinsiyet, kaygı, sosyoekonomik düzey) açısından DMF'ye neden olup olmadığını belirlemeye yönelik çalışmalar yapılabilir.
5. Maddelerin test içerisinde farklı sıralarda yerleştirildiği farklı test formlarının verilmesinin alt yetenek düzeyleri açısından DMF yaratıp yaratmadığını araştıran çalışmalar yapılabilir.



## KAYNAKÇA

- Aamodt, M. G. & Teige, M. (1992). A Meta-Analytic Investigation of the Effect of Various Test Item Characteristics on Test Scores and Test Completion Times. *Public Personel Management*, 21(2), 151- 161. Eriřim adresi: <http://ppm.sagepub.com/>
- Airasian, P. W. (1997). *Classroom assessment*. New York: McGraw Hill.
- Albano, A.D. (2013). Multilevel modeling of item position effects. *Journal of Educational Measurement*, 50(4), 408-426. doi: 10.1111/jedm.12026
- Allison, D. E. (1970). Test anxiety, stress, and intelligence-test performance. *Canadian Journal of Behavioural Science* , 2(1), 26-37. Eriřim adresi: <http://www.apa.org/pubs/journals/cbs/>
- Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In P. W. Holland, and H. Wainer (Eds.), *Differential item functioning* (pp. 3-23) . Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Aiken, L. R. (2000). *Psychological Testing and Assessment* ( 10 th ed.).Boston: Allyn and Bacon
- Atılgan, H. (2014). *Eđitimde ölçme ve deęerlendirme*. Ankara: Anı.
- Ayan. C. (2010). *PISA 2009 fen okuryazarlıęı alt testinin deęiřen madde fonksiyonu aęısından incelenmesi*. Yüksek lisans tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Barcikovski, R. S. & Olsen, H. (1975). Test Item Arrangement and Adaptation Level. *The Journal of Psychology*, 90(1), 87-93. doi: 10.1080/00223980.1975.9923929
- Baykul, Y. (2010). *Eđitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: Pegem.
- Benito, J. G. & Ara, M. J. N. (2000). A comparison of  $\chi^2$ , RFA and IRT based procedures in the detection of DIF.

*Quality ve Quantity*, 34(1), 17-31. Erişim adresi: <http://link.springer.com/journal/11135>

Bertrand, R. & Boiteau, N. (2003). Comparing the stability of IRT-Based and non IRT-based DIF methods in different cultural contexts using TIMSS data (Eric Document Reproduction No. ED 476 924).

Brenner, M. H. (1964). Test Difficulty, Reliability, and Discriminations of Item Difficulty Order. *Journal of Applied Psychology*, 48(2), 98-100. doi: 10.1037/h0045738

Bulut, O. (2015). An Empirical Analysis of Gender-Based DIF Due To Test Booklet Effect. *European Journal of Research on Education*, 3(1), 7-16. Erişim adresi: <http://iassr2.org/rs/030102.pdf>

Camilli, G. (2006). *Test fairness*. In Brennan, R. L. (Ed). Educational Measurement (p. 221–257) . CT: Wesport. American Council on Education.

Camilli, G. & Shepard, L. A. (1994). *Methods for Identifying Biased Test Items*. Hollywood: Sage.

Carlson, J. L. & Ostrosky, A. L. (1992). Item Sequence and Student Performance on Multiple-Choice Exams: Further Evidence . *Journal of Economic Education*, 2(3), 232-235. doi: 10.2307/1183225 Chiu, P.C., & Irwin, P.M. (2011). Chronological Item Ordering: Does It Make a Difference on a State History and Government Assessment? Paper presented at the Annual Meeting of the American Educational Research Association, Kansas. Erişim adresi: [https://cete.ku.edu/sites/cete.drupal.ku.edu/files/docs/Presentations/2011/History\\_Item\\_Order\\_7-12-2011.pdf](https://cete.ku.edu/sites/cete.drupal.ku.edu/files/docs/Presentations/2011/History_Item_Order_7-12-2011.pdf)

Chiu, P. (2012) . The Effect Of Item Position On State Mathematics Assessment. Paper presented at the Annual Meeting of the American Educational Research Association, Kansas. Erişim adresi: [https://aai.ku.edu/sites/cete.ku.edu/files/docs/Presentations/2012\\_04\\_ChIU\\_Math\\_Item\\_Ordering\\_AERA\\_2012.pdf](https://aai.ku.edu/sites/cete.ku.edu/files/docs/Presentations/2012_04_ChIU_Math_Item_Ordering_AERA_2012.pdf)

Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Ohio: Cengage Learning.

Cronbach, J. L. (1990). *Essentials of Psychological Testing*. New York: Harper & Row.

- Clauser, B. E. & Mazor, K. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement, Issues and Practice*, 17(1), 31–44. doi: 10.1111/j.1745-3992.1998.tb00619.x
- Clauser, B. E., Mazor, K. & Hambleton, R. K. (1993). The effects of purification of the matching criterion on the identification of DIF using the Mantel-Haenszel procedure. *Applied Measurement in Education*, 6(4), 269–279. doi:10.1207/s15324818ame0604\_2
- Davis, J. & Ferdous, A. (2005). *Using item difficulty and item position to measure test fatigue*. Washington, DC: American Institutes for Research.
- Debeer, D. & Janssen R. (2013). Modeling Item-Position Effects Within an IRT Framework. *Journal of Educational Measurement*, 50(2), 164-185. doi: 10.1111/jedm.12009
- Doğan, N. ve Öğretmen, T. (2008). Değişen madde fonksiyonunu belirlemede Mantel- Haenszel, Ki-kare ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 33, 100-112. Erişim adresi: <http://egitimvebilim.ted.org.tr/index.php/EB>
- Dorans, N. J. & Holland, P.W. (1993). DIF detection and description: Mantel-Haenszel and Standardization. *Differential item functioning*. P.W. Holland & H.Wainer (Eds.), *Differential item functioning* (s. 35-66) içinde. Hillsdale, NJ: Lawrence Erlbaum .
- Eignor, D. R. & Cook, L. L. (1983, April). *An investigation of the feasibility of using item response theory in the pre-equating of aptitude tests* (Report No. TM 830 59). Paper presented at the meeting of the American Educational Research Association, Montreal, Canada. (ERIC Document Reproduction Service No. ED 235.197). Erişim adresi: <http://files.eric.ed.gov/fulltext/ED235197.pdf>
- Erkuş, A. (2003). *Psikometri Üzerine Yazılar*. Ankara: Türk Psikologlar Derneği.
- Gerow, J. R. (1980). Performance on Achievement Test as a Function of The Order of Item Difficulty. *Teaching of Psychology*, 7(2), 93-94. doi: 10.1207/s15328023top0702\_7
- Flaugher, R. L., Melton, R. S., & Myers, C. T. (1968). Item rearrangement under typical test conditions. *Educational and Psychological Measurement*, 28(3), 813-824. doi: 10.1177/001316446802800310

- French, J. L. & Greer, D. (1964). Effect of Test Item Arrangement on Physiological and Psychological Behavior in Primary-School Children. *Journal of Educational Measurement*, 1(2), 151- 153. doi: 10.1111/j.1745-3984.1964.tb00174.x
- Gierl, M., Khaliq, S. N., & Boughton, K. (1999, June). Gender differential item functioning in mathematics and science: Prevalence and policy implications. Paper presented at the Improving Large-Scale Assessment in Education Symposium at the Annual Meeting of the Canadian Society for the Study of Education, Canada. Erişim adresi: [http://www.education.ualberta.ca/educ/psych/crame/files/dif\\_csse99.pdf](http://www.education.ualberta.ca/educ/psych/crame/files/dif_csse99.pdf)
- Gomez-Benito, J. & Navas-Ara, M. J. (2000). A comparison of Ki-kare, RFA and IRT based procedures in the detection of DIF. *Quality ve Quantity*, 34(1),17–31. Erişim adresi: <http://link.springer.com/journal/11135>
- Gondal, M. B.(2001). *Differential item functioning analysis of 4Th graders' science and urdu (national language) achievement test items in Pakistan*. Unpublished doctoral dissertation, Middle East Technical University, Ankara.
- Gök, B., Kelecioğlu, H., ve Doğan, N. (2010). Değişen madde fonksiyonunu belirlemede Mantel Haenszel ve Lojistik Regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim Dergisi*, 35(156), 3-16. doi: 10.21031/epod.71099
- Hambleton, K. R. & Rogers, J. H. (1989). Detecting potentially biased test items: comparison of IRT area and MH methods. *Applied Measurement in Education*, 2(4), 313–334.Erişim adresi: <http://www.tandfonline.com/loi/hame20#.V4pFRtKLTIU> Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. California: Sage.
- Hambleton, R. K., & Traub, R. E. (1974). The Effects of Item Order on Test Performance and Stres. *Journal of Experimental Education*, 43(1), 40-46. Erişim adresi: <http://eric.ed.gov/?id=ED017960>
- Hidalgo, M. D., & Lopez-Pina, J. A. (2004). Differential item functioning detection an effect size: A comparison between logistic regression and Mantel-Haenszel procedures. *Educational and Psychological Measurement*, 64(6), 903-915. doi: 10.1177/0013164403261769

- Hodson, D. (1984). The Effect of Changes in Item Sequence on Student Performance in a Multiple-Choice Chemistry Test. *Journal of Research in Science Teaching*, 21(5), 489-495. doi: 10.1002/tea.3660210506
- Hohensinn, C., Kubinger, K. D., Reif, M., Schleicher, E., & Khorramdel, L. (2011). Analysing item position effects due to test booklet design within large-scale assessment. *Educational Research and Evaluation*, 17(6), 497-509. doi: 10.1080/13803611.2011.632668
- Holland, P. W. & Thayer, D. T. (1988). *Differential item functioning detection and the Mantel-Haenszel procedure*. In H. Wainer and H. Braun (Eds.), *Test Validity* (pp. 129–145) içinde. Hillsdale, NJ: Lawrence Erlbaum.
- Holland, P. W. & Wainer, H. (1993). *Differential Item Functioning*. Hillside, New Jersey: Lawrence Erlbaum.
- Huck, S. W. & Bowers, N. D. (1972). Item Difficulty Level and Sequence Effects in Multiple Choice Achievement Tests. *Journal of Educational Measurement*, 9(2), 105-111. Erişim adresi: <http://www.jstor.org/stable/1433800>
- Jodoin, M. G. & Gierl, M.J. (2001). Evaluating Type I error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14(4), 329-349. doi:10.1207/S15324818AME1404\_2
- Karami H. & Nodoushan M. A. S. (2011). Differential Item Functioning (DIF): Current Problems and Future Directions. *International Journal of Language Studies*, 5(3), 133-142. Erişim adresi: <http://eric.ed.gov/?id=ED521872>
- Kestenbaum, J. M. & Weiner, B. (1970). Achievement performance related to achievement motivation and test anxiety. *Journal of Consulting and Clinical Psychology*, 34(3), 343-344. Erişim adresi: <http://www.apa.org/pubs/journals/ccp/>
- Kingston, N. M. & Dorans, N. J. (1984). Item location effects and their implications for IRT equating and adaptive testing. *Applied Psychological Measurement*, 8(2), 147–154. doi: 10.1177/014662168400800202

- Kleinke, D. J. (1980). Item Order, Response Location, and Examinee Sex and Handedness on Performance on Multiple-Choice Tests. *Journal of Educational Research*, 73(4), 225–229. doi:10.1080/00220671.1980.10885240
- Klimko, I. P. (1984). Item arrangement, cognitive entry characteristics, sex, and test anxiety as predictors of achievement examination performance. *Journal of Experimental Education*, 52(4), 214-219. doi: 10.1080/00220973.1984.11011896
- Klosner, N. C. & Gellman, E. K. (1973). The Effect of Item Arrangement on Classroom Test Performance: Implications for Content Validity. *Educational and Psychological Measurement*, 33(2), 413-418. doi: 10.1177/001316447303300224
- Kubinger, K. D. (2008). On the revival of the Rasch model-based LLTM: From composing tests by item generating rules to measuring item administration effects. *Psychology Science Quarterly*, 50(3), 311–327. Erişim adresi: <http://www.psychologie-aktuell.com/index.php?id=204>
- Laffittee, R. G. (1984). Effects of item order on achievement test scores and students' perceptions of test difficulty. *Teaching of Psychology*, 11(4), 212 - 214. doi: 10.1177/009862838401100405
- Leary, L. F. & Dorans, N. J. (1982). The Effects of Item Rearrangement on Test Performance: The Review of the Literature, Princeton NJ: Educational Testing Service. doi: 10.1002/j.2333-8504.1982.tb01316.x
- Leary, L. F. & Dorans, N. J. (1985). Implications for Altering the Context In Which Test Items Appear: A Historical Perspective on Immediate Concern. *Review of Educational Research*, 55(33), 387-413. doi: 10.3102/00346543055003387
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass.: Addison-Welsley Publishing Company.
- Louisa, N. (2013). Effect of Item Arrangement on Test Reliability Coefficients: Implications for Testing. *Journal of Research in Education and Society*, 4(3), 54-62. Erişim adresi: <http://www.journalres.com/eboard.php>

- MacNichol, K. (1960). *Effects of varying order of item difficulty in an unspeeded verbal test*. Unpublished manuscript, Princeton, NJ: Educational Testing Service.
- Magnusson, D. (1967). *Test Theory*. Massachusetts: Addison-Wesley
- Marso, N. R. (1970). Test Item Arrangement, Testing Time, and Performance. *Journal of Educational Measurement*, 7(2), 113-118. doi: 10.1111/j.1745-3984.1970.tb00704.x
- Mazor, K. M., Clauser, R. E. & Hambleton, R. K. (1994). Identification of nonuniform differential item functioning using a variation of the Mantel-Haenszel procedure. *Educational and Psychological Measurement*, 54(2), 284–291. doi: 10.1177/0013164494054002003
- McNamara, T. & Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell
- Miller, S. K. (1989). *Interaction effects of gender and item arrangement on test and item performance* (Doktora tezi) ProQuest Dissertations and Thesis veri tabanından erişildi. (UMI No. 303780321)
- Mollenkopf, W. G. (1950). An experimental study of the effects on item-analysis data of changing item placement and test time limit. *Psychometrika*, 15(3), 291–315. doi: 10.1007/BF02289044
- Narayanan, P., & Swaminathan, H. (1994). Performance of the Mantel-Haenszel and simultaneous item bias procedures for detecting differential item functioning. *Applied Psychological Measurement*, 18(4), 315-338. doi: 10.1177/014662169401800403
- Narayanan, P., & Swaminathan, H. (1996). Identification of items that show non-uniform DIF. *Applied Psychological Measurement*, 20(3), 257–274. doi: 10.1177/014662169602000306
- Newman, D. L, Kundert, D. K, Lane D. S. & Bull K. S. (1988). Effect of Varying Item Order on Multiple-Choice Test Scores: Importance of Statistical and Cognitive Difficulty. *Applied Measurement in Education*, 1 (1), 89-97. doi: 10.1207/s15324818ame0101\_8
- Ollennu, S. N. N.& Etsey, Y. K. A. (2015). The Impact of Item Position in Multiple-Choice Test on Student Performance at the Basic Education Certificate Examination (BECE) Level, *Universal Journal of Educational Research*, 3(10),718-723 . doi: **10.13189/ujer.2015.031009**

Osterlind, S. J. (1983). *Test item bias*. Beverly Hills and London: Sage.

Osterlind, S. ve H. Everson (2009). *Differential item functioning*. Londra: Sage.

Pang, X. L. and et all. (1994, April). Performance of Mantel-Haenszel and Logistic Regression DIF

Procedures over Replications Using Real Data. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. Erişim adresi: <http://eric.ed.gov/?id=ED374153>

Penfield, R. D. & Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao & S. Sinharay (Eds.), *Handbook of Statistics Psychometrics* (26, pp. 125–167). Amsterdam: Elsevier.

Plake, B. (1980). Item Arrangement and Knowledge of Arrangement on Test Scores. *Journal of Experimental Education*, 49(1), 56-58. Erişim adresi: <http://www.jstor.org/stable/20151376>

Plake, B. S., Ansorge, C. J., Parker, C. S. & Lowry, S. R. (1982). Effects of item arrangement, knowledge of arrangement, test anxiety, and sex on test performance. *Journal of Educational Measurement*, 19(1), 49–57. doi: 10.1111/j.1745-3984.1982.tb00114.x

Plake, B. S., Melican, G. J., Carter, L. & Shaughnessy, M. (1983). Differential performance of males and females on easy to hard item arrangements: Influence of feedback at the item level. *Educational and Psychological Measurement*, 43(4), 1067-1075. doi: 10.1177/001316448304300416

Ryan, K. E. & Chiu, S. (2001). An examination of item context effects, DIF, and gender DIF. *Applied Measurement in Education*, 14 (1), 73–90. doi:10.1207/S15324818AME1401\_06

Sax, G. & Cromack, T. A. (1966). The effects of various forms of item arrangements on test performance. *Journal of Educational Measurement*, 3 (4), 309-311. doi: 10.1111/j.1745-3984.1966.tb00896.x

Smouse, A. D. & Munz, D. C. (1968). The Effects of Anxiety and Item Difficulty Sequence on Achievement Testing Scores. *The Journal of Psychology*, 68(2), 181-184. doi: 10.1080/00223980.1968.10543421



- Swaminathan, H. & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361-370. doi: 10.1111/j.1745-3984.1990.tb00754.x
- Tan, Ş. (2009). Soru Sırasının Madde Güçlüğü ve Ayırıcılık Gücüne Etkisi. *e-Journal of New World Sciences Academy*, 4(2), 486-493. Erişim adresi: [http://www.newwsa.com/download/gecici\\_makale\\_dosyalari/NWSA-1091-3-2.pdf](http://www.newwsa.com/download/gecici_makale_dosyalari/NWSA-1091-3-2.pdf)
- Tan, Ş. (2013). *Öğretimde ölçme ve değerlendirme: KPSS el kitabı* (Sekizinci Baskı). Ankara: Pegem.
- Tippets, E. & Benson, J. (1989). The Effect of Item Arrangement on Test Anxiety. *Applied Measurement in Education*, 2(24), 289-296. doi:10.1207/s15324818ame0204\_2
- Towle, N. J. & Merrill, P. F. (1975). Effects of anxiety type and item difficulty sequencing on mathematics test performance. *Journal of Educational Measurement*, 12(4) , 241-249. Erişim adresi: <http://www.jstor.org/stable/1434151>
- Turgut, M.F. (1995). *Eğitimde ölçme ve değerlendirme metotları* (Onuncu Baskı). Ankara: Yargıcı.
- Whitely, S. E. & Dawis, R. V. (1976). The influence of test context on item difficulty. *Educational and Psychological Measurement*, 36(2), 329-337. doi: 10.1177/001316447603600211
- Wiberg, M. (2007). *Measuring and detecting differential item functioning in criterion-referenced licensing test: a theoretic comparison of methods* (EM No. 60). Umea, Sweden: Umea University, Department of Educational Measurement.
- Yergin, F. (2007). *Çoktan Seçmeli Testlerde Madde Sıralamanın Madde Ve Test İstatistikleri Üzerine Etkisi*, (Yüksek lisans Tezi) Mersin Üniversitesi, Sosyal Bilimler Enstitüsü, Mersin.
- Yen, W. M. (1980). The extent, causes and importance of context effects on item parameters for two latent trait models. *Journal of Educational Measurement*, 17(4), 297-311. doi: 10.1111/j.1745-3984.1980.tb00833.x

- Yılmaz, H. B. (2014,Haziran). *Madde Ve Seçenek Sıralamasının Değişen Madde Fonksiyonuna Etkileri* [Öz]. IV. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresinde sunulan bildiri, Hacettepe Üniversitesi, Ankara. Erişim adresi:<http://epod-online.org/wp-content/uploads/2014/07/IV.-Ulusal-E%C4%9Fitimde-ve-Psikolojide-%C3%96l%C3%A7me-ve-De%C4%9Ferlendirme-Kongresi.pdf>
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P.W. Holland ve H. Wainer (Eds.), *Differential item functioning* (pp. 337-347) içinde. Hillsdale NJ: Erlbaum.
- Zumbo, B. D. A. (1999). *Handbook on the theory and methods of differential item functioning: Logistic regression modeling as a unitary framework for binary and likert-type item scores*. Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense. (<http://www.educ.ubc.ca/faculty/zumbo/DIF/handbook.pdf>)
- Zumbo, B. D. (2009). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF) Scores*. Ottawa: ON: Directorate of Human Resources Research and Education, Department of National Defense.
- Zumbo, B. D., Perlino, A. H. & Lind, D. L. (1998). Context Effects on Examinations: The Effects of Time, Item Order, Item Difficulty. *Canadian Psychology*, 39(4), 299-314.Erişim adresi: <http://www.apa.org/pubs/journals/cap/>
- Zumbo, B. D. & Thomas, D. R. (1997). *A measure of effect size for a model-based approach for studying DIF* (Working paper of the Edgeworth Laboratory for Quantitative Behavioral Science). Prince George, Canada: University of Northern British Columbia.

## EKLER

### EK: 1: Uzman Görüşleri Örnek Madde Bilgi Formu

Değerli Uzman,

“Maddelerin Test İçerisinde Güçlük Düzeylerine Göre Farklı Şekillerde Sıralanmasının Etkisinin Değişen Madde Fonksiyonuyla İncelenmesi” konulu yüksek lisans tez çalışmam kapsamında Mersin Üniversitesi Meslek Yüksek Okullarında Temel Matematik Dersi alan öğrencilerin Sayılar Ünitesine ait Alt Öğrenme Alanı olan “Kareköklü Sayılar ve Kareköklü Sayılarla Dört İşlem” Ünitesindeki Matematik becerilerini ortaya koymak amacıyla çoktan seçmeli paralel test geliştirme çalışması yapılmaktadır. Bu amaçla Kareköklü Sayılar ve Kareköklü Sayılarla Dört İşlem becerisini ölçtüğü düşünülen testler için belirli sayıda soru geliştirilmiştir. Çalışmanın sürdürülebilmesi için geliştirilen soruların bazı ölçütler bakımından incelenmesine ihtiyaç duyulmaktadır.

Bu form ekte yer alan “Kareköklü Sayılar ve Kareköklü Sayılarla Dört İşlem Ünitesi Başarı Testlerinin ” maddelerinin belirlenen ölçütler doğrultusunda tarafınızca değerlendirilmesi ve görüşlerinize dayanarak bir takım düzenlemeler yapılması amacıyla hazırlanmıştır. Sizden istenen testlerde yer alan maddeleri Form 1 de yer alan ifadeler doğrultusunda çarpı koyarak (X) değerlendirmenizdir (evet, hayır, düzeltme gerekli). Düzeltme gerektiğini düşündüğünüz maddeler için ekte yer alan test formları üzerine açıklama yazabilirsiniz.

İlginiz ve değerli katkılarınız için teşekkür ederim.

Arş. Gör. Ebru BALTA

Mersin Üniversitesi

Eğitim Fakültesi

Eğitimde Ölçme ve Değerlendirme A.B.D



**Ek 2: Uzmanlar Arası Uyum Veri Toplama Formu**

Sayın Matematik Eğitimi Uzmanı/Ölçme ve Değerlendirme Uzmanı,

Sizden istenen Test 1 ve Test 2’de yer alan her maddeyi karşılaştırıp testlerin her bir maddesinin paralel olup olmadığını belirlemenizdir.

<b>Madde Numarası</b>	<b>1 İle 2</b>	
	<b>Paralel</b>	<b>Paralel Değil</b>
<b>1</b>		
<b>2</b>		
<b>3</b>		
<b>4</b>		
<b>5</b>		
<b>6</b>		
<b>7</b>		
<b>8</b>		
<b>9</b>		
<b>10</b>		
<b>11</b>		
<b>12</b>		
<b>13</b>		
<b>14</b>		
<b>15</b>		
<b>16</b>		
<b>17</b>		
<b>18</b>		
<b>19</b>		
<b>20</b>		

**Ek 3: 40 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testinin (Test 1) Madde Analizleri**

Madde Numarası	Test 1		
	Madde Güçlüğü	Madde Ayırt ediciliği	Madde Standart Sapması
1	0,68	0,84	0,46
2	0,55	0,63	0,50
3	0,59	0,78	0,49
4	0,60	0,78	0,49
5	0,43	0,59	0,50
6	0,46	0,76	0,50
7	0,18	-0,18	0,38
8	0,66	0,79	0,47
9	0,46	0,72	0,50
10	0,73	0,72	0,44
11	0,57	0,79	0,50
12	0,44	0,77	0,50
13	0,51	0,75	0,50
14	0,30	0,52	0,46
15	0,52	0,74	0,50
16	0,42	0,84	0,49
17	0,51	0,51	0,50
18	0,45	0,73	0,50
19	0,44	0,72	0,50
20	0,11	-0,11	0,31
21	0,13	-0,13	0,33
22	0,49	0,76	0,50
23	0,62	0,77	0,49
24	0,43	0,72	0,50
25	0,47	0,74	0,50
26	0,55	0,75	0,49
27	0,44	0,70	0,50
28	0,41	0,69	0,49
29	0,51	0,71	0,50
30	0,58	0,74	0,49
31	0,58	0,68	0,49
32	0,47	0,79	0,50
33	0,38	0,76	0,49
34	0,52	0,75	0,50
35	0,44	0,74	0,50
36	0,48	0,82	0,50
37	0,50	0,71	0,50
38	0,53	0,79	0,50
39	0,49	0,76	0,50
40	0,14	-0,14	0,34

## Ek 4: 20 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testi (Test 1)

### KAREKÖKLÜ SAYILAR VE KAREKÖKLÜ SAYILARLA DÖRT İŞLEM ÜNİTESİ TESTİ

**Kazanım 1:** Kareköklü sayılarla toplama ve çıkarma işlemi yapar.

1) Hale ile Erkan her akşam yürüyüş yapmaktadır. Bu iki kişi yürüyüşe başladıktan sonra Hale  $1875\sqrt{2}$  m, Erkan ise  $1500\sqrt{2}$  m yol alıyor. Buna göre, bu iki kişinin akşam yürüyüşünde aldığı toplam yol kaç m'dir?

- A)  $2675\sqrt{2}$     B)  $2775\sqrt{2}$     C)  $3375\sqrt{2}$   
D)  $2875\sqrt{3}$     E)  $3000\sqrt{6}$

**Kazanım 2:** Tam kare doğal sayılarla bu sayıların kareköklü arasında ilişkiyi modelleriyle açıklar ve kareköklerini belirler.

2) Aşağıdaki sayılardan hangisi tam karedir?

- A) 121    B) 129    C) 133    D) 149    E) 194

3) 5, 8, 14, 16, 26, 38, 49, 121, 144 sayılarından kaç tanesi tam kare sayıdır?  
A) 4    B) 5    C) 6    D) 7    E) 8

**Kazanım 1:** Kareköklü sayılarla toplama ve çıkarma işlemi yapar.

4)  $3\sqrt{5} + 2\sqrt{5} + 5\sqrt{5} - 2\sqrt{5}$  işleminin sonucu kaçtır?  
A)  $3\sqrt{5}$     B)  $4\sqrt{5}$     C)  $7\sqrt{5}$   
D)  $8\sqrt{5}$     E)  $9\sqrt{5}$

**Kazanım 3:** Kareköklü bir sayıyı  $a\sqrt{b}$  şeklinde yazar ve  $a\sqrt{b}$  şeklindeki ifadede katsayıyı kök içine alır.

5)  $\sqrt{48}$  sayısı aşağıdakilerden hangisine eşittir?  
A)  $2\sqrt{3}$     B)  $3\sqrt{3}$     C)  $4\sqrt{3}$   
D)  $5\sqrt{3}$     E)  $6\sqrt{3}$

**Kazanım 4:** Kareköklü sayıları karşılaştırır ve sıralar.

6) Aşağıda kenar uzunlukları verilen dikdörtgenlerden hangisinin alanı en büyüktür?  
A)  $\sqrt{5}$  cm,  $2\sqrt{5}$  cm    B)  $3\sqrt{2}$  cm,  $4\sqrt{2}$  cm  
C)  $4\sqrt{3}$  cm,  $3\sqrt{3}$  cm    D)  $4\sqrt{6}$  cm,  $2\sqrt{6}$  cm  
E)  $5\sqrt{7}$  cm,  $2\sqrt{7}$  cm

**Kazanım 5:** Paydasında köklü ifade bulunan bir kesrin paydasını kökten kurtararak kesrin paydasını rasyonel yapar ve rasyonel sayıların karekökünü alır.

7)  $\frac{1}{\sqrt{6}-\sqrt{7}} - \frac{1}{\sqrt{6}+\sqrt{7}}$  işleminin sonucu aşağıdakilerden hangisidir?  
A)  $-\sqrt{12}$     B)  $-\sqrt{26}$     C)  $-2\sqrt{7}$   
D)  $-5\sqrt{7}$     E)  $-\sqrt{46}$

**Kazanım 3:** Kareköklü bir sayıyı  $a\sqrt{b}$  şeklinde yazar ve  $a\sqrt{b}$  şeklindeki ifadede katsayıyı kök içine alır.

8) Aşağıda verilen eşitliklerden hangisi doğrudur?

- A)  $\sqrt{18} = 9\sqrt{2}$       B)  $\sqrt{25} = 5\sqrt{5}$   
 C)  $\sqrt{30} = 3\sqrt{10}$       D)  $\sqrt{45} = 3\sqrt{5}$   
 E)  $\sqrt{60} = 4\sqrt{15}$

**Kazanım 6:** Ondalık kesirlerin kareköklerini belirler.

9)  $\frac{\sqrt{0,09} + \sqrt{0,04}}{\sqrt{0,36} - \sqrt{0,01}}$  işleminin sonucu kaçtır?

- A)  $\frac{1}{7}$       B)  $\frac{1}{5}$       C) 1      D) 2      E) 3

**Kazanım 2:** Tam kare doğal sayılarla bu sayıların kareköklü arasında ilişkiyi modelleriyle açıklar ve kareköklerini belirler.

10) 24 ile 85 arasında kaç tane tam kare sayı vardır?

- A) 3      B) 4      C) 5      D) 6      E) 7

**Kazanım 1:** Kareköklü sayılarla toplama ve çıkarma işlemi yapar.

11)  $\sqrt{500} - \sqrt{5}$

Yukarıda verilen işlemin sonucu kaçtır?

- A)  $7\sqrt{5}$       B)  $8\sqrt{5}$       C)  $9\sqrt{5}$   
 D)  $10\sqrt{5}$       E)  $11\sqrt{5}$

**Kazanım 7:** Kök dışına köklü bir ifadenin mutlak değerini çıkarır ve köklü ifadeyi üslü ifade olarak yazar.

12)  $\sqrt{(-3)^2} - \sqrt{3^2} - (-2)^3$  işleminin sonucu kaçtır?

- A) -17      B) -9      C) 0      D) 6  
 E) 8

**Kazanım 5:** Paydasında köklü ifade bulunan bir kesrin paydasını kökten kurtararak kesrin paydasını rasyonel yapar ve rasyonel sayıların karekökünü alır.

13)  $\sqrt{\frac{1}{81}} + \sqrt{\frac{1}{16}} - \sqrt{\frac{1}{36}}$  işleminin sonucu kaçtır?

- A)  $\frac{7}{4}$       B)  $\frac{7}{8}$       C)  $\frac{7}{16}$       D)  $\frac{7}{24}$       E)  $\frac{7}{36}$

**Kazanım 3:** Kareköklü bir sayıyı  $a\sqrt{b}$  şeklinde yazar ve  $a\sqrt{b}$  şeklindeki ifadede katsayıyı kök içine alır.

14)  $-4\sqrt{6}$

sayısı aşağıdakilerden hangisine eşittir?

- A)  $\sqrt{-96}$       B)  $\sqrt{96}$       C)  $-\sqrt{96}$

- D)  $-\sqrt{24}$       E)  $-\sqrt{12}$

15) Aşağıdaki işlemlerden hangisi  $6\sqrt{3}$  'e eşit değildir?

- A)  $2\sqrt{27}$       B)  $3\sqrt{12}$       C)  $3 \cdot 2\sqrt{3}$

- D)  $\sqrt{108}$       E)  $\sqrt{8} \cdot 4\sqrt{2}$

**Kazanım 1:** Kareköklü sayılarla toplama ve çıkarma işlemi yapar.

16)  $4\sqrt{28} + 3\sqrt{7} - (\sqrt{28} + \sqrt{7})$  işleminin sonucu kaçtır?

- A)  $2\sqrt{7}$       B)  $4\sqrt{7}$       C)  $6\sqrt{7}$

- D)  $8\sqrt{7}$       E)  $10\sqrt{7}$



**Kazanım 6:** Ondalık kesirlerin kareköklerini belirler.

17)  $K = \sqrt{0,09}$  ve  $L = \sqrt{0,01}$  olduğuna göre,

$\frac{K + L}{K \times L}$  işleminin sonucu kaçtır?

- A)  $\frac{5}{3}$     B)  $\frac{10}{3}$     C)  $\frac{25}{3}$     D)  $\frac{40}{3}$   
E)  $\frac{90}{3}$

**Kazanım 5:** Paydasında köklü ifade bulunan bir kesrin paydasını kökten kurtararak kesrin paydasını rasyonel yapar ve rasyonel sayıların karekökünü alır.

18)  $4^2 \times \left(\frac{1}{\sqrt{4}}\right)^4$  işlemin sonucu kaçtır?

- A)  $\frac{1}{16}$     B)  $\frac{1}{4}$     C) 1    D) 4    E) 16

**Kazanım 8:** Kareköklü sayılarla çarpma ve bölme işlemi yapar.

19) Eni  $4\sqrt{5}$  br, boyu  $3\sqrt{5}$  br olan bir dikdörtgenin alanı kaç  $\text{br}^2$ 'dir?

- A)  $6\sqrt{5}$     B)  $7\sqrt{5}$     C)  $12\sqrt{5}$     D) 12  
E) 60

**Kazanım 2:** Tam kare doğal sayılarla bu sayıların kareköklü arasında ilişkiyi modelleriyle açıklar ve kareköklerini belirler.

20) Aşağıda ifade edilen sayılardan hangisi 625'in kareköküne eşittir?

- A) 5'in karesi    B) 15'in karesi  
C) 125'in karesi    D) 25'in karesi  
E) 5'in küpü

**Ek 5: 40 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testlerinin (Test 2-Test 3) Madde Analizleri**

Madde Numarası	Madde Güçlüğü		Madde Ayırt ediciliği		Madde Standart Sapması	
	Test 2	Test 3	Test 2	Test 3	Test 2	Test 3
1	0,66	0,60	0,67	0,79	0,48	0,49
2	0,60	0,44	0,63	0,77	0,49	0,50
3	0,67	0,65	0,66	0,66	0,47	0,48
4	0,69	0,51	0,76	0,81	0,46	0,50
5	0,54	0,55	0,70	0,71	0,50	0,50
6	0,45	0,48	0,75	0,75	0,50	0,50
7	0,38	0,39	0,77	0,72	0,49	0,49
8	0,77	0,75	0,46	0,77	0,42	0,43
9	0,50	0,55	0,72	0,70	0,50	0,50
10	0,71	0,71	0,54	0,86	0,45	0,45
11	0,57	0,56	0,73	0,81	0,49	0,50
12	0,51	0,50	0,57	0,63	0,50	0,50
13	0,51	0,51	0,78	0,74	0,50	0,50
14	0,21	0,29	0,64	0,53	0,41	0,46
15	0,56	0,55	0,72	0,83	0,50	0,50
16	0,32	0,53	0,78	0,82	0,47	0,50
17	0,51	0,43	0,42	0,49	0,50	0,49
18	0,32	0,34	0,56	0,58	0,47	0,47
19	0,34	0,37	0,74	0,71	0,48	0,48
20	0,56	0,51	0,69	0,75	0,50	0,50
21	0,44	0,53	0,77	0,73	0,50	0,50
22	0,43	0,51	0,75	0,77	0,49	0,50
23	0,70	0,70	0,76	0,86	0,46	0,46
24	0,51	0,37	0,51	0,75	0,50	0,48
25	0,50	0,60	0,75	0,69	0,50	0,49
26	0,45	0,49	0,70	0,73	0,50	0,50
27	0,47	0,49	0,64	0,77	0,50	0,50
28	0,46	0,49	0,80	0,75	0,50	0,50
29	0,53	0,39	0,75	0,67	0,50	0,49
30	0,56	0,55	0,77	0,78	0,50	0,50
31	0,55	0,56	0,66	0,79	0,50	0,50
32	0,38	0,49	0,70	0,70	0,48	0,50
33	0,40	0,45	0,74	0,73	0,49	0,50
34	0,41	0,47	0,78	0,84	0,49	0,50
35	0,44	0,42	0,79	0,80	0,50	0,49
36	0,57	0,55	0,80	0,80	0,49	0,50
37	0,55	0,60	0,64	0,68	0,50	0,49
38	0,60	0,56	0,78	0,79	0,49	0,50
39	0,53	0,51	0,78	0,80	0,50	0,50
40	0,58	0,55	0,72	0,78	0,49	0,50

## Ek 6: 20 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testi (Test 2)

### KAREKÖKLÜ SAYILAR VE KAREKÖKLÜ SAYILARLA DÖRT İŞLEM TESTİ

1) 4, 12, 15, 25, 36, 49, 111, 120, 144 sayılarından kaç tanesi tam kare sayıdır?

- A)5 B)6 C)7 D)8 E)9

2) Mehmet ile Hakan her sabah yürüyüş yapmaktadır. Bu iki kişi yürüyüşe başladıktan sonra Mehmet  $1275\sqrt{3}$  m, Hakan ise  $1500\sqrt{3}$  m yol alıyor. Buna göre, bu iki kişinin sabah yürüyüşünde aldığı toplam yol kaç m' dir?

- A)  $1775\sqrt{3}$  B)  $2775\sqrt{2}$  C)  $1775\sqrt{6}$   
D)  $2875\sqrt{3}$  E)  $3000\sqrt{3}$

3)  $3\sqrt{2} + 5\sqrt{2} + 8\sqrt{2} - 12\sqrt{2}$  işleminin sonucu kaçtır?

- A)  $\sqrt{2}$  B)  $2\sqrt{2}$  C)  $4\sqrt{2}$   
D)  $6\sqrt{2}$  E)  $8\sqrt{2}$

4) Alanı  $529 \text{ m}^2$  olan kare şeklindeki bir bahçenin bir kenarının uzunluğu kaç m 'dir?

- A)13 B)23 C) 33 D)43 E)53

5) Aşağıdaki sayılardan hangisi tam karedir?

- A) 200 B)222 C)225 D)250 E)275

6) Aşağıda verilen eşitliklerden hangisi doğrudur?

- A)  $\sqrt{36} = 6\sqrt{6}$  B)  $\sqrt{24} = 6\sqrt{4}$   
C)  $\sqrt{12} = 4\sqrt{3}$  D)  $\sqrt{16} = 4\sqrt{4}$   
E)  $\sqrt{45} = 3\sqrt{5}$

7)  $\sqrt{0,04} + \sqrt{0,49}$  işleminin sonucu kaçtır?

- A) 0,05 B) 0,07 C) 0,9  
D) 1,4 E) 1,9

8)  $\sqrt{12 + \sqrt{11 + \sqrt{23 + \sqrt{3 + \sqrt{1}}}}}$  işleminin sonucu kaçtır?

- A)1 B)2 C)3 D)4 E)5

9)  $\underbrace{\sqrt{5} + \sqrt{5} + \dots + \sqrt{5}}_{6TANE\sqrt{5}}$  işleminin sonucu kaçtır?

- A)  $\sqrt{30}$  B)  $5\sqrt{6}$  C)  $6\sqrt{5}$  D)  
 $6 + \sqrt{5}$  E)36

10)  $3\sqrt{20} + 2\sqrt{5} - (\sqrt{20} + \sqrt{5})$  işleminin sonucu kaçtır?

- A)  $\sqrt{5}$  B)  $2\sqrt{5}$  C)  $3\sqrt{5}$   
D)  $4\sqrt{5}$  E)  $5\sqrt{5}$

11)  $\sqrt{300} - \sqrt{3}$

Yukarıda verilen işlemin sonucu kaçtır?

- B)  $7\sqrt{3}$  B)  $8\sqrt{3}$  C)  $9\sqrt{3}$   
D)  $10\sqrt{3}$  E)  $11\sqrt{3}$

12) Aşağıdaki işlemlerden hangisi  $3\sqrt{8}$ 'e eşit değildir?

- A)  $\sqrt{72}$       B)  $6\sqrt{2}$       C)  $2\sqrt{18}$   
 D)  $2 \cdot 2\sqrt{3}$       E)  $\sqrt{9} \cdot 2\sqrt{2}$

13)  $a=\sqrt{42}$ ,  $b=7$ ,  $c=\sqrt{45}$ ,  $d=\sqrt{40}$

sayıları aşağıdaki seçeneklerin hangisinde doğru olarak sıralanmıştır?

- A)  $b < d < a < c$       B)  $b < a < d < c$   
 C)  $d < b < a < c$       D)  $d < a < b < c$   
 E)  $d < a < c < b$

14) Aşağıdaki ifadelerden hangisi yanlıştır?

- A)  $\sqrt[6]{5} = 5^{1/6}$       B)  $\sqrt{(-2)^6} = 2^3$   
 C)  $\sqrt[3]{4^7} = 4^{7/3}$       D)  $\sqrt[8]{8^3} = 8^{3/8}$   
 E)  $\sqrt{6^9} = 6^{2/9}$

15)  $\frac{\sqrt{0,04} + \sqrt{0,16}}{\sqrt{0,81} - \sqrt{0,25}}$  işleminin sonucu kaçtır?

- A)  $\frac{3}{4}$       B) 1      C)  $\frac{3}{2}$       D) 2      E) 3

16) Eni  $3\sqrt{6}$  br, boyu  $2\sqrt{6}$  br olan bir dikdörtgenin alanı kaç  $br^2$ 'dir?

- A)  $4\sqrt{6}$       B)  $5\sqrt{6}$       C)  $6\sqrt{6}$   
 D) 33      E) 36

17)  $\sqrt{0,12}$  ifadesi aşağıdakilerden hangisine eşittir?

- A)  $\frac{\sqrt{3}}{5}$       B)  $\frac{2\sqrt{3}}{5}$       C)  $\frac{3\sqrt{3}}{5}$

D)  $\frac{4\sqrt{3}}{5}$       E)  $\frac{6\sqrt{3}}{5}$

18)  $\sqrt{3}$  sayısı  $\sqrt{3} + 1$  sayısının kaç katıdır?

- A)  $\frac{3-\sqrt{3}}{2}$       B)  $\frac{1+\sqrt{3}}{2}$       C)  $\frac{3+\sqrt{3}}{2}$   
 D)  $\frac{4+\sqrt{3}}{2}$       E)  $\frac{6+\sqrt{3}}{2}$

19)  $\frac{\sqrt{8} + \sqrt{8} + \sqrt{8}}{\sqrt{2} \times \sqrt{2} \times \sqrt{2}}$  işleminin sonucu kaçtır?

- A) 2      B)  $2\sqrt{2}$       C) 3      D)  $3\sqrt{2}$       E) 4

20) Aşağıda ifade edilen sayılardan hangisi 64'ün kareköküne eşittir?

- A) 8'in karesi      B) 6'nın karesi  
 C) 16'in karesi      D) 4'ün karesi  
 E) 2'nin küpü

## Ek 7: 20 Maddelik Kareköklü Sayılar ve Kareköklü Sayılarla İşlemler Testi (Test 3)

### KAREKÖKLÜ SAYILAR VE KAREKÖKLÜ SAYILARLA DÖRT İŞLEM ÜNİTESİ TESTİ

1) Aşağıda ifade edilen sayılardan hangisi 81'in kareköküne eşittir?

- A) 3'ün karesi    B) En küçük tek asal sayı  
C) 9'un karesi    D) 81'in karesi  
E) 3'ün küpü

2)  $\frac{\sqrt{10} + \sqrt{10} + \sqrt{10}}{\sqrt{5} \times \sqrt{5} \times \sqrt{5}}$  işleminin sonucu kaçtır?

- A)  $\frac{\sqrt{2}}{5}$     B)  $\frac{2\sqrt{2}}{5}$     C)  $\frac{3\sqrt{2}}{5}$   
D)  $\frac{4\sqrt{2}}{5}$     E)  $\sqrt{2}$

3)  $\sqrt{7}$  sayısı  $\sqrt{7} + 1$  sayısının kaç katıdır?

- A)  $\frac{7 - \sqrt{7}}{6}$     B)  $\frac{1 + \sqrt{7}}{6}$     C)  $\frac{7 + \sqrt{7}}{6}$   
D)  $\frac{8 + \sqrt{7}}{6}$     E)  $\frac{9 + \sqrt{7}}{6}$

4)  $\sqrt{0,32}$  ifadesi aşağıdakilerden hangisine eşittir?

- A)  $\frac{\sqrt{2}}{5}$     B)  $\frac{2\sqrt{2}}{5}$     C)  $\frac{3\sqrt{2}}{5}$   
D)  $\frac{4\sqrt{2}}{5}$     E)  $\frac{6\sqrt{2}}{5}$

5) Eni  $5\sqrt{7}$  br, boyu  $4\sqrt{7}$  br olan bir dikdörtgenin alanı kaç  $br^2$ 'dir?

- A)  $8\sqrt{7}$     B)  $10\sqrt{7}$     C)  $18\sqrt{7}$   
D) 120    E) 140

6)  $\frac{\sqrt{0,25} + \sqrt{0,01}}{\sqrt{0,16} - \sqrt{0,04}}$  işleminin sonucu kaçtır?

- A) 1    B)  $\frac{3}{2}$     C) 2    D) 3    E) 4

7) Aşağıdaki ifadelerden hangisi yanlıştır?

- A)  $\sqrt[8]{6} = 6^{1/8}$     B)  $\sqrt{(-5)^4} = 5^2$   
C)  $\sqrt[6]{7^7} = 7^{7/6}$     D)  $\sqrt[2]{10^5} = 10^{2/5}$   
E)  $\sqrt{3^7} = 5^{7/2}$

8)  $a = \sqrt{80}$ ,  $b = 9$ ,  $c = \sqrt{85}$ ,  $d = \sqrt{79}$

sayıları aşağıdaki seçeneklerin hangisinde doğru olarak sıralanmıştır?

- A)  $b < d < a < c$     B)  $b < a < d < c$   
C)  $d < b < a < c$     D)  $d < a < b < c$   
E)  $d < a < c < b$

9) Aşağıdaki işlemlerden hangisi  $12\sqrt{2}$ 'ye eşit değildir?

- A)  $6\sqrt{8}$     B)  $3\sqrt{32}$     C)  $4\sqrt{18}$   
D)  $\sqrt{288}$     E)  $\sqrt{6} \times 2 \times 2\sqrt{4}$

10)  $\sqrt{600} - \sqrt{6}$

Yukarıda verilen işlemin sonucu kaçtır?

- A)  $7\sqrt{6}$     B)  $8\sqrt{6}$     C)  $9\sqrt{6}$   
D)  $10\sqrt{6}$     E)  $11\sqrt{6}$

11)  $3\sqrt{32} + 2\sqrt{2} - (\sqrt{32} + \sqrt{2})$  işleminin sonucu kaçtır?

- A)  $9\sqrt{2}$     B)  $10\sqrt{2}$     C)  $11\sqrt{2}$   
D)  $12\sqrt{2}$     E)  $13\sqrt{2}$

12)  $\underbrace{\sqrt{3} + \sqrt{3} + \dots + \sqrt{3}}_{8TANE \sqrt{3}}$  işleminin sonucu kaçtır?

- A)  $\sqrt{24}$     B)  $3\sqrt{8}$     C)  $8\sqrt{3}$     D)  $8 + \sqrt{3}$   
E) 24

13)  $\sqrt{\sqrt{\sqrt{\sqrt{21 + \sqrt{13 + \sqrt{7 + \sqrt{2 + \sqrt{4}}}}}}}}$  işleminin sonucu kaçtır?

- A)1    B)2    C)3    D)4    E)5

14)  $\sqrt{1,44} + \sqrt{0,49}$  işleminin sonucu kaçtır?

- A) 0,07    B)0,12    C)0,7  
D)1,9    E)2

15) Aşağıda verilen eşitliklerden hangisi doğrudur?

- A)  $\sqrt{81} = 2\sqrt{9}$     B)  $\sqrt{128} = 2\sqrt{64}$   
C)  $\sqrt{99} = 3\sqrt{11}$     D)  $\sqrt{80} = 8\sqrt{10}$   
E)  $\sqrt{65} = 5\sqrt{13}$

16) Aşağıdaki sayılardan hangisi tam karedir?

- A) 136    B)146    C)196    D)216    E)296

17) Kare şeklindeki bir arazinin alanı  $784 \text{ m}^2$  'dir. Buna göre bu arazinin bir kenarının uzunluğu kaç m'dir?

- A)22    B)24    C) 26    D)28    E)32

18)  $2\sqrt{7} + 4\sqrt{7} + 3\sqrt{7} - 3\sqrt{7}$  işleminin sonucu kaçtır?

- A)  $3\sqrt{7}$     B)  $4\sqrt{7}$     C)  $5\sqrt{7}$   
D)  $6\sqrt{7}$     E)  $8\sqrt{7}$

19) Bir bisiklet parkurunda Leyla ile Ersan bisiklet gezisi yapmaktadır. Bu iki kişi geziye başladıktan sonra Leyla  $1385\sqrt{6}$  m,Ersan ise  $1575\sqrt{6}$  m yol alıyor.Buna göre,bu iki kişinin bisiklet gezisinde aldığı toplam yol kaç m'dir?

- A)  $2675\sqrt{3}$     B)  $2965\sqrt{3}$     C)  $4375\sqrt{2}$   
D)  $2965\sqrt{6}$     E)  $3000\sqrt{6}$

20) 7, 9, 18, 28, 36, 52, 126, 169, 256 sayılarından kaç tanesi tam kare sayıdır?

- A)4    B)5    C)6    D)7    E)8

## Ek 8:Yönerge Sayfasının Ekran Görüntüsü

### Kareköklü Sayılar Ve Kareköklü Sayılarla İşlemler Ünitesi Başarı Testi-1

Sevgili öğrenciler,

Bu test sizin Matematik dersindeki Sayılar Ünitesine ait alt öğrenme alanı olan "Kareköklü Sayılar ve Kareköklü Sayılarla Dört İşlem" konusundaki becerinizi ölçmek amacıyla hazırlanmıştır. Testte 5 seçenekli 20 adet çoktan seçmeli soru bulunmaktadır. Testi cevaplamanız için verilen süre 30 dakikadır. Bu süre içinde bütün soruları cevaplayınız.

Testte yer alan soruları görebilmek için şu anda okumakta olduğunuz sayfadaki "**Sınavı Şimdi Uygula**" butonuna farenin sol tuşu ile basınız. Sorulara verdiğiniz cevap program tarafından kaydedilecektir. Her sorunun sadece bir doğru cevabı bulunmaktadır. Soruları cevaplarken, her soru için karşınıza çıkan sayfanın başındaki soruyu dikkatlice okuyunuz, sonra doğru cevap seçeneği için ayrılmış dairenin üzerine gelerek farenin sol tuşuna basınız. Bir sonraki soruya geçmek için "**Sonraki**" tuşuna basınız.

**Bir soruyu cevaplayıp "Sonraki" tuşuna bastıktan sonra tekrar o soruya dönülemez. Eğer "Sonraki" tuşuna basmadan önce işaretlediğiniz seçeneği değiştirmek isterseniz tekrardan doğru cevap seçeneği için ayrılmış dairenin üzerine gelerek farenin sol tuşuna basarak işaretleme yapınız.**

Testin bütün sorularının cevaplamaını bitirdiğinizde "**Tümünü gönder ve bitir**" butonuna basarak sistemden çıkış yapınız.

**Testte yer alan her soru eşit puandadır. Yanlış cevaplarınız doğru cevaplarınızı etkilemeyecektir.**

İzin verilen uygulama: 1

Bu sınav 29 Şubat 2016, Pazartesi, 17:05 tarihinde başladı

Sınav bu tarihte bitecek: 15 Haziran 2016, Çarşamba, 17:05

Zaman sınırı: 30 dk

Sınavı şimdi uygula

## Ek 9: Odak Grubunun Test 2' yi, Referans Grubunun Test 3' ü Alması Durumuna Göre MH Analiz Sonuçları

Detection of Differential Item Functioning using Mantel-Haenszel method without continuity correction and without item purification

Results based on asymptotic inference

No set of anchor items was provided

Mantel-Haenszel Chi-square statistic:

	Stat.	P-value
Cevap..1	1.6836	0.1944
Cevap..2	0.0648	0.7991
Cevap..3	0.0072	0.9325
Cevap..4	0.0061	0.9376
Cevap..5	0.0554	0.8140
Cevap..6	0.3718	0.5420
Cevap..7	6.5341	0.0106 *
Cevap..8	1.6782	0.1952
Cevap..9	0.1367	0.7115
Cevap..10	0.2306	0.6311
Cevap..11	0.4433	0.5055
Cevap..12	0.6063	0.4362
Cevap..13	0.3501	0.5541
Cevap..14	0.6634	0.4154
Cevap..15	1.8249	0.1767
Cevap..16	0.0493	0.8242
Cevap..17	1.9241	0.1654
Cevap..18	1.5771	0.2092
Cevap..19	4.6908	0.0303 *
Cevap..20	0.1819	0.6697

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Detection threshold: 3.8415 (significance level: 0.05)

Items detected as DIF items:

Cevap..7  
Cevap..19

Effect size (ETS Delta scale):

Effect size code:

'A': negligible effect  
'B': moderate effect  
'C': large effect

	alphaMH	deltaMH	
Cevap..1	0.7034	0.8267	A
Cevap..2	0.9312	0.1674	A
Cevap..3	1.0253	-0.0586	A
Cevap..4	1.0210	-0.0487	A
Cevap..5	1.0705	-0.1600	A



Cevap..6	0.8402	0.4091	A
Cevap..7	2.0501	-1.6871	C
Cevap..8	1.4643	-0.8963	A
Cevap..9	1.1046	-0.2338	A
Cevap..10	0.8848	0.2877	A
Cevap..11	0.8023	0.5177	A
Cevap..12	0.8000	0.5245	A
Cevap..13	0.8296	0.4389	A
Cevap..14	0.7962	0.5356	A
Cevap..15	0.6483	1.0184	B
Cevap..16	0.9315	0.1667	A
Cevap..17	0.6404	1.0472	B
Cevap..18	1.6512	-1.1786	B
Cevap..19	1.8604	-1.4589	B
Cevap..20	1.1811	-0.3912	A

Effect size codes: 0 'A' 1.0 'B' 1.5 'C'  
(for absolute values of 'deltaMH')



## Ek 10: Odak Grubunun Test 3' ü, Referans Grubunun Test 2'yi Alması Durumuna Göre MH

### Analiz Sonuçları

Detection of Differential Item Functioning using Mantel-Haenszel method without continuity correction and without item purification

Results based on asymptotic inference

No set of anchor items was provided

Mantel-Haenszel Chi-square statistic:

	Stat.	P-value	
Cevap..1	0.4763	0.4901	
Cevap..2	0.1265	0.7221	
Cevap..3	1.7910	0.1808	
Cevap..4	0.4526	0.5011	
Cevap..5	5.2957	0.0214	*
Cevap..6	0.6696	0.4132	
Cevap..7	7.9903	0.0047	**
Cevap..8	1.3655	0.2426	
Cevap..9	1.0466	0.3063	
Cevap..10	1.4356	0.2308	
Cevap..11	0.0143	0.9048	
Cevap..12	0.9437	0.3313	
Cevap..13	13.2429	0.0003	***
Cevap..14	0.0963	0.7564	
Cevap..15	0.0125	0.9109	
Cevap..16	6.2051	0.0127	*
Cevap..17	0.0062	0.9373	
Cevap..18	0.5678	0.4511	
Cevap..19	7.0855	0.0078	**
Cevap..20	2.0238	0.1548	

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Detection threshold: 3.8415 (significance level: 0.05)

Items detected as DIF items:

Cevap..5  
Cevap..7  
Cevap..13  
Cevap..16  
Cevap..19

Effect size (ETS Delta scale):

Effect size code:

'A': negligible effect  
'B': moderate effect  
'C': large effect

	alphaMH	deltaMH	
Cevap..1	1.2154	-0.4585	A
Cevap..2	0.9070	0.2294	A
Cevap..3	0.6787	0.9110	A
Cevap..4	1.2027	-0.4337	A
Cevap..5	2.0171	-1.6490	C

Cevap..6	1.2501	-0.5246	A
Cevap..7	0.4359	1.9512	C
Cevap..8	1.3927	-0.7785	A
Cevap..9	0.7567	0.6551	A
Cevap..10	1.3697	-0.7393	A
Cevap..11	1.0349	-0.0806	A
Cevap..12	1.3178	-0.6485	A
Cevap..13	2.8806	-2.4863	C
Cevap..14	1.0909	-0.2045	A
Cevap..15	1.0332	-0.0768	A
Cevap..16	0.4663	1.7931	C
Cevap..17	1.0232	-0.0540	A
Cevap..18	0.7497	0.6769	A
Cevap..19	0.4385	1.9373	C
Cevap..20	0.5911	1.2354	B

Effect size codes: 0 'A' 1.0 'B' 1.5 'C'  
(for absolute values of 'deltaMH')



## Ek 11: Odak Grubunun Test 2' yi, Referans Grubunun Test 3' ü Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları

Detection of uniform Differential Item Functioning  
using Logistic regression method, without item purification  
and with LRT DIF statistic

Matching variable: test score

No set of anchor items was provided

Logistic regression DIF statistic:

	Stat.	P-value	
Cevap..1	1.8671	0.1718	
Cevap..2	0.0046	0.9462	
Cevap..3	0.1425	0.7058	
Cevap..4	0.0125	0.9112	
Cevap..5	0.0265	0.8706	
Cevap..6	0.6506	0.4199	
Cevap..7	7.4592	0.0063	**
Cevap..8	1.3626	0.2431	
Cevap..9	0.0023	0.9621	
Cevap..10	0.2429	0.6221	
Cevap..11	0.6050	0.4367	
Cevap..12	0.1646	0.6850	
Cevap..13	0.4898	0.4840	
Cevap..14	0.3064	0.5799	
Cevap..15	1.9838	0.1590	
Cevap..16	0.0012	0.9728	
Cevap..17	1.1475	0.2841	
Cevap..18	2.8021	0.0941	.
Cevap..19	3.9403	0.0471	*
Cevap..20	0.1515	0.6971	

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Detection threshold: 3.8415 (significance level: 0.05)

Items detected as uniform DIF items:

Cevap..7  
Cevap..19

Effect size (Nagelkerke's R<sup>2</sup>):

Effect size code:

'A': negligible effect  
'B': moderate effect  
'C': large effect

	R <sup>2</sup>	ZT	JG
Cevap..1	0.0076	A	A
Cevap..2	0.0000	A	A
Cevap..3	0.0005	A	A
Cevap..4	0.0000	A	A
Cevap..5	0.0001	A	A
Cevap..6	0.0021	A	A
Cevap..7	0.0262	A	A
Cevap..8	0.0044	A	A

Cevap..9 0.0000 A A  
Cevap..10 0.0011 A A  
Cevap..11 0.0024 A A  
Cevap..12 0.0007 A A  
Cevap..13 0.0017 A A  
Cevap..14 0.0014 A A  
Cevap..15 0.0075 A A  
Cevap..16 0.0000 A A  
Cevap..17 0.0054 A A  
Cevap..18 0.0164 A A  
Cevap..19 0.0181 A A  
Cevap..20 0.0011 A A

Effect size codes:

Zumbo & Thomas (ZT): 0 'A' 0.13 'B' 0.26 'C' 1  
Jodoin & Gierl (JG): 0 'A' 0.035 'B' 0.07 'C' 1



## Ek 12: Odak Grubunun Test 3'ü, Referans Grubunun Test 2' yi Alması Durumuna Göre Tek Biçimli DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları

Detection of uniform Differential Item Functioning using Logistic regression method, without item purification and with LRT DIF statistic

Matching variable: test score

No set of anchor items was provided

Logistic regression DIF statistic:

	Stat.	P-value	
Cevap..1	1.1469	0.2842	
Cevap..2	0.0313	0.8597	
Cevap..3	1.4358	0.2308	
Cevap..4	0.4039	0.5251	
Cevap..5	5.6473	0.0175	*
Cevap..6	0.5689	0.4507	
Cevap..7	7.3505	0.0067	**
Cevap..8	1.7917	0.1807	
Cevap..9	0.7034	0.4016	
Cevap..10	2.3237	0.1274	
Cevap..11	0.0351	0.8514	
Cevap..12	0.6506	0.4199	
Cevap..13	12.4331	0.0004	***
Cevap..14	0.0190	0.8904	
Cevap..15	0.1317	0.7167	
Cevap..16	8.3330	0.0039	**
Cevap..17	0.2932	0.5882	
Cevap..18	0.4227	0.5156	
Cevap..19	6.9808	0.0082	**
Cevap..20	0.8308	0.3620	

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Detection threshold: 3.8415 (significance level: 0.05)

Items detected as uniform DIF items:

Cevap..5  
Cevap..7  
Cevap..13  
Cevap..16  
Cevap..19

Effect size (Nagelkerke's R<sup>2</sup>):

Effect size code:

'A': negligible effect  
'B': moderate effect  
'C': large effect

	R <sup>2</sup>	ZT	JG
Cevap..1	0.0043	A	A
Cevap..2	0.0001	A	A
Cevap..3	0.0052	A	A
Cevap..4	0.0014	A	A
Cevap..5	0.0163	A	A

Cevap..6	0.0019	A	A
Cevap..7	0.0212	A	A
Cevap..8	0.0061	A	A
Cevap..9	0.0024	A	A
Cevap..10	0.0087	A	A
Cevap..11	0.0001	A	A
Cevap..12	0.0028	A	A
Cevap..13	0.0481	A	B
Cevap..14	0.0001	A	A
Cevap..15	0.0005	A	A
Cevap..16	0.0376	A	B
Cevap..17	0.0013	A	A
Cevap..18	0.0022	A	A
Cevap..19	0.0290	A	A
Cevap..20	0.0043	A	A

Effect size codes:

Zumbo & Thomas (ZT): 0 'A' 0.13 'B' 0.26 'C' 1

Jodoin & Gierl (JG): 0 'A' 0.035 'B' 0.07 'C' 1



### Ek 13: Odak Grubunun Test 2'yi, Referans Grubunun Test 3' ü Alması Durumuna Göre Tek Bıçimli Olmayan DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları

Detection of nonuniform Differential Item Functioning  
using Logistic regression method, without item purification  
and with LRT DIF statistic

Matching variable: test score

No set of anchor items was provided

Logistic regression DIF statistic:

	Stat.	P-value
Cevap..1	0.3793	0.5380
Cevap..2	0.9301	0.3348
Cevap..3	0.1047	0.7463
Cevap..4	1.2872	0.2566
Cevap..5	0.0950	0.7579
Cevap..6	0.2731	0.6012
Cevap..7	0.8373	0.3602
Cevap..8	0.0987	0.7534
Cevap..9	1.4139	0.2344
Cevap..10	0.2347	0.6280
Cevap..11	1.2426	0.2650
Cevap..12	0.2574	0.6119
Cevap..13	1.5713	0.2100
Cevap..14	2.9674	0.0850
Cevap..15	3.9923	0.0457
Cevap..16	4.3166	0.0377
Cevap..17	0.8756	0.3494
Cevap..18	0.8359	0.3606
Cevap..19	0.6384	0.4243
Cevap..20	1.0268	0.3109

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Detection threshold: 3.8415 (significance level: 0.05)

Items detected as nonuniform DIF items:

Cevap..15  
Cevap..16

Effect size (Nagelkerke's R<sup>2</sup>):

Effect size code:

'A': negligible effect  
'B': moderate effect  
'C': large effect

	R <sup>2</sup>	ZT	JG
Cevap..1	0.0015	A	A
Cevap..2	0.0036	A	A
Cevap..3	0.0004	A	A
Cevap..4	0.0045	A	A
Cevap..5	0.0003	A	A
Cevap..6	0.0009	A	A
Cevap..7	0.0029	A	A
Cevap..8	0.0003	A	A
Cevap..9	0.0051	A	A



Cevap..10 0.0010 A A  
Cevap..11 0.0050 A A  
Cevap..12 0.0011 A A  
Cevap..13 0.0056 A A  
Cevap..14 0.0135 A A  
Cevap..15 0.0149 A A  
Cevap..16 0.0211 A A  
Cevap..17 0.0041 A A  
Cevap..18 0.0049 A A  
Cevap..19 0.0029 A A  
Cevap..20 0.0071 A A

Effect size codes:

Zumbo & Thomas (ZT): 0 'A' 0.13 'B' 0.26 'C' 1  
Jodoin & Gierl (JG): 0 'A' 0.035 'B' 0.07 'C' 1



## Ek 14: Odak Grubunun Test 3'ü, Referans Grubunun Test 2' yi Alması Durumuna Göre Tek Bıçimli Olmayan DMF'yi Belirlemek İçin Yapılan Analiz Sonuçları

Detection of nonuniform Differential Item Functioning  
using Logistic regression method, without item purification  
and with LRT DIF statistic

Matching variable: test score

No set of anchor items was provided

Logistic regression DIF statistic:

	Stat.	P-value	
Cevap..1	0.0504	0.8223	
Cevap..2	0.2410	0.6235	
Cevap..3	2.4919	0.1144	
Cevap..4	0.1214	0.7275	
Cevap..5	0.0750	0.7842	
Cevap..6	8.0753	0.0045	**
Cevap..7	1.4429	0.2297	
Cevap..8	0.8552	0.3551	
Cevap..9	0.2647	0.6069	
Cevap..10	3.5740	0.0587	.
Cevap..11	6.0795	0.0137	*
Cevap..12	1.7662	0.1839	
Cevap..13	1.4313	0.2316	
Cevap..14	0.8620	0.3532	
Cevap..15	1.9959	0.1577	
Cevap..16	0.0040	0.9498	
Cevap..17	4.9415	0.0262	*
Cevap..18	0.6751	0.4113	
Cevap..19	0.7676	0.3810	
Cevap..20	0.2986	0.5848	

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Detection threshold: 3.8415 (significance level: 0.05)

Items detected as nonuniform DIF items:

Cevap..6  
Cevap..11  
Cevap..17

Effect size (Nagelkerke's R<sup>2</sup>):

Effect size code:

'A': negligible effect  
'B': moderate effect  
'C': large effect

	R <sup>2</sup>	ZT	JG
Cevap..1	0.0002	A	A
Cevap..2	0.0009	A	A
Cevap..3	0.0089	A	A
Cevap..4	0.0004	A	A
Cevap..5	0.0002	A	A
Cevap..6	0.0259	A	A
Cevap..7	0.0041	A	A

Cevap..8 0.0029 A A  
Cevap..9 0.0009 A A  
Cevap..10 0.0133 A A  
Cevap..11 0.0217 A A  
Cevap..12 0.0076 A A  
Cevap..13 0.0054 A A  
Cevap..14 0.0035 A A  
Cevap..15 0.0071 A A  
Cevap..16 0.0000 A A  
Cevap..17 0.0209 A A  
Cevap..18 0.0036 A A  
Cevap..19 0.0031 A A  
Cevap..20 0.0016 A A

Effect size codes:

Zumbo & Thomas (ZT): 0 'A' 0.13 'B' 0.26 'C' 1  
Jodoin & Gierl (JG): 0 'A' 0.035 'B' 0.07 'C' 1



## ÖZGEÇMİŞ

**Adı, Soyadı:** Ebru BALTA

**Doğum Yeri, Tarihi :** İskenderun/ 02.08.1986

**e-mail:** [ebrubalta2@gmail.com](mailto:ebrubalta2@gmail.com)

### Eğitim Bilgileri

**Lisans Öğrenimi:** Hacettepe Üniversitesi Eğitim Fakültesi Fizik Eğitimi ABD

**Yüksek Lisans Öğrenimi:** Hacettepe Üniversitesi Fizik Öğretmenliği (Tezsiz)-Mersin Üniversitesi  
Ölçme ve Değerlendirme (Tezli)

### İş Deneyimi

**Çalıştığı Kurumlar:** Selçuk Üniversitesi Eğitim Fakültesi Fizik Eğitimi ABD (2011-2013)

Mersin Üniversitesi Eğitim Fakültesi Ölçme ve Değerlendirme ABD (2013-...)