



Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Eğitim Bilimleri Anabilim Dalı

Eğitimde Ölçme ve Değerlendirme Bilim Dalı

**PISA 2009 FEN OKURYAZARLIĞI ALT TESTİNİN DEĞİŞEN
MADDE FONKSİYONU AÇISINDAN İNCELENMESİ**

Cansu AYAN

Yüksek Lisans Tezi

Ankara, 2011

PISA 2009 FEN OKURYAZARLIĞI ALT TESTİNİN DEĞİŞEN MADDE FONKSİYONU
AÇISINDAN İNCELENMESİ

Cansu AYAN

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Eğitim Bilimleri Anabilim Dalı

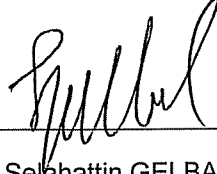
Eğitimde Ölçme ve Değerlendirme Bilim Dalı

Yüksek Lisans Tezi

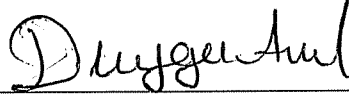
Ankara, 2011

KABUL VE ONAY

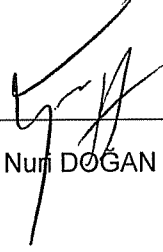
Cansu AYAN tarafından hazırlanan "PISA 2009 Fen Okuryazarlığı Alt Testinin Değişen Madde Fonksiyonu Açısından İncelenmesi" başlıklı bu çalışma, 27.12.2011 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.



Prof. Dr. Selahattin GELBAL (Başkan)



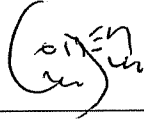
Doç. Dr. Duygu ANIL (Danışman)



Doç. Dr. Nuri DOĞAN



Doç. Dr. Hülya KELECIOĞLU



Yrd. Doç. Dr. Ömer KUTLU

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. İrfan ÇAKIN

Enstitü Müdürü

BİLDİRİM

Hazırladığım tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kağıt ve elektronik kopyalarının Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

- ☒ Tezimin/Raporumun tamamı her yerden erişime açılabilir.
- ☐ Tezim/Raporum sadece Hacettepe Üniversitesi yerleşkelerinden erişime açılabilir.
- ☐ Tezimin/Raporumunyıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

27.12.2011

Cansu AYAN



TEŞEKKÜR

Tez çalışmam sürecinde öncelikle, desteğini ve güvenini hep yanımda hissettiğim, yoğunluğuna rağmen değerli vaktini bana ayıran ve her buluşmamızda moral veren sevgili danışmanım Doç. Dr. Duygu ANIL'a, ardından bu süreçte her sorum ve sıkıntım için bıkmadan uğraşan içlerinde hep aile ortamındaki samimiyeti ve sevgiyi hissettiğim başta Doç. Dr. Nuri DOĞAN olmak üzere kıymetli hocalarım Prof. Dr. Selahattin GELBAL'a ve Doç. Dr. Hülya KELECİOĞLU 'na çok teşekkür ederim.

Çalışmamda bana vakit ayırıp yardımcı olan sevgili Dilara BAKAN KALAYCIOĞLU'na, başım ne zaman sıkışsa hemen hiç şikayet etmeden tüm samimiyetiyle yardımına koşan sevgili arkadaşlarım Neşe ÖZTÜRK ve Şeyma YÜKSEL 'e, karşısına oturduğum günden beri her türlü sıkıntıda yardımına koşan hep destek olan, yol gösteren, sevgili oda arkadaşım Tuğçe ÖZMEN'e, ilk günden beri her türlü sıkıntımı şikayetimi heyecanımı paylaşan ve benden hiç bıkmayan canım arkadaşlarım Aylin BALLIDAĞ, Zeynep ULUDAĞ, Ayşe Nur DEMİR, Özgür CENİK, Arda GÜLER, Ulaş AKBAŞ'a çok teşekkür ederim.

Hayatım boyunca yaptığım her işte bana tereddütsüz güvendiklerini tüm kalbimle hissettiğim, ve verdikleri bu güç sayesinde hayata karşı çok daha dik durabilmemi sağlayan, varlıklarının ve sevgilerinin bedelini ödeyemeyeceğim annem Nuray AYAN, babam Erol AYAN ve sevgili kardeşim Cansel AYAN'a

Ve maddi desteklerinden ötürü TÜBİTAK 'a

Ne kadar teşekkür etsem azdır.

ÖZET

Ayan, Cansu. *PISA 2009 Fen Okuryazarlığı Alt Testinin Değişen Madde Fonksiyonu Açısından İncelenmesi*, Yüksek Lisans Tezi, Ankara, 2011.

Bu çalışma kapsamında, Uluslararası Öğrenci Değerlendirme Programı (PISA) fen okuryazarlığı alt testinde yer alan maddelerle, cinsiyet değişkenine göre, değişen madde fonksiyonu (DMF) analizi yürütülmüştür. DMF analizinde lojistik regresyon (LR) ve Mantel Haenszel (MH) yöntemleri kullanılmış ve ardından bu iki teknikten elde edilen sonuçlar karşılaştırılmıştır.

Çalışmanın veri setini, PISA 2009 uygulaması Türkiye örnekleminde 12. Kitapçığın uygulandığı 338 öğrencinin fen okuryazarlığı alt testindeki maddelere verdikleri cevaplar oluşturmaktadır. Ayrıca Fen okuryazarlığı alt testine ait betimsel istatistikler ve güvenilirlik analizleri hesaplanarak DMF analizinden önce grup dağılımı hakkında genel bilgi edinilmiştir.

Cinsiyet değişkenine göre yürütülen 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.

Anahtar Sözcükler

PISA, değişen madde fonksiyonu, lojistik regresyon, Mantel Haenszel

ABSTRACT

Ayan, Cansu. *PISA 2009 Analysis of Science Test in Terms of Changing Item Function*
Masters Thesis, Ankara 2011.

In this study, Differential items Function (DIF) analysis was conducted with the help of the items in The Programme for International Student Assessment (PISA) science test according to gender variable. Logistic regression (LR) and Mantel Haenszel (MH) methods were used for DIF analysis which is followed by the comparison of the results.

The data of the work includes the answers of 338 students who was applied 12. booklet from Turkey which is one of 65 countries attending PISA 2009. Also, some research on the distribution of group were done by calculating descriptive statistics and reliability analysis before making DIF analysis.

The analysis of the data collected in Turkey in terms of gender demonstrated nonuniform DIF in 4 items when the LR method was used, and in 4 items when MH method was employed. The three items of these favored female students and one item favored males.

When the results of MH and LR techniques are compared, it is seen that MH and LR procedures were similar with respect to the size of their respective Chi-square values and were different in terms of the number of items with DIF values they generated.

In addition to this, when total item numbers that were decided as DIF were compared, it was identified that MH technique is more sensitive in uniform DIF identification, because none of uniform DIF which were decided after MH usage could be identified with LR procedure.

Key Words

PISA, differential item functioning, logistic regression, Mantel-Haenszel

İÇİNDEKİLER

KABUL VE ONAY.....	i
BİLDİRİM.....	ii
TEŞEKKÜR	iii
ÖZET.....	iv
ABSTRACT.....	v
İÇİNDEKİLER.....	vi
TABLolar	viii
ŞEKİLLER	ix
I. BÖLÜM	1
GİRİŞ	1
1.1 Problem Durumu	1
1.1.1 Yanlılık.....	5
1.1.2 Değişen Madde Fonksiyonu.....	7
1.1.3 Tarihsel Süreçte Yanlılık Belirleme Teknikleri.....	10
1.1.4 Olasılık Tablosuna Dayalı Yöntemler.....	14
Mantel Haenszel.....	16
Lojistik Regresyon.....	18
1.2 Araştırmanın Amacı ve Önemi.....	21
1.3 Problem Cümlesi.....	22
1.3.1 Alt Problemler.....	22
1.4 Sayıtlar.....	22
1.5 Sınırlılıklar.....	22
1.6 İlgili Araştırmalar.....	22
2. BÖLÜM.....	30
YÖNTEM.....	30
2.1 Araştırmanın Türü.....	30
2.2 Çalışma Grubu.....	30
2.3 PISA’da Kullanılan Veri Toplama Araçları.....	31

2.4 Verilerin Çözümlemesi.....	33
3. BÖLÜM.....	38
BULGULAR VE YORUM.....	38
3.1 Birinci Alt Probleme İlişkin Bulgular ve Yorumlar.....	38
3.2 İkinci Alt Probleme İlişkin Bulgular ve Yorumlar.....	40
3.3 Üçüncü Alt Probleme İlişkin Bulgular ve Yorumlar.....	42
4. BÖLÜM.....	45
SONUÇ VE ÖNERİLER.....	45
4.1 Sonuçlar.....	45
LR ve MH Yöntemleri ile DMF Analiz Sonuçları.....	45
İki Yöntemin Karşılaştırılmasına Yönelik Analiz Sonuçları.....	45
4.2 Öneriler.....	46
KAYNAKÇA.....	47
ÖZGEÇMİŞ.....	52
EKLER.....	53
Ek-1: LR etki büyüklüğünü hesaplamada kullanılan syntax.....	53
Ek-2: MH analiz sonuçları.....	58
Ek-3: LR analiz sonuçları.....	57
Ek-4: LR yöntemine ait R^2 değerleri.....	62

TABLOLAR

Tablo 1: Fen maddelerinin ölçtüğü alt boyutlar.....	3
Tablo 2: J yetenek seviyesindekiler için bir maddenin olasılık tablosu.....	15
Tablo 3: J yetenek seviyesindekiler için bir maddenin olasılık tablosu.....	15
Tablo 4: Herhangi bir m seviyesi için hazırlanmış ki kare tablosu.....	16
Tablo 5: MH_{D-DIF} değerlerini yorumlama ölçütleri.....	18
Tablo 6: Zumbo ve Thomas 'ın ΔR^2 değerlerini yorumlama ölçütleri.....	20
Tablo 7: Jodoin ve Gierl 'in ΔR^2 değerlerini yorumlama ölçütleri.....	20
Tablo 8: Çalışma Grubuna ait frekans ve yüzde değerleri.....	31
Tablo 9: Kitapçıklara ait kişi ve fen okuryazarlığı madde sayısı.....	33
Tablo 10: Çalışma Grubuna ait betimsel istatistikler.....	35
Tablo 11: Cinsiyet değişkenine göre betimsel istatistikler.....	36
Tablo 12: Fen okuryazarlığı alt testine ait cinsiyet değişkenine göre Mantel-Haenszel analiz sonuçları.....	38
Tablo 13: Cinsiyet değişkenine göre Mantel-Haenszel analiz sonuçları.....	39
Tablo 14: Fen okuryazarlığı alt testine ait cinsiyet değişkenine göre Lojistik Regresyon analiz Sonuçları.....	40
Tablo 15: Cinsiyet değişkenine göre etki büyüklüğü değerleri.....	42
Tablo 16: LR ve MH yöntemleriyle elde edilen DMF sonuçları.....	43

ŞEKİLLER

Şekil 1: Tek Biçimli ve Tek Biçimli Olmayan DMF.....	9
Şekil 2: Kızların Dağılım Grafiği.....	36
Şekil 3: Erkeklerin Dağılım Grafiği.....	36

I. BÖLÜM

GİRİŞ

Bu bölümde problem durumu, problem cümlesi, alt problemler, sınırlılıklar, sayıltılar, araştırmanın amacı ve önemi ile ilgili araştırmalar yer almaktadır.

1.1 PROBLEM DURUMU

Ülkemizde öğrenci başarısını belirlemek, belirlenen başarıya dayalı kararlar vermek ya da mevcut eğitim sistemindeki durumu değerlendirmek adına ulusal ve uluslararası düzeyde birçok sınav yapılmaktadır. T.C. Milli Eğitim Bakanlığı (MEB) ulusal boyutta yaptığı öğrenci başarısını belirleme çalışmalarını uluslararası boyutta da sürdürmekte, kendi öğrencilerinin başarı düzeylerini ve eğitim sistemini, diğer ülkelerin verileri ile karşılaştırarak güçlü ve iyileşmeye açık yönlerini belirlemek için uluslararası çalışmalara katılmaktadır.

Söz konusu uluslararası çalışmalardan biri de Uluslararası Öğrenci Değerlendirme Programı (The Programme for International Student Assessment-PISA)'dır. Ekonomik İşbirliği ve Kalkınma Teşkilâtı-OECD (Organisation for Economic Co-Operation and Development) tarafından düzenlenen PISA, öğrencilerin, matematik, fen ve okuma becerileri alanlarındaki bilgi ve becerilerinin değerlendirildiği uluslararası en büyük eğitim araştırmalarından biridir. Üç yılda bir yapılan bu araştırmayla, OECD üyesi ülkeler ve diğer katılımcı ülkelerdeki 15 yaş grubu öğrencilerin modern toplumda yerlerini alabilmeleri için gereken temel bilgi ve becerilere ne ölçüde sahip oldukları değerlendirilmektedir (MEB, 2010).

PISA projesinde zorunlu eğitimin sonuna gelen 15 yaş grubu öğrencilerin yalnızca öğrendiklerinin ne kadarını hatırlayabildiklerinin değil, aynı zamanda öğrendiklerini okulda ve okul dışı yaşamlarında kullanabilme yeterliklerinin; karşılaştıkları yeni durumları anlamak, sorunları çözmek, bilmedikleri konularda tahminde bulunmak ve muhakeme yapabilmek için bilgi ve becerilerinden ne ölçüde yararlanabildiklerinin

belirlenmesi hedeflenmektedir. Bu amaç, PISA'yı diğer değerlendirme yaklaşımlarından ayırmaktadır.

PISA'nın değerlendirme çerçevesi ve kavramsal temelleri, projeye katılan ülkelerdeki uzmanlar tarafından belirlenmekte, yapılan görüşmeler sonrasında katılımcı ülkelerin fikir birliğiyle onaylanmaktadır. Bu çerçevede, “okuryazarlık” kavramına ilişkin yeni bir anlayış ortaya çıkmıştır:

“Okuryazarlık kavramı; öğrencilerin bilgilerini günlük yaşamda kullanmak, mantıksal çıkarımlar yapmak, çeşitli durumlarla ilgili problemleri yorumlamak ve çözmek için öğrendiklerinden çıkarımlar yapma kapasitesi olarak tanımlanmaktadır” (MEB, 2010).

PISA projesinde her dönem okuma becerileri, matematik ve fen okuryazarlığı alanlarından birine temel alan olarak ağırlık verilmektedir. Bu bağlamda PISA uygulaması çerçevesinde kazanılan yeni kavramlardan biri de fen okuryazarlığı kavramıdır:

“Fen okuryazarlığı kavramı; bir bireyin sahip olduğu fen bilgisi ve bu bilginin soruları tanımlamak, yeni bilgi edinmek, bilimsel olguları açıklamak, fen ile ilgili konularda kanıtlara dayalı sonuçlar çıkarmak için kullanımı; bilgi edinme ve araştırma amacıyla fenin karakteristik özelliklerini anlayışı, fen ve teknolojinin maddî, düşünsel ve kültürel çevremizi nasıl şekillendirdiğinin farkına varması ve duyarlı bir vatandaş olarak bilimle ilgili konulara ve bilimsel fikirlere ilgi göstermesi şeklinde tanımlanmaktadır” (MEB, 2010).

PISA 2009 uygulamasında fen okuryazarlığı alt testine 90 dakikalık bir uygulama süresi ayrılmıştır. Bu süreçte öğrencilerin fen okuryazarlığındaki yeterliklerinin hem bilişsel hem de duyuşsal yönleri değerlendirilmeye çalışılmaktadır.

Tablo 1’de fen okuryazarlığı değerlendirme çerçevesine ilişkin hazırlanan maddelerin ölçtüğü alt boyutların temel tanımları sunulmakta ve bu alt boyutlara örnekler verilmektedir(OECD, 2009).

Tablo 1: Fen Maddelerinin Ölçtüğü Alt Boyutlar

Tanımlama ve Ayırtedici Özellikler	Bireyin aşağıdaki becerilere ne derece sahip olduğu: • Bilimsel bilgiye sahiptir ve bu bilgiyi fen ile ilgili konularda soruları tanımlama, yeni bilgiler edinme, bilimsel olguları açıklama ve kanıta dayalı sonuçlar çıkarma amacıyla kullanır, • Fenin karakteristik özelliklerini, insana ait bilgi birikimi ve sorgulamanın bir biçimi olarak algılar, • Fen ve teknolojinin maddi, düşünsel ve kültürel çevremizi nasıl şekillendirdiğinin farkına vardığını gösterir, • Duyarlı bir vatandaş olarak bilimle ilgili konularla ve bilimsel fikirlerle ilgilenir.
Bilgi Alanı	Fen alanı bilgisi, örneğin; • “Fiziksel Sistemler” • “Canlılar ile İlgili Sistemler” • “Yerküre ve Uzay Sistemleri” • “Teknoloji Sistemleri” <i>Bilimsel yöntem bilgisi, örneğin:</i> • “bilimsel sorgulama” • “bilimsel açıklamalar
İlgili Beceriler ve Düşünme Süreçleri	Fen ile ilgili görev ya da süreç türleri; • Bilimsel durumları ayırt etme • Olguları bilimsel olarak açıklama • Bilimsel kanıtları kullanma
Bağlam ve Konular	Fenin; • “Sağlık” • “Doğal Kaynaklar” • “Çevre” • “Afetler” • “Fen ve teknolojinin sınırları” gibi konularda kişisel, sosyal ve küresel ortamlarla ilişkili kullanımları üzerinde yoğunlaşan uygulama alanı.

Tablo 1’de görüldüğü gibi PISA 2009 uygulamasında fen okuryazarlığı alt testine ait maddeler genel olarak öğrencilerin alanda sahip olması gereken bilgiler, bu alana ilişkin düşünme süreçleri, öğrencilerin bilimsel problemlerde karşılaştığı bağlamlar, öğrencilerin öğrenmeye yönelik tutum ve eğilimleri gibi boyutları ölçmeyi hedeflemektedir. Bu amaçla geliştirilen ölçme araçları değerlendirme üniteleri biçiminde hazırlanmaktadır ve öğrencilerin günlük hayatta karşılaşılabilecekleri türden görevler içermektedir. Bu değerlendirme ünitelerinde metinler, şekiller, tablolar ve/veya grafiklerden oluşan ortak bir madde kökü ile ardından gelen metnin, şeklin, tablonun ya da grafiğin farklı yönleriyle ilgili maddeler bulunmaktadır.

Başlangıçta OECD’ye üye ülkelerin katıldığı PISA, günümüzde dünya genelinde yapılan en önemli araştırmalardan biri haline gelmiştir. PISA 2009’a, 65 katılımcı ülkeden, 15 yaş grubu yaklaşık 26 milyon öğrenciyi temsilen, 475.460 öğrenci katılmıştır. Bu değerlendirmeye daha sonra ek olarak OECD üyesi olmayan 9 ülkeden, 15 yaş grubu yaklaşık 2 milyon öğrenciyi temsilen 50.000 öğrenci daha katılmıştır (MEB,2010).

Dünya çapında uygulanan ve sonuçlarıyla ülkelerin eğitim sistemlerini etkileyen bu denli önemli bir sınavdan çıkan sonuçların anlamlı olabilmesi için ölçme sonuçları minimum miktarda hata içermelidir. Eğitim ve psikoloji gibi sosyal alanlarda yapılan ölçmelerde ise bu durum üzerinde daha hassasiyetle çalışılması gereken bir nokta olmaktadır. Çünkü canlıların davranışsal özelliklerini ve bunlardan da özellikle bilişsel ve duyuşsal yönü baskın olanları ölçme söz konusu olunca, test puanlarının testle ölçüldüğü varsayılan yapıdan farklı bir değişkenlik kaynağından etkilenmesi durumu kaçınılmazdır. Elle tutulur, gözle görülür ve kararlı özellikler üzerindeki ölçümlerin bile tümüyle hatadan arınmış olması olanaksızdır. Eğer bu doğru olmasaydı, puanlar mükemmel bir biçimde güvenilir ve geçerli olurdu ki; özellikle sosyal bilimlerde bu durumun sağlanması olası değildir. Ancak bütün bunlara karşın, uygulamada yeterli sayılabilecek doğrulukta ölçüler verecek kadar yüksek geçerlikte ölçekler geliştirilebilir (Crocker ve Algina,1986; Acar,2008; Özçelik,2011). Geçerliğe karşı önemli tehditlerden birisi madde ve test yanlılığıdır (Clauser ve Mazor, 1998).

Test maddelerinde, yanlılık ölçmedeki sistematik hatanın varlığı olarak düşünülür (Osterlind,1983). Yanlılıkta önemli olan nokta söz konusu hata kaynaklarının alt

popülasyonlardan birine (erkekler ya da diğer alt grup olan kadınlar gibi) haksız bir yarar sağlaması durumudur (Crocker ve Algina,1986). Test maddelerinin, testi alan aynı yetenek düzeyinde olup, farklı gruplara mensup bireyler için daha zor ya da daha kolay cevaplanabilir olmaması gerekir. Bu durum, farklı alt gruplardan gelen aynı yetenek düzeyindeki bireylerin test puanlarında farklılıklar oluşturur ve test puanlarındaki bu farklılıklar bireylerin belli bir yetenek ölçüsünde sıralamalarında adil olmayan, yanlış kararlar alınmasına neden olur (Zumbo, 1999).

Son yıllarda, eğitimdeki ilerlemenin değerlendirilmesinde standart başarı testlerinin kullanımının yaygınlaşmış ve bu testlere duyulan güvenle birlikte bu testlerde yapılan yanlılık çalışmaları da büyük ilgi çekmiştir (Finch ve French,2007). Dünya çapında sonuçlar doğuran ve ülkelerin eğitim sistemlerini etkileyen bu denli büyük ölçekli bir sınavda da maddelerin değişen madde fonksiyonu içerip içermeme durumunun belirlenmesi önemli görülmektedir. Konuyla ilgili alanyazın incelendiğinde özellikle fen ve matematik gibi sayısal alanlardaki performanslarda gözlenen cinsiyet farklılıklarına ve sonucunda tespit edilen Değişen Madde Fonksiyonu (DMF) gösteren maddelere dikkat çekildiği görülmektedir. Bu nedenlerle bu çalışma kapsamında, PISA 2009 uygulaması Türkiye örnekleminde Fen Okuryazarlığı alt testinde yer alan maddelerin cinsiyete göre değişen madde fonksiyonu gösterip göstermediği incelenmiştir.

1.1.1 Yanlılık

Test yanlılığı belli bir grubun üyeleri için, test ölçümlerinin, geçersiz ya da sistematik hatalı olması olarak tarif edilir. Azınlık-çoğunluk, zenci-beyaz gruplar gibi etnik özellikler, kırsal-kentsel köken gibi farklı yaşam bölgelerinde bulunma, konuşulan dildeki farklılıklar, kadın-erkek olma gibi cinsiyet koşulları ölçümleri alınan grupların sistematik özellikleridir. Yanlılık belli bir grubun üyeleri için test sonuçlarında bir bozulmaya neden olması nedeniyle sistematiktir. Bir grup insanın bir koşudaki ölçme hızlarının kronometre ile ölçüldüğünü düşünelim. Bu grup siyahlar ve beyazlar olmak üzere iki alt gruba sahip olsun ve siyahlar için kullanılan kronometre yavaş çalışsın. Bu durumda yapılan ölçme sonucunda siyah ve beyaz grupların kendi içindeki sıralamaları görece doğru olabilir. Ancak, gruplardaki en iyi performansı gösteren bireylerin ve iki

grubun ortalama performanslarının karşılaştırılmaları için ölçme aracının yanlı olduğu belirtilir (Camilli ve Shepard, 1994; Korkmaz, 2005).

Örnekte de görüldüğü gibi test maddeleri belli bir grup için ölçülen yapı ile ilgisiz güçlük kaynakları içermesi nedeniyle yanlı olmaktadır. Yanlılıkla ilgili akılda tutulması gereken önemli bir nokta da, Gruplardan hiçbirinin etkilenmemesinin yanında her iki grubun da eşit etkilenmesi durumunun da yanlılık olarak tespit edilemeyecek olmasıdır. Yani eğer bir madde herkes için tehdit oluşturuyorsa, yanlılık olarak belirlenemeyecektir (Zumbo,1999).

Burada gözden kaçırılmaması gereken diğer bir nokta yanlılığın yalnızca iki grup arası puan farkı anlamına gelmemesidir. Eğer öyle olsaydı, her madde cinsiyet, etnik grup vs. gibi herhangi bir gruba karşı yanlılık içerir ve yalnızca grupların yeniden tanımlanmasıyla yanlılıkla ilgili peş peşe farklı anlamlar çıkarılabilirdi (Osterlind,1983; Osterlind ve Everson,2009).

Hangi alt grubun üyesi olduklarından bağımsız olarak, aynı evrende eşit yeteneğe sahip bireyler için maddeyi doğru cevaplama olasılığı aynı olduğu zaman bir test maddesinin yanlılık göstermediği söylenebilir; ancak aynı yetenek düzeyinde farklı grup üyeleri için bir maddeyi doğru cevaplama olasılığı farklılaşıyorsa maddenin kesinlikle yanlı olduğunu söylemek doğru olmaz. Eğer bu gruplar için test puanı ortalamaları farklı bulunursa, ardından bu farklılığın testteki yanlılığı yansıtıp yansıtmadığı sorusu ve testte hangi belli maddelerin bu yanlılığın kaynağı olabileceği sorusu ortaya çıkar (Osterlind, 1983; Zumbo, 2007).

Yanlılığın tarihsel gelişim sürecinde de yukarıda açıklanan noktada kavramsal bir karmaşa olduğu erken fark edilmiş, yanlılığın yargı ve yorum içeren anlam ve kullanımından farklı bir parça olan istatistiksel gözlem kısmını anlatmak için yanlılıktan farklı bir terim kullanılmasını söyleyen bir takım önerilerde bulunulmuştur. Sonunda, değişen madde fonksiyonu terimi farklı gruplarda farklı istatistiksel özellikler gösteren maddelerin basit bir gözlemini anlatmak için alanyazına girmiştir (Holland ve Wainer, 1993).

1.1.2 Değişen Madde Fonksiyonu (DMF)

Değişen madde fonksiyonu, test ile ölçülen yapı açısından eşleştirilmiş iki grup için bir madde fonksiyonunun psikometrik olarak nasıl farklılaştığını gösterir. Holland ve Wainer (1993), DMF istatistiğini katılımcılardan karşılaştırılabilir iki grup arasındaki madde fonksiyonları arasındaki fark olarak tanımlarken, Robitzsch ve Rupp (2008) ise, DMF'nin aynı popülasyondan farklı alt gruptaki katılımcılar için madde işletim özelliklerindeki (çok yaygın olarak madde güçlüğü) koşullu bir farkı yansıttığını belirtmiştir. DMF basit olarak, bir gruptaki kişilerin bir maddeyi doğru cevaplamalarının başka bir grupta eşit bilgi düzeyindeki kişilerden daha sıklıkla gerçekleşmesi durumunu tarif etmek için kullanılan istatistiksel bir terimdir (Robitzsch ve Rupp, 2008).

Özünde, tüm DMF teknikleri bir takım test maddelerinin sistematik olarak katılımcıların kişisel özellikleri ya da testin merkezinde yer alan yapıyla ilgisiz olarak işaretlenmiş olup olmadığını belirlemeyi merkeze almıştır (Osterlind ve Everson, 2009). Test maddelerinin bir grup aleyhine problem oluşturup oluşturmadığını belirlemek için belirli bir eşleştirme ölçütüne göre farklı gruplardan katılımcılar eşlenir ve sonra tek tek test maddelerindeki farklı grupların nasıl davrandığı incelenir (Zumbo, 1999). Farklı grup fakat aynı yetenek seviyesinden bireyler maddeyi doğru cevaplamak için farklı olasılıklara sahip olduğunda bu maddeler için DMF tespit edilir (Finch ve French, 2007).

DMF incelemesi, test puanlarının yorumlanması ve test maddelerinin geliştirilmesinin daha eksiksiz bir biçimde yapılması beklentisi ile test maddelerinde ayrıntılı bir değerlendirme yapılmasına olanak sunar. Ancak, DMF ölçülen ya da değer biçilen yapılardaki gerçek grup farklılığını ortaya çıkarmak için tasarlanmış değildir. Sonuçta, gerçek farklılığı ölçmek ilk etapta yapılan değerlendirmenin amacıdır. Testi alan popülasyondaki alt gruplar arasındaki ortalama puan farkları kanıtlanmasına gerek görülmeden DMF olarak belirlenmemeli hatta DMF'nin bir işareti olarak da alınmamalıdır. DMF gözlenen puan farklılıklarının bir tahmini değildir. Bu fikirle yakından ilgili bir kavram ise DMF keşfindeki grup farklılığının kaynağının ölçülen yapı ile her zaman ilişkisiz olmasıdır. Yani; bir grubun performansını testteki

hedeflenen yapıdan ayrı bir faktör etkilerken diğer grubu etkilememiştir. Örneğin, eğer DMF kızlar ve erkekler arasında matematik maddelerinde ortaya çıkıyor ise bunun nedeni cinsiyetin ölçülen yapı ile ilişkisiz olmasıdır. Fakat, bu noktada önemli olan DMF analizi ile niçin cinsiyet ya da başka bir değişkenin performansta fark oluşturduğu bilinmemesi, yalnızca alt gruplarda sistematik olarak fark olduğunun ve bu farkın anlamlı olduğunun ortaya konulabilmesidir (Osterlind ve Everson,2009).

Osterlind ve Everson (2009), DMF sonuçlarının yorumlanmasıyla ilgili olarak test maddelerinin analizinin sadece psikometrik ve istatistiksel bilgiler gerektiren bir uygulama değil aynı zamanda iyi bir karar verme yeteneği hatta bilgelik gerektiren bir uygulama olduğunu belirtmişlerdir. Bu beceri ve yargılama gücüne nispeten basit olan madde analizinde bile ihtiyaç vardır. Ancak gruplar arası DMF analizi yapılırken bu beceri ve karar verme yetenekleri daha önemli hale gelmektedir. Çünkü, DMF çalışmasındaki bulguları yorumlamak her zaman kolay olmamaktadır. Yaygın olarak, farklı DMF belirleme stratejileri kullanıldığında farklı sonuçlar gözlenmekte ve istatistiksel olarak DMF gösteren maddelerin içeriğine bakıldığında DMF göstermesi için bir sebep bulunamamaktadır. DMF'nin basit bir aritmetik hesaplama probleminde gözlenmesi durumunda olduğu gibi.

O halde, bir maddede DMF bulmak o maddenin yanlı yani gruplardan birine mutlaka adaletsiz davrandığı anlamına gelmediğinin üzerinde durulması gerekmektedir. Ironson (1982) bir kek pişirmek için gereken sıcaklığı soran bir madde örneği vermiştir. Madde istatistiksel olarak erkeklere karşı yanlı olsa bile eğer aşçı seçmek için kullanılan bir testteki madde ise oldukça uygun olabilir (Holland ve Wainer,1993).

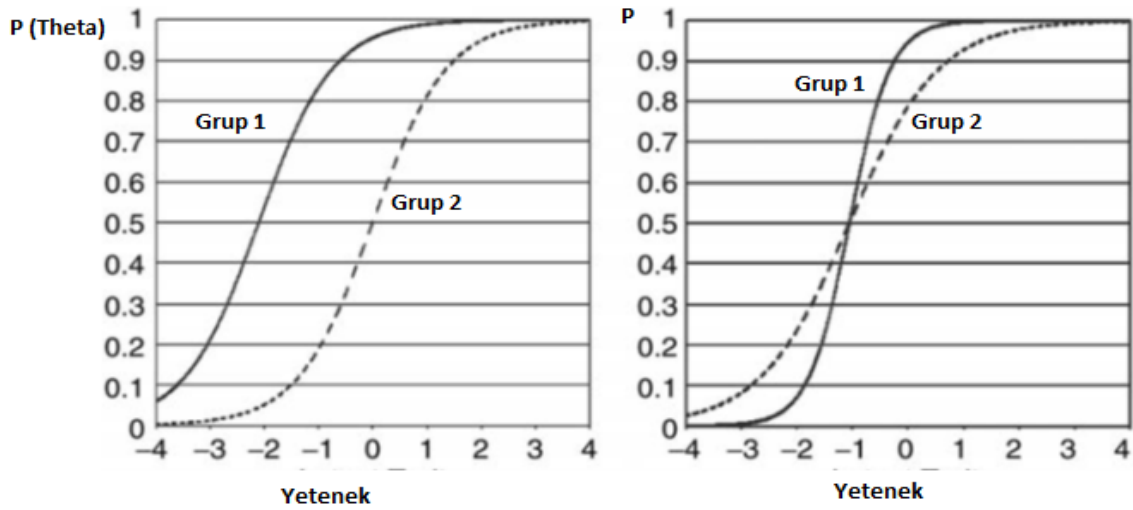
Yanlılık ve etki arasında ayırım yapmak önemlidir (Holland ve Wainer,1993; Robitzsch ve Rupp, 2008). Madde etkisi, maddeler tarafından ölçülen yetenek bakımından gruplar arasında gerçek bir farklılık olması nedeniyle DMF'nin ortaya çıktığı durumları anlatmak için kullanılır. Madde yanlılığı ise ilgilenilen söz konusu yetenek ile ilgisiz olarak, test maddelerinin bazı karakteristik özellikleri nedeniyle DMF'nin ortaya çıktığı durumlarda kullanılır (Zumbo,2007). DMF madde yanlılığı için gerekli ancak yeterli olmayan bir şarttır. Yani, eğer bir madde için DMF bulunamıyorsa o maddede yanlılık da yoktur. Ancak, eğer DMF var ise, bu durum madde yanlılığını bildirmek için yeterli değildir; hatta madde yanlılığının varlığını belirleyebilmek için, takip eden madde

yanlılığı analizlerine (içerik analizi, ampirik değerlendirmeler gibi) başvurulması zorunludur (Zumbo,2009).

DMF ile ilgili başka bir önemli husus, araştırılan puan farklarının sistematik olması gerekliliğidir. Yani, alt gruplar arası farklılık aynı ve öngörülebilir bir biçimde tüm (ya da neredeyse tüm) kişileri etkilemesi gerekir. Analiz edilebilecek grup sayısı DMF belirleme yöntemine bağlıdır; fakat pratikte sıklıkla bir grubun referans diğer grubun odak grup olarak ifade edildiği iki gruba analiz yapılır. DMF bir test maddesindeki bireyin performansına başvurmaz, o bir grup fenomenidir (Osterlind ve Everson, 2009).Çünkü; kişi bir testte bildiğini göstermekte güçlük yaşamış olabilir, ancak bu problem bütün katılımcılar için ya da belirli bir alt grup için yaşanmadıkça testle ilgili bir problem olduğu söylenemez.

Tek Biçimli(TB) ve Tek Biçimli Olmayan (TBO) DMF

DMF'nin yaygın olarak kullanılan tek biçimli (TB) ve tek biçimli olmayan (TBO) şeklinde iki tipi olduğunu söylenebilir. DMF'nin etkisi inceleme altında olan madde tarafından bir grup katılımcının sürekli haksız yere dezavantajlı olması anlamına gelen TB DMF, ya da ölçeğin bazı noktalarında bu ilişkinin tersine dönmesi anlamına gelen TBO DMF olabilir (Robitzsch ve Rupp, 2008). Başka bir deyişle, TB DMF tüm yetenek düzeyleri boyunca bir grubun sabit olarak daha avantajlı olması anlamına gelirken, TBO DMF'de avantaj sabit değildir ve bir gruptan diğerine geçiş yapabilir (Finch ve French, 2007).



Şekil 1: Tek Biçimli ve Tek Biçimli Olmayan DMF

Şekil 1’de görüldüğü gibi, madde karakteristik eğrisi (MKE) iki farklı grubun cevaplarını karşılaştırabilmenin bir yolunu sunar. MKE’ler arasındaki fark sınava giren kişilerden oluşan iki grubun test maddesini doğru cevaplayabilme olasılığının eşit olmadığını gösterir. TB DMF’de işaretli madde θ (yetenek) dağılımı boyunca referans grup için tutarlı bir avantaj sağlarken, TBO DMF ise koşullu bağımlı değişkenin, sürekli θ değişkeni üzerinde farklı noktalarda derece ve yön değiştirdiği zamanlarda ortaya çıkar. Bu durumda TB DMF’de iki grup için MKE’ler farklı ve kesişmezken TBO DMF’de ise aksine MKE’ler θ ölçeğindeki bazı noktalarda iki grup için farklı fakat keşmiş konumdadır. İki MKE arasındaki alan ise DMF’nin derecesini gösterir.

Test geliştiricilerin sıklıkla keşfettikleri gibi, bir madde θ ’nın bir seviyesinde referans gruba küçük ve ihmal edilebilir bir avantaj sağlarken θ ’nın daha yüksek bir seviyesinde daha büyük bir avantaj sağlayabilir. Ya da bir maddede Odak grubun θ ölçeğinin bir ucunda göreceli avantaja sahip iken, θ sürekli değişkeninin diğer ucunda da referans gruba bir avantaj sağlayabilir. O halde denebilir ki; TBO DMF grup üyeleri ve yetenek seviyesi (θ) arasında bir etkileşim olduğu zaman ortaya çıkar (Osterlind ve Everson, 2009).

Şekil 1’de, her iki grubun madde tepki fonksiyonları aynı ayırıcılık değerine sahiptir; fakat yetenek ekseninde öylesine belirlenmiş herhangi bir noktada, grup 2’nin maddeyi doğru cevaplama olasılığı grup 1’den daha düşüktür. Madde 2 ise çaprazlanmış DMF gösterir, yani; yetenek eksenindeki düşük değerler için grup 1’in maddeyi doğru cevaplama olasılığı grup 2’den daha yüksekken, yüksek yetenek değerlerinde tam tersi bir durum ortaya çıkar (Finch ve French, 2007).

1.1.3 Tarihsel Süreçte Yanlılık Belirleme Teknikleri

Madde yanlılığı çalışmalarının ilk olarak 1910 yılında Alfred Binet ile başladığı düşünülmektedir. Çünkü; Binet, bir IQ testindeki bazı test maddelerinin zihinsel kapasiteden daha çok kültürel eğitimin etkilerini ölçtüğünden şüphelenmiş ve düşük sosyoekonomik düzeyden çocuklarla çalışarak bu durumu test etmeye çalışmıştır. Ayrıntılı araştırmalar sonucunda ise belli kategorilerdeki maddeleri testten çıkarmıştır. Genel olarak atılan maddelerin tamamen evdeki eğitim, dikkat, dil, resimlere bakma alışkanlığı ve bilimsel egzersizlere bağlı olduğu düşünülmüştür.

Ardından, 1912’de William Stern Almanya’daki sınıf farklılıkları üzerine çalışmalar yapmıştır. O da Binet gibi önemli düzeyde farklılıklar tespit etmiştir.

Test yanlılığındaki modern çalışmaların başlangıcı ise 1951 yılında Eells, Davis, Herrick ve Tyler tarafından Şikago’daki Eells Üniversitesi’nde tez çalışma konusu olarak başlatılmıştır. Eells ve arkadaşları ilk olarak ölçme sonucunda elde edilen bazı farklılıkların yeteneğin doğru bir yansıması olamayabileceği olasılığı üzerine odaklanmışlardır. Çocukların IQ’larında gözlenen farklılıkların testteki maddelerin belirli bir içeriğinden kaynaklanıyor olabileceğini ve çocuklarla ilgili ölçmesi hedeflenen herhangi önemli bir yeteneği doğru yansıtmıyor olabileceğini belirtmişlerdir. Araştırmalarında artık önerilmeyen bir yöntem olan, analiz edilen maddelere verilen doğru cevap yüzdelerindeki grup farklılıklarını kullanmalarına rağmen, yine de IQ testinin içerdiği maddelerle ilgili önemli sorular sormuş, güzel noktalara değinmişlerdir. Ayrıca Eells ve ark. sosyal sınıflardaki etnik grupları karşılaştırarak test performanslarındaki etnik geçmiş ve sosyal sınıf etkilerine ilk değinme girişimini gösteren araştırmacılar olmuşlardır.

Sonraki süreçte ilgili tartışmaların devam etmesine rağmen, Eells ve ark. çalışması gibi birçok çalışmanın ortaya çıkması ve uygulanan metotlardaki gelişme ile testlerdeki yanlı maddelerin belirlenmesi ve testlerden arındırılması işlemlerinin rutin olarak uygulanması süreçlerindeki önemli gelişmeler, sivil haklar hareketinin istihdam ve eğitim fırsatlarına erişimin eşit olmadığını iddia edip, bu konuda testlerin önemini fark etmesiyle, ancak yirmi yıl sonra gerçekleşebilmiştir. 1960 ve 1970’lerin sonlarında yanlılık çalışmaları sadece IQ testlerinde değil izleme ya da yerleştirme yapmak için kullanılan eğitimsel testlerde de yaygın olarak kullanılmaya başlanmıştır (Camilli ve Shepard, 1994).

Bir testteki madde yanlılığının varlığını ve derecesini değerlendirmek için yapılan ilk formal prosedür 1963 de SAT sınavına giren katılımcılardan beyaz ve siyah gruplar ile madde etkileşimini test eden Cardall ve Coffman (1964) tarafından uygulanan varyans analizi prosedürüdür. Aynı prosedür Cleary ve Hilton (1968) ve ardından Angoff ve Sharon (1974)’nın çalışmalarıyla takip edildiyse de, yöntemin alanda çok geniş bir kabul görmediği söylenebilir.

Ardından oldukça önemli bir gelişme olarak, F.M. Lord'un (1952) tezinde "Madde Tepki Kuramı" olarak açıklanan test kuramı tarihindeki temel bir buluş yayınlanmıştır. Sonrasındaki kısa bir zaman içinde bu modelin değişen madde fonksiyonu çalışmalarında kullanılmasının da uygun olabileceği belirlenmiştir. İyi bilindiği üzere kuramın temeli farklı gruplara ait belirli bir θ yetenek düzeyindeki bireyler için çizilen madde tepki fonksiyonuna dayanmaktadır. MTK kuramında eğer maddeler DMF gösteriyorsa, madde karakteristik eğrileri de gruplar için tanımlanabilir bir şekilde farklı olacaktır. Madde karakteristik eğrileri iki yaygın yolla farklı çıkabilir. İlki, eğriler yalnızca eşik parametrelerinde farklılaşabilir (güçlük gibi), ve bundan dolayı θ sürekli değişkeni üzerindeki konumları bir kayma ile değişir. İkincisi eğriler yalnızca güçlükte değil aynı zamanda ayırıcılıkta da farklılaşabilir (ve/veya tahmin şans parametresinde) ve bu nedenle eğrilerde kesişim görülür. Bu bağlamda, ilk tarif edilen TB DMF ve son sunulan ise TBO DMF'dir (Holland ve Wainer, 1993; Zumbo, 2007).

Diğer madde yanlılığı yöntemlerinin aksine 3 parametrelili MTK kullanımı yalnızca gruplar arasındaki farkı madde güçlüğüne göre değil aynı zamanda ayırıcılık ve şans başarısına göre de belirleyen kapsamlı bir çalışmadır. Ancak yine de tutarlı bulgular elde etmek için büyük örneklemelere ihtiyaç duyması, maliyetinin yüksek olması, modelin kendisinin karmaşık olması ve modelin varsayımlarını sağlamak için kullanılan operasyonel testlerde sıklıkla hata yapılması nedenlerinden ötürü standart bir yöntem olarak kullanımı kısıtlanmaktadır.

Başka bir DMF belirleme yöntemi olan MTK'nın tek parametrelili bir modeli olan Rasch modeli (1960) tüm maddelerin eşit ayırıcılık gücüne sahip olduğu ve şans eseri doğru cevaplanmadığı varsayımları üzerine kuruludur. Bu nedenle bu tekniğin maddelerdeki yanlılığın tespiti için kullanımı madde güçlüklerindeki farkın değerlendirilmesiyle sınırlıdır. Bu yöntemin sınırlılığı a ve c parametrelerinde herhangi bir farklılık olmadığı varsayımından ötürü, bunlardaki herhangi bir gerçek farklılığın da bu model ile tespit edilemeyecek olmasıdır. Daha da önemlisi a ya da c parametresinde gerçekten bir farkın var olması yüksek ihtimalle DMF değeri oluşmasına neden olacaktır. Shepard ve arkadaşları (1981), Rasch modeli indeksi ile dönüştürülmüş madde güçlüğü indeksi arasında çok yüksek bir korelasyon (0.99) bulmuşlardır. Bu durum Rasch modelinden elde edilen b parametresi ile dönüştürülmüş madde güçlüğü yönteminden elde edilen

delta değerinin doğrudan orantılı olduğunu göstermektedir (Camilli ve Shepard, 1994 ; Holland ve Wainer, 1993).

1979’da Scheuneman DMF değerlendirmek için ki kareye benzeyen bir yöntem yayınlamış ancak bu yöntem daha sonra Baker (1981) tarafından örneklem büyüklüğünden ve bilinmeyen bir örneklem dağılımından çok etkilenmesi nedeniyle eleştirilmiştir. Ardından düzeltilmiş istatistiği olan tam ki kare yöntemi Bishop, Fienberg ve Holland (1975) tarafından bulunmuştur. Bu istatistik, grupları toplam puan aralıklarına bölmekte (tavsiye edilen 4 ya da 5 aralıktır) ve her bir aralık için 2X2’lik tablolar hazırlamaktadır. Tablonun bir ekseninde grup üyelikleri diğer ekseninde ise maddeyi doğru cevaplayıp cevaplayamama durumu gösterilmektedir.

Scheuneman’ın (1979) gelişmesinden dokuz yıl sonra Holland ve Thayer (1988), Mantel ve Haenszel (1959) tarafından geliştirilen bir tekniğin DMF araştırmaları için de kullanılabileceğini belirtmişlerdir (Holland ve Wainer, 1993).

Tarihsel süreçten de anlaşıldığı gibi DMF incelemeleri için tek bir teknik ya da yöntem söz konusu değildir. Ancak, hangi yöntem dikkate alınırsa alınsın, yöntemlerin hepsi grupları maddeler tarafından ölçülen yeterlilikte doğrudan ya da dolaylı olarak eşleştirmek ve ardından farklı gruplara fakat aynı yetenek seviyesine ait bireylerin doğru cevap verme olasılıklarının kıyaslanması üzerine tasarlanmıştır. (Osterlind ve Everson, 2009; Holland ve Wainer, 1993; Bakan Kalaycıoğlu ve Berberoğlu, 2010).

Alanyazında DMF yöntemlerini sınıflamada farklı bakış açıları ve buna bağlı oluşan farklı sınıflama biçimleri mevcuttur.

Zumbo (2009) sınıflandırmayı;

- olasılık tablosuna dayalı yöntemler
- madde tepki kuramına dayalı yöntemler
- çok boyutlu yöntemler biçiminde yaparken,

Camilli ve Shepard (1994) ise;

- varyans analizine dayalı yöntemler
- MTK’na dayalı yöntemler
- olasılık tablosuna dayalı yöntemler olmak üzere üçe ayırmıştır.

Tekniklerin birbirlerine göre pek çok üstünlüğü ya da sınırlılığı olduğu göz önüne alındığında, bir tekniğin diğerine göre daha tercih edilebilir olması görüşünün, durumu belirlemede kesin bir yol olmadığı sonucuna varılabilir. Son olarak ilgili alan yazından elde edilen sonuçlar değerlendirildiğinde daha güvenilir sonuçlar elde etmek için birden çok DMF belirleme tekniğinin bir arada kullanılması ve farklı tekniklerin hepsinde tespit edilen maddelerin DMF gösteren madde olarak karşılaştırılması gerektiği ifade edilebilir.

1.1.4 Olasılık Tablosuna Dayalı Yöntemler

Sınıflandırmalarında olasılık tablosu terimini kullanan Camilli ve Shepard (1993), bu başlık altındaki yöntemlerin özel bir analiz modeli olduğundan değil verinin tablolaştırılmasını öneren bir biçimi olmasından bu terimi kullandıklarını belirtmişlerdir. Bu yöntemlerde gözlenemeyen örtük özelliklerden çok ölçekle gözlenen değerler sunulmaktadır ve MTK modellerinin gerektirdiği yüksek düzeyde matematik ön bilgisi gerektirmemektedir. Bu durumun DMF analizi sonuçlarının kitlelere duyurulması noktasında bir avantaj sağlayabileceği düşünülmektedir.

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 katılımcı ile ilgili

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

bilgilerinin bilinmesi zorunludur. Toplam puan her bir katılımcı 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 katılımcılar birlikte gruplandırılmakta ve her bir grup için sıklıkla tercih edilen biçimi ile 3 yollu olasılık tablosu biçiminde tablo organize edilmektedir. Söz konusu veri her bir katılımcının odak ya da referans grup üyesi olmasına ve maddeyi doğru ya da yanlış cevaplamasına

göre çapraz sınıflandırılır. Örneğin Tablo 2 2X2'lik olasılık tablosudur ve toplam puan sürekli değişkeninin bir j kategorisini sunar.

Tablo 2: j yetenek seviyesindekiler için bir maddenin olasılık tablosu

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

Tablo 2’de R referans grup, F odak grup, A_j referans gruptan C_j odak gruptan maddeyi doğru cevaplayanların sayısı B_j referans grup D_j ise odak gruptan maddeyi yanlış cevaplayanların sayısını M_{1j} j yetenek seviyesindeki tüm gruptan maddeyi doğru cevaplayanların sayısını M_{0j} yanlış cevaplayanların sayısını N_{rj} j yetenek seviyesinde referans grupta bulunan N_{fj} odak grupta bulunankişi sayısını ve T_j ise j yetenek seviyesindeki kişilerin toplam sayısını göstermektedir. Tüm katılımcıların aynı toplam puana (j) sahip olması nedeniyle bu katılımcıların kendi aralarında kıyaslanabilir olduğu düşünülmektedir. Toplam puan şartı kontrol altına alındıktan sonra, ardından odak ve referans gruptan bireylerin maddeyi doğru cevaplama şanslarının eşit olup olmadığını sorusu ile karşılaşılır. Tablo 2’deki veri üzerinden yapılan basit bir takım hesaplamalardan sonra ise bu soruya cevap olacak uygun bilgilere ulaşılır.

$P_{rj} = A_j/N_{rj}$ referans grubun soruyu doğru cevaplama oranını ve benzer biçimde $q_{rj} = B_j/N_{rj}$ yanlış cevaplama oranını gösterecek biçimde $P_{rj} + q_{rj} = 1$ ‘dir. Bu hesaplamalar odak grup için de yapıp P_{fj} ve q_{fj} değerleri de hesaplandığında ise Tablo 3 elde edilir.

Tablo 3 : j yetenek seviyesindekiler için bir maddenin olasılık tablosu

	1	0	Toplam
R	P_{rj}	q_{rj}	1
F	P_{fj}	q_{fj}	1

Veri düzenlemesinin son noktasında Tablo 2’deki veri Tablo 1.3’deki oranların gösterildiği bir tablo haline gelmektedir. Olasılık tablosuna dayalı metotlar

kategorisaltına giren tüm DMF belirleme metotlarında, MH ve LR de dahil, bu tablonun kullanımı faydalı olmaktadır (Camilli ve Shepard, 1994).

Mantel Haenszel (MH)

Mantel ve Haenszel (1959) tarafından eşleştirilmiş gruplarla çalışmak için yeni bir yöntem tanıtılmıştır. Holland (1985) ve sonrasında da Holland ve Thayer (1988) tarafından ise bu yöntem değişen madde fonksiyonu değerlendirilmesine uyarlanmıştır. Bu uyarılama ETS’de ilk DMF belirleme aracı olarak kullanılmıştır (Holland ve Wainer,1993).

MH yöntemi olarak, DMF değerlendirilmesinde eşleştirilmiş gruplar şartı için gerekli ayarlamaların yapılması adına odak ve referans gruplar, toplam test puanlarına dayalı olarak yetenek ya da yeterlik tabakalarına bölünür. Genellikle 4 ya da 5 yetenek seviyesi geliştirilir. Sonra her yetenek seviyesi için bir ki kare olasılık tablosu hazırlanır. Tabloda her bir yetenek seviyesindeki gruplar için doğru ve yanlış cevapların frekansları sunulur.

Tablo 4 : Herhangi bir m seviyesi için hazırlanmış ki kare tablosu;

GRUP	MADDE PUANI		
	DOĞRU	YANLIŞ	TOPLAM
ODAK GRUP	D_{om}	Y_{om}	T_{om}
REFERANS GRUP	D_{rm}	Y_{rm}	T_{rm}
TOPLAM	D_{tm}	Y_{tm}	T_{tm}

Tablo 4’te sırası ile D_{om} ve D_{rm} m yetenek seviyesinde odak gruptan ve referans gruptan maddeyi doğru cevaplayanların sayısını benzer biçimde Y_{om} ve Y_{rm} ise odak ve referans gruptan maddeyi yanlış cevaplayanların sayısını, D_{tm} m yetenek seviyesinde maddeyi doğru cevaplayanların Y_{tm} ise maddeyi yanlış cevaplayanların sayısını, T_{om} m yetenek seviyesinde odak gruptaki T_{rm} ise referans gruptaki toplam kişi sayısını ve son olarak T_{tm} ise m yetenek seviyesindeki toplam kişi sayısını göstermektedir. Buna göre sunulan bilgilere bağlı olarak geliştirilen H_0 hipotezi $m=1, \dots, M$ ve M =toplam puanı gösterecek şekilde aşağıdaki biçimde açıklanabilir;

$$H_0 = \frac{D_{rm}/Y_{rm}}{D_{om}/Y_{om}} = 1$$

$$H_0 = \frac{D_{rm}}{Y_{rm}} = \frac{D_{om}}{Y_{om}}$$

Sözel biçimiyle ise, eşleştirme değişkenindeki (toplam puan) tüm m değerleri boyunca odak ve referans gruplarının her ikisinin de eşleştirme değişkeninin verilen seviyesindeki maddeyi doğru cevaplama olasılığı aynıdır denilebilir.

Ho hipotezine alternatif olarak geliştirilen hipotez isem=1,...,M ve $\alpha \neq 1$ olacak biçimde şu şekilde belirlenir;

$$H_1 = \frac{D_{rm}}{Y_{rm}} = \alpha \frac{D_{om}}{Y_{om}}$$

Hipotezlerden de anlaşıldığı üzere denebilir ki; eğer gruplar arasında hiç fark yoksa, odak ve referans gruplar arası denkliliği gösterir biçimde bu olasılık oranı 1'e eşit olur ve DMF gözlenmemiştir diye yorumlanır. Ancak, $\alpha_m > 1$ ise m yetenek seviyesinde söz konusu maddede referans grubun performansı odak gruptan anlamlı bir biçimde daha iyidir ve tersi olarak $\alpha_m < 1$ ise odak grup anlamlı bir biçimde yetenek seviyesinde referans grubu geride bırakmıştır biçiminde yorumlanır.

Mantel ve Haenszel tarafından α sürekli olasılık oranının bir kestiricisi elde edilmiştir. α_{mh} biçiminde gösterilen bu kestiricinin eşitliği şu biçimdedir;

$$\alpha_{MH} = \frac{\sum_m \frac{D_{rm}Y_{om}}{T_{tm}}}{\sum_m \frac{D_{om}Y_{rm}}{T_{tm}}}$$

Bu değer 1'e eşit olduğu zaman Ho'ı işaret etmek üzere, sıfır sonsuz aralığında değerler alabilen DMF etki büyüklüğünün bir kestiricisidir. Bu değer yorumlanması zordur bu nedenle logaritmik dönüşümlerle daha uygun bir ölçeğe dönüştürülür. Bu yeni ölçek (MH_{D-DIF}), sıfır civarında simetrik ve yorumlanması daha kolaydır. $MH_{D-DIF} = 0$ DMF'nin yokluğu olarak yorumlanır. MH_{D-DIF} pozitif değerdeyse madde odak grup lehine çalışmıştır, negatif değerde ise madde referans grup lehine ve odak grup aleyhinde çalışmıştır denilmektedir (Osterlind ve Everson, 2009; Holland ve Wainer, 1993).

Çeşitli derecelerde DMF gösterdiği belirlenmiş test maddelerini belirlemede MH_{D-DIF} ölçümü kullanılırken de ETS’de test geliştirmede maddeleri 3 kategoriden birine yerleştirmede kullanmak üzere bir sınıflama planı geliştirilmiştir. Bu üç kategori önemsiz (A), orta derce (B) ve büyük (C) olarak belirlenmiştir. Bu kategoriler ve karar verme ölçütleri Tablo 3’te sunulmuştur (Holland ve Wainer, 1993).

Tablo 5: MH_{D-DIF} Değerlerini Yorumlama Ölçütleri

Kategoriler	ΔR^2 Değeri	Yorum
A	$ MH_{D-DIF} < 1$	DMF yoktur veya göz ardı edilebilecek düzeydedir.
B	$1 \leq MH_{D-DIF} \leq 1,5$	Orta düzeyde DMF vardır
C	$ MH_{D-DIF} > 1,5$	Yüksek düzeyde DMF vardır.

MH ve LR arasındaki bir belirgin fark olarak; MH yöntemi etkileşim olmadığını varsayarken LR etkileşime izin vermektedir (Zumbo, 2007).

Lojistik Regresyon (LR)

Lojistik Regresyon bağımlı değişkenin sürekli ya da nicel bir değişken olmadığı, bir diğer deyişle kategorik ya da sınıflamalı olduğu durumlar için uygun bir analiz türüdür. DMF çalışmasında LR yöntemi, maksimum olabilirlik yöntemleri ile tahmin edilen olasılık fonksiyonuna başvurur (Osterlind ve Everson, 2009). Grup üyeliği ve bir ölçüt ya da şart değişkenine bakarak bir maddeye doğru cevap verme olasılığının istatistiksel modellenmesine dayanır. Bu ölçüt değişkeni genellikle ölçek ya da alt ölçeğin toplam puanı, fakat bazen de aynı değişkenin farklı bir ölçümüdür.

LR tekniğinde madde cevapları (1 ya da 0) bağımlı, grup değişkeni (referans grup=1 ve odak grup=2 olacak biçimde kodlanmış), her bir grup için toplam ölçek puanı ve grup ile toplam ölçek puanının etkileşimini de bağımsız değişkenler olarak kullanılmaktadır. TB DMF için grubun etkisini, TBO DMF için ise grup-toplam puan etkileşimini test eden bir DMF testi sağlayacaktır.

LR eşitliği aşağıdaki biçimdedir;

$$Y = b_0 + b_1TOP + b_2GRUP + b_3TOP*Grup + \varepsilon_i$$

Burada Y değişkeni olasılık oranının doğal bir logaritmasını göstermektedir. Eşitlik başka bir anlatımla şöyledir;

$$\ln \left[\frac{pi}{(1 - pi)} \right] = b0 + b1tot + b2group + b3(tot * group)$$

Eşitlikte p örtük özellik yönündeki maddeyi doğru yanıtlayan bireylerin oranını göstermektedir. Sonrasında TB ve TBO DMF'nin her ikisi için de 2 serbestlik derecesinde ki kare testi ile test edilebilir (Zumbo,1999).

LR'de 3 aşamalı bir modelleme söz konusudur ve modelin testi 2 adımda gerçekleşir. Adım 1'de model 1 model 2'ye karşı test edilir. Model 1'de model 2'den farklı olarak etkileşim etkisi adı verilen bir fazla parametre vardır. Modeller arası fark serbestlik derecesi 1 olan LR test istatistiği ile test edilir. Etkileşim terimi manidar çıkmazsa diğer adıma geçilir.

Adım 2'de model 2 model 3'e karşı test edilir. Burada da model 2 model 3'ten bir fazla katsayıya sahiptir ve özünde madde cevaplarını tahmin etmede bu eklenen terimin gerekli olup olmadığı test edilir. Yine serbestlik derecesi 1 olan bir LR testi kurulur. Buradaki soru, yeteneğin ötesinde grup üyeliğinin maddeyi doğru cevaplamaya bir etkisinin olup olmamasıdır. Bu istatistik grup ve yetenek etkileşimi terimini içermediği için β_{mh} terimi ile benzerlik göstermektedir ve her ikisi de TB DMF varlığını belirlemektedir.

Söz konusu aşamaların her birinde ki-kare değeri elde edildiği gibi aynı zamanda R^2 değeri de hesaplanır. İki model arasındaki R^2 değerlerinin farkı olan ΔR^2 istatistiği DMF miktarının yorumlanmasında kullanılan bir etki büyüklüğü belirtir.

ΔR^2 'nin DMF yorumlanmasında kullanılmak üzere farklı biçimde kategorilendirilmeler mevcuttur. Zumbo ve Thomas (1996) tarafından önerilen sınır değerler Tablo 6'da görüldüğü gibidir;

Tablo 6: Zumbo ve Thomas 'ın ΔR^2 değerlerini yorumlama ölçütleri

Kategoriler	ΔR^2 Değeri	Yorum
A	$\Delta R^2 < 0.13$	DMF yoktur veya göz ardı edilebilecek düzeydedir.
B	$0.13 \leq \Delta R^2 \leq 0.26$	Orta düzeyde DMF vardır.
C	$\Delta R^2 \geq 0.26$	Yüksek düzeyde DMF vardır.

Tablo 7'de ise Jodoin ve Gierl (2001) tarafından belirlenen sınır değerler verilmektedir.

Tablo 7: Jodoin ve Gierl'in ΔR^2 değerlerini yorumlama ölçütleri

Kategoriler	ΔR^2 Değeri	Yorum
A	$\Delta R^2 < 0.035$	DMF yoktur veya göz ardı edilebilecek düzeydedir.
B	$0.035 \leq \Delta R^2 \leq 0.070$	Orta düzeyde DMF vardır.
C	$\Delta R^2 \geq 0.070$	Yüksek düzeyde DMF vardır.

Zumbo (1999), LR yönteminin kullanımındaki yararları şu biçimde sıralamaktadır;

- Ölçüt değişkeni (çoğunlukla toplam puan) kategori etmeye ihtiyaç yoktur.
- TB ve/veya TBO DIF i modelleyebilir
- İkili puanlanan maddeler için kullanılan LR modeli sıralı puanlanan maddeler için de yaygınlaştırılabilir.

Diğer yandan ampirik çalışmalardaki kullanımında, olasılık tablosuna dayalı yaklaşımlar sıklıkla MTK ile bilinen DMF belirleme yöntemleri kadar iyi çalıştığı tespit edilmiştir. Bu nedenle, DMF hesaplamak için kullanılan olasılık tablosu yöntemleri kullanıcıların MTK yöntemlerini uygulamak için gereken bilgisayar kaynakları ve teknik bilgi eksikliği sebebi ile uygulamada tercih nedeni olabileceği düşünülmektedir.

Bu çalışmada DMF analizi için Mantel Haenszel ve Lojistik Regresyon teknikleri tercih edilmiştir. Literatürde bu tekniklere ait yararlar;

- Pratikte çokça kullanılmaları
- Yaygın yazılım paketleri sayesinde uygulamada kolaylık sağlamaları
- MTK yöntemlerine göre daha küçük örneklem büyüklükleri gerektirmeleri

Biçiminde sıralanmaktadır (Clauser ve Mazor,1998; Camilli ve Shepard,1994; Zumbo,2009).

1.2 ARAŞTIRMANIN AMACI VE ÖNEMİ

Bu çalışmada, Uluslararası Öğrenci Değerlendirme Programı fen okuryazarlığı alt testindeki maddelerin cinsiyet değişkenine göre değişen madde fonksiyonu gösterip göstermediğinin araştırılması amaçlanmıştır.

Dünya genelinde politika belirleyiciler kendi ülkelerindeki öğrencilerin bilgi ve beceri düzeylerini, projeye katılan diğer ülkelerdeki öğrencilerin bilgi ve beceri düzeyleri ile karşılaştırmak, eğitim düzeyinin yükseltilmesi amacı ile standartlar oluşturmak ve eğitim sistemlerinin güçlü ve zayıf yönlerini belirlemek için PISA sonuçlarını kullanmaktadır. Diğer yandan ölçekte yer alan maddeler için tespit edilecek DMF'nin yanlılığın bir işareti olabileceği, yanlılığın ise gruplar arasında sistematik hata yapılmasına sebep olarak ölçülmesi hedeflenen yapı ile ilişkisiz değişkenler karıştırarak, geçerliğe zarar verdiği bilinmektedir. Geçerliği sağlanamamış bir ölçekten anlam çıkarmak ve yorumlamalara gitmek ise faydasız ve anlamsız bir uğraştır.

Diğer yandan literatürde sayısal alanlardan özellikle de matematik alanında cinsiyet grupları arasında ilgi ve yapıdan kaynaklanan farklar olduğunun bu nedenle yanlılık analizlerinde ilgili konuya (geometri, uzamsal zeka gerektiren maddeler ya da basit algoritmik problemler) bağlı olarak DMF gösteren maddelerin tespit edildiğinin altını çizen çalışmaların sayısı oldukça çoktur. Benzer bir durumun fen alanı için de geçerli olduğu düşünülmektedir.

Bu bağlamda sonuçları ile ülke eğitim politikalarını etkileyen bu denli önemli bir sınav için, PISA fen okuryazarlığı alt testinde, DMF incelemelerinde ve eğitim politikalarında ana odak noktalardan biri olan cinsiyet değişkeni bakımından DMF incelemesi yapmak önem taşımaktadır. Ayrıca DMF analizlerinin, ölçeklerdeki geçerlik kanıtı olarak ölçme değerlendirme çalışmalarının temel gerekliliği haline gelmesinin de birçok açıdan faydalı olacağı düşünülmektedir.

1.3 PROBLEM CÜMLESİ

Uluslararası Öğrenci Değerlendirme Programı (PISA) fen okuryazarlığı alt testindeki maddeler Türkiye örnekleminde cinsiyet gruplarına göre değişen madde fonksiyonu göstermekte midir?

1.3.1 Alt Problemler

- 1) PISA 2009 yılı uygulamasında fen okuryazarlığı alt testindeki maddeler cinsiyete göre Mantel Haenszel yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
- 2) PISA 2009 yılı uygulamasında fen okuryazarlığı alt testindeki maddeler cinsiyete göre Lojistik Regresyon yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?
- 3) Lojistik Regresyon ve Mantel Haenszel yöntemleriyle yapılan analizlerde DMF ile ilgili elde edilen sonuçlar uyum göstermekte midir?

1.4 SAYILTILAR

Öğrencilerin okuma becerileri testindeki maddelere verdikleri cevapların gerçek durumlarını yansıttığı varsayılmıştır.

1.5 SINIRLILIKLAR

Bu araştırmada değişen madde fonksiyonu analizi Mantel Haenszel ve Lojistik Regresyon yöntemleri ile sınırlı tutulmuştur.

Araştırmada, PISA uygulamasına katılan her öğrencinin aynı soruyu görmemesi durumundan kaynaklı, DMF analizine alınan maddeler bir kitapçıkla sınırlı tutulmuştur.

1.7 İLGİLİ ARAŞTIRMALAR

Doolittle ve Welch (1989), Amerika’da uygulanan üniversite akademik yeterlik değerlendirilmesi sınavının beş farklı alt testi üzerinde yaptıkları çalışmalarında MH yöntemi kullanarak katılımcı performanslarındaki cinsiyete dayalı farklılaşmayı incelemişlerdir. Sonuçta matematik testinde erkekler lehine bir farklılaşma tespit

edilirken çoktan seçmeli ve kompozisyon temelli yazma becerileri testinde ise performanstaki farklılaşmanın kadınlar lehine olduğunu tespit etmişlerdir. Diğer yandan okuma ve eleştirel düşünme testlerinde kadın ve erkeklerin performanslarında farklılaşma tespit edilememiştir.

Pang, Tian ve Boss (1994), çalışmalarında bir matematik testi üzerinden gerçek data kullanarak MH ve LR yöntemlerini örneklem büyüklüğü ve kriter değişkeni temelli kıyaslamışlardır. 1000, 500, 250, ve 100 ‘er kişilik gruplar ve toplam puan ile alt test toplam puanları olmak üzere iki farklı kriter değişken kullanılmıştır. Sonuçta her iki tekniğinde hem örneklem büyüklüğünden hem de kriter değişkenden etkilendiği sonucuna ulaşılmıştır. Yöntemler arasında her bir durum için belirlenen DMF değerinde yüksek bir uyum gözlenirken alt test puanlarının kriter değişken olarak kullanıldığı analizlerde DMF gösteren madde sayısında bir azalma tespit edilmiştir.

Öğretmen (1995), Türkiye’deki üniversite giriş sınavının birinci basamağının sözel yetenek bölümündeki maddelerin madde yanlılığı analizlerini MTK çerçevesinde dört farklı alan indeksi kullanarak, maddelerin konu alanları ve taksonomik seviyelerine göre cinsiyet değişkenini baz alarak karşılaştırmıştır. Ağırlıksız alan indeksleri Türkçe ve Sosyal Bilimler testlerindeki çoğu maddelerin farklı yetenek seviyelerinde cinsiyetlere göre farklı madde karakteristik eğrileri gösterdiğini ortaya çıkarmıştır. Türkçe testinde ağırlıklı ve ağırlıksız indekslere göre madde yanlılığı gözlenmesine rağmen Sosyal Bilimler testinde sadece ağırlıksız indekslere göre madde yanlılığı olduğu bulunmuştur. Sosyal bilimler alt testinde tarih ve coğrafya maddeleri erkek öğrenciler lehine çalışırken, psikoloji ve sosyoloji maddelerinin kız öğrenciler lehine çalıştığı tespit edilmiştir.

Hamilton ve Snow (1998), çalışmalarında 12 sınıf düzeyinde çoktan seçmeli ve açık uçlu cevaplandırmaya dayalı iki farklı fen bilgisi testi üzerinden cinsiyete dayalı DMF belirleme metotlarını araştırmaktadır. Her iki testte de uygulanan DMF istatistiği sonuçları test maddelerini tamamlayan öğrencilerle yapılan görüşmelerle de kanıtlanarak desteklenmiştir. Her iki test türünde de okul dışı bilgi ve deneyime dayalı ve görsellik içerdiği düşünülen maddelerin erkekler lehine farklı fonksiyonlaştığı tespit edilmiştir.

Gierl, Khaliq ve Boughton (1999), Matematik ve Fen bilgisi alanlarında uygulanan başarı testlerinde MH SIBTEST ve LR olmak üzere 3 yöntem kullanarak cinsiyete dayalı DMF analizi yapmışlardır. Elde edilen sonuçlara göre İncelenen 469 matematik maddesinden MH %7.3, SIBTEST %8.5 ve LR ise %10.5 TB, %0.2 ise TBO DIF belirlemiş, fen bilgisinde ise daha çok sayıda DMF gösteren madde tespit etmişlerdir. İncelenen 428 fen bil. Maddesinden MH %8.4, SIBTEST %14.7 ve LR de %17.3 TB, %0.5 TBO DMF içerdiği belirlenmiştir. Her iki alan da dikkate alındığında genel olarak maddelerin çoğunun B ve C düzeyinde DMF göstermediği ve Fen Bilgisi maddelerinin Matematik maddelerinden daha fazla DMF gösterme eğiliminde olduğu sonuçlarına ulaşılmıştır.

Henderson (2001), 1998 Ocak ve Haziran aylarında Alberta'da uygulanan dört farklı Alberta Eğitimi Diploma Sınavının İngilizce, sosyal bilimler, matematik ve biyoloji alt testlerindeki maddelerde cinsiyete dayalı DMF gözlenip gözlenmediğini araştırmıştır. Sonuçların alanyazındaki cinsiyete dayalı DMF ile ilgili bazı hipotezleri desteklerken bazılarını desteklemediği belirtilmiştir. Erkeklerin geometri ve matematiksel problem çözmeyi içeren sorularda kızlardan daha iyi performans göstermediği belirlenmiştir. Diğer yandan ikili puanlanan maddelerin çoğunluğu erkekler lehine DMF gösterirken çoklu puanlanan maddelerde tespit edilen DMF ise kadınlar lehine bulunmuştur. Bu bulgu sonucunda cinsiyet grupları ile madde formları arasında bir etkileşim olabileceği çıkarımına ulaşılmıştır.

Zenisky, Hambleton ve Robin (2003), yaptıkları çalışmada büyük ölçekli fen bilimleri sınavlarında kullanılan maddeleri 3 farklı eğitim seviyesi (ilkokul, ortaokul, lise) ve 2 farklı test biçimi (çoktan seçmeli ve açık cevaplı) olmak üzere toplamda 6 farklı sınıfta cinsiyet bakımından DMF gösterip göstermediğini incelemişlerdir. Sonuçlara göre çoktan seçmeli test maddeleri erkeklerin lehine çalışmaya eğilimli ve acık uçlu maddeler ise kadınlar lehine çalışmaya eğilimli bulunmuştur. Konu bazında yapılan incelemede ise dünya, uzay ve teknoloji ile ilgili maddeler ile görsel-uzamsal zeka gerektiren (grafikler, tablolar, diyagramlar, haritalar içeren) maddelerin erkek öğrenciler lehine çalıştığı belirlenmiştir.

Yurdugül (2003), çalışmasında 2001 yılı ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavında yer alan maddeleri testi alan öğrencilerin cinsiyet ve yerleşim yeri alt gruplarına göre DMF açısından MH ve LR tekniklerini kullanarak incelemiştir. Türkçe ve Sosyal bilimler maddelerinde DMF tespit edilememiş, Matematik maddelerinde ise cinsiyete ve yerleşim yerine göre DMF içeren birer adet maddeye rastlanmıştır. Fen bilgisi maddelerinde ise cinsiyete göre 2 adet maddede DMF tespit edilmiştir.

Hidalgo ve Lopez-Pina (2004), çalışmalarında madde parametreleri ve DMF büyüklüklerini manipüle ederek Lojistik Regresyon ve Mantel Haenszel tekniklerinin DMF belirlemedeki etkililiklerini, ve bu tekniklerden elde edilen etki büyüklüklerini kıyaslayan bir simülasyon çalışması yürütmüşlerdir. Yöntemler hem TB hem de TBO DMF için kıyaslanmıştır. Araştırma sonucunda DMF'nin TB olduğu durumda LR modelinin daha duyarlı olduğu ancak etki büyüklüğü temel alınarak yapılan karşılaştırmalarda MH yönteminin daha duyarlı sonuçlar ürettiği tespit edilmiştir.

Öğretmen ve Doğan (2004), Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı Matematik alt testindeki maddelerin madde yanlılığı analizlerini cinsiyet değişkenine göre Madde Tepki Kuramı çerçevesinde işaretli ve işaretsiz alan indekslerini kullanarak yapmışlardır. Araştırma sonucunda matematik alt testindeki maddelerin cinsiyete göre yanlılık gösterdiğini belirlemişlerdir. İşaretli alan indeksleri testteki çoğu maddenin kızların lehine yanlı davrandığını göstermiştir. Ayrıca, matematik konu alanlarına göre işaretli ve işaretsiz alan indekslerinin büyüklükleri arasında anlamlı bir fark bulunamamıştır.

Stoneberg (2004), 2003 Idaho Standart Başarı Testlerinde 4, 8 ve 10 sınıflarda matematik, dil sanatları ve okuma alanlarında cinsiyete ve etnik yapıya dayalı DIF analizi çalışmasında MH ve SIBTEST yöntemlerini kullanmış ve ortadan yüksek seviyeye kadar DMF gösteren maddelerin oranının %7 ile %37 oranında değiştiğini belirlemiştir.

Bekçi (2007), Milli Eğitim Bakanlığı'nın 2005 yılında yapmış olduğu orta öğretim kurumları öğrenci seçme ve yerleştirme sınavında yer alan maddelerin cinsiyete ve okul türüne göre değişen madde fonksiyonu içerip içermediğini MH ve LR tekniklerini

kullanarak incelemiştir. Çalışma sonucunda; Sosyal bilgiler alt testinde DMF olmadığı gözlenmiştir. Aynı şekilde fen bilgisi alt testinde okul türü alt grubuna göre, 1 maddede DMF bulunduğu, matematik alt testinde de 4 maddede DMF olduğu gözlenmiştir. Elde edilen DMF'lerin madde etkisinden mi, yoksa madde yanlılığından mı kaynaklandığı araştırılmış, konu alanı uzmanlarından görüşleri alınmıştır. Sonuçta OKS'de yer alan maddelerin yanlılık içermediği belirlenmiştir.

Doğan ve Öğretmen (2008), çalışmalarında değişen madde fonksiyonunu belirleme tekniklerinden ki-kare, Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırarak uygulamada ortaya çıkan benzerlik veya farklılıklarını belirlemişlerdir. Tekniklere göre DMF gösteren madde sayısının 0 ile 15 arasında değiştiği gözlemlenmiş, tekniklere göre elde edilen yanlı madde oranları arasında manidar bir fark bulunmuştur. En çok DMF'li madde sayısı MH tekniğinde elde edilirken en az DMF'li madde sayısı ise LR ve işaretli ki-kare tekniklerinden elde edilmiştir. Diğer taraftan teknikler arasında bakılan korelasyon değerleri dikkate alındığında tekniklerde elde edilen değerlerin birbirleriyle büyüklük sırası bakımından paralel olduğu belirlenmiştir. Özet olarak çalışma sonucunda MH, ki-kare ve lojistik regresyon tekniklerinin ki-kare değerlerinin büyüklüğü bakımından benzer; DMF verdiği belirlenen madde sayısı bakımından farklı sonuçlar ürettiği tespit edilmiştir.

Hauger ve Sireci (2008), 1999 TIMMS sınavına katılan öğrencilerden Singapur, İran ve Amerika Birleşik Devletleri örneklemelerinde farklı dil gruplarında DMF varlığını araştırmışlardır. Araştırmalarında LR tekniğini kullanılmış ve yalnızca bir madde de önemsiz düzeyde DMF tespit edilmiştir.

Bakan Kalaycıoğlu (2008), 2005 yılı Öğrenci Seçme Sınavında (ÖSS) yer alan maddelerin cinsiyet ve okul türü değişkenlerine göre farklı fonksiyonlaşıp fonksiyonlaşmadığı ve tespit edilen farklı fonksiyonlaşma durumunun yanlılıktan kaynaklanıp kaynaklanmadığını belirlemek üzere çalışmıştır. DMF analizi için Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) yöntemleri kullanılmıştır. Araştırma sonucunda Türkçe alt testini oluşturan maddelerin farklı fonksiyonlaşmadığı Sosyal bilimler alt testinde 7, matematik ve fen bilimleri alt testlerinde de 3 maddenin farklı

fonksiyonlaştığı ve bu maddelerden de 1 fen bilimleri maddesinin yanlılık gösterdiği belirlenmiştir.

Ateşok Deveci (2008), çalışmasında ÜDS-Sosyal Bilimler-İngilizce, 2005 mart dönemi, testi maddelerinin eğitim alanları ve cinsiyetine göre, yanlılık gösterip göstermediğini incelemiştir. Verilerin çözümlenmesinde Madde Tepki Kuramı'na dayalı, madde ayrışık fonksiyonu (MAF) belirleme yöntemleri olan, b parametre farkları, işaretli ve işaretsiz alan indeksleri yöntemleri kullanılmıştır. Elde edilen sonuçlara göre doğum oranları, ekonomi, para ve sanat olan maddeler, kız yanıtlayıcı grubuna, askerlik, döviz kuru, tarih ve ticaret konu içerikli maddeler ise erkek yanıtlayıcı grubuna avantaj sağladığı tespit edilmiş, Genel olarak ise ÜDS-SB İngilizce testinin, tüm katılımcılar bazında, kızlar lehine avantaj sağladığı sonucuna ulaşılmıştır.

Le (2009), PISA fen bilgisi testinde yer alan maddelerin ülkeler ve test dillerinin cinsiyete dayalı DMF ile arasındaki ilişkiyi araştırmıştır. DMF analizlerinde MTK yöntemleri kullanılmıştır. Test dil gruplarında ortalama, maddelerin %5.6'sı önemli ölçüde erkekler lehine, %2.8 i ise büyük ölçüde kadınlar lehine çalıştığı tespit edilmiştir. Genel olarak her bir test dili grubunda %2.8 ila %5.6 arasındaki madde kadınlar ya da erkekler lehine DMF gösteriyor olarak belirlenmiştir.

Liu ve Wilson (2009), PISA 2000 ve 2003 yıllarının matematik testindeki maddelerin Amerika örnekleminde madde tipi ve konu alanları sınıflarında cinsiyet bakımından farklılaşmasını araştırmışlardır. Sonuçta genel olarak hafif ama tutarlı bir erkek avantajının olduğu belirtilmiştir. Şekil ve uzay maddelerinde erkekler lehine oldukça büyük farklılık gözlemlenirken, daha önce yapılan birçok araştırmanın aksine çoktan seçmeli maddelerde ise kadın ve erkekler arasında bir farklılık belirlenememiştir.

Wiberg (2009), çalışmasında log. lineer model MH ve LR yöntemlerini kıyaslamıştır. Araştırma sonucunda MH yönteminin diğer yöntemlerden daha kapalı sonuçlar ürettiği gözlemlenmiş, Ayrıca üç model arasında elde edilen korelasyon değerlerinin de yüksek olduğu belirtilmiştir. LR ve MH yöntemleri arasında elde edilen korelasyon değeri 0.749 olarak bulunmuştur.

Yıldırım ve Berberoğlu (2009), çalışmalarında 2003 PISA verisinin Türkiye ve Amerika örneklemi üzerinde kültürler arası bir DMF analizi yürütmüşlerdir. Araştırmada sınırlandırılmış faktör analizi, MH, ve MTK ‘ya dayalı olabilirlik oran analizi yöntemleri kullanılmıştır. Yapılan değerlendirmeler sonucunda, maddelerin ölçtüğü bilişsel yetenekler, nicel kelime kullanımları ve çeviri hataları PISA 2003’te DMF ‘ye sebep olabilecek önemli hata kaynakları olarak belirlenmiştir.

Gök, Kelecioğlu ve Doğan (2010), çalışmalarında değişen madde fonksiyonunu belirlemede kullanılan Mantel- Haenszel (MH) ve lojistik regresyon (LR) tekniklerini karşılaştırmaktırlar. Araştırmada, 2005 yılında yapılan Ortaöğretim Kurumları Sınavı (OKS)’nın matematik ve fen bilgisi alt testleri kullanılmıştır. Test maddelerinin cinsiyet ve okul türüne göre fonksiyonlarının farklılaşıp farklılaşmadığı incelenmiştir. Araştırma sonucunda cinsiyet ve okul türü değişkenine göre matematik ve fen bilgisi alt testindeki maddelerin tümünde tespit edilen DMF büyüklüğü ihmal edilebilir düzeydedir. Ancak bu iki teknik arasında düşük düzeyde DMF bulmaları haricinde düşük düzeyde bir benzerlik tespit edilmiştir. Diğer yandan genel olarak MH tekniğine göre elde edilen DMF’li madde sayısının, LR tekniğinden daha fazla olduğu görülmüştür. Bu sonuca göre MH tekniğinin LR tekniğinden daha duyarlı olduğu sonucuna ulaşılmıştır.

Bakan Kalaycıoğlu ve Berberoğlu (2010), çalışmalarında Türkiye’deki Üniversite Giriş Sınavının sayısal bölümündeki maddelerinin cinsiyet grupları arasında DMF gösterip göstermediğini belirlemişler, DMF ‘nin olası kaynaklarını belirleyebilmek için madde içerik analizi yapmışlardır. MH, LR, sınırlandırılmış faktör analizi ve MTK ya dayalı olasılık oranı yöntemi olmak üzere dört ayrı yöntemle inceleme yapılmış araştırma sonucunda; yüksek seviyeden bilişsel yetenekler ve şekil ya da grafik içeren maddeler erkek öğrenciler lehine DMF kaynağı gösterirken, rutin algoritmik hesaplamaların olduğu maddelerin ise erkeklere karşı DMF ürettiği görülmüştür. Buna rağmen, DMF gösteren maddelerin ÖSS sınavının sayısal bölümünün ölçüt dayalı geçerliğine bir tehdit oluşturmadığı belirlenmiştir.

Adedoyin (2010), çalışmasında 2004 Botswana ortaokul sınavının Matematik bölümündeki maddelerin cinsiyete göre yanlılık içerip içermediği araştırmıştır. 3 parametrelili MTK İstatistiği kullanılmıştır. Toplamda 38 maddeden oluşan testte 16

madde bu y nteme uyum saėlamıř ve 5 maddede cinsiyet bakımından yanlılık tespit edilmiřtir.

 epni (2011),  alıřmasında Akademik Personel ve Lisans st  Eėitimi Giriř Sınavı Sayısal Yetenek testlerinde cinsiyete ve adayların mezun olduėu/olacalı b l me g re deėiřen madde fonksiyonu ve deėiřen madde grubu fonksiyonu i erip i ermediėi belirlenmiřtir. Sayısal 1 Testi'nde    maddenin erkek  ėrenciler lehine, d rt maddenin ise kız  ėrenciler lehine iřlediėi, Sayısal 2 Testi'nde ise bir maddenin erkek  ėrenciler lehine,    maddenin ise kız  ėrenciler lehine iřlediėi tespit edilmiřtir. DMF g steren maddelerin farklı gruplar i in farklı iřlemelerinin nedenleri uzman kanıları yardımıyla incelenmiřtir. Genel olarak, rutin algoritmik iřlemlerle c z len, cebirsel ifadelerle soyut olarak verilen maddeler kız  ėrenciler lehine islerken, ger ek hayat durumlarının verildiėi kelime problemlerinin erkek  ėrenciler lehine iřlediėi tespit edilmiřtir.

Alanyazın incelendiėinde  zellikle fen ve matematik gibi sayısal alanlarda maddelerin ilgili olduėu konular doėrultusunda cinsiyet deėiřkenine g re DMF'ye rastlandıėı g r lm řtir. Buna baėlı olarak erkek ve kızların farklı ilgi alanları ve yeteneklerinin olabileceėi ve bunun DMF g zlenmesine sebep olabileceėi d ř n lm řtir. Ayrıca Gierl, Khaliq ve Boughton (1999)  alıřmalarında fen maddelerinin matematik maddelerine g re DMF g stermeye daha eėilimli olduėu bulgusuna ulařmıřlardır. Bu noktalar deėerlendirildiėinde arařtırma kapsamında PISA Fen okuryazarlıėı alt testindeki maddelerde DMF analizi yapılmasına karar verilmiřtir.

II. BÖLÜM

YÖNTEM

Bu bölümde araştırmanın türü, çalışma grubu, verilerin elde edilmesi, verilerin analizi ve verilerin analizi sırasında kullanılan bilgisayar programları üzerinde durulmuştur.

2.1 ARAŞTIRMANIN TÜRÜ

Araştırma, PISA 2009 sınavının fen okuryazarlığı alt testinde yer alan maddelerinin cinsiyete göre değişen madde fonksiyonu içerip içermediğini incelemeye yöneliktir. Var olan durumu tam ve dikkatli bir biçimde ortaya çıkarmak ve ayrıntısıyla incelemek amacı ile tasarlanması nedeninden ötürü betimsel çalışma özelliği göstermekte, ayrıca araştırma kapsamında kullanılan yöntemlerin benzerlik ve farklılıkları yönünden bilgi vermesi sebebiyle kuramsal bir çalışma özelliği de göstermektedir.

2.2 ÇALIŞMA GRUBU

PISA öğrenci evreni, okul türüne bakılmaksızın okullarda öğrenim gören, değerlendirmenin yapılacağı tarih itibarıyla yaşları 15 yıl 3 ay ve 16 yıl 2 ay arasında değişen, en az altı yıllık örgün eğitimi tamamlamış öğrencilerden oluşmaktadır.

PISA 2009'a, 65 katılımcı ülkeden, 15 yaş grubu yaklaşık 26 milyon öğrenciyi temsilen, 475.460 öğrenci katılmıştır. Türkiye örnekleminde ise 12 istatistiki bölge biriminden 56 il ve okul türlerine göre tabakalandırılarak PISA uluslararası merkez tarafından seçkisiz yöntemle belirlenen toplam 170 okuldan 4996 öğrenci yer almaktadır.

PISA uygulamasında öncelikle, ulusal merkezler tarafından tabakalı örnekleme yöntemine göre örneklemin çatısı oluşturulur. Bu çatıda, öğrencilerin ve okulların hangi kriterlere göre örnekleme dâhil edileceği (bölge, program türü, okul türü vb.) belirlenir. Oluşturulan örneklem dosyasında, belirlenen tabakalara göre okul ve öğrenci sayıları yer almaktadır. Daha sonra ulusal merkezler bu dosyayı uluslararası kuruluşa gönderirler. Belirlenen tabakalara ve öğrencilerin temsil edilme oranlarına göre okullar belirlenir ve ülkelere gönderilir. Ülkeler belirlenen okullardaki 7. sınıf ve üzeri sınıflarda okuyan tüm 15 yaş grubu öğrencilerin arasından seçkisiz yöntemle 35 öğrenciyi PISA'ya katılmak üzere seçer. 35'ten daha az sayıda öğrencisi olan

okullardaki tüm 15 yaş grubu öğrenciler seçilir. Bu liste kontrol edilmek üzere okullara gönderilir ve okullardan onay alındıktan sonra PISA uygulamasına katılacak okul ve öğrenciler kesinleşmiş olur.

PISA 2009’a katılan okul türleri arasında; ilköğretim okulu, genel lise, Anadolu lisesi, fen lisesi, Anadolu öğretmen lisesi, Anadolu güzel sanatlar lisesi, meslek lisesi, Anadolu meslek lisesi, teknik lise, Anadolu teknik lisesi ve çok programlı liseler yer almaktadır. Bölgeleri temsil eden okul sayıları, her bölgede bulunan okul sayıları ile orantılı olarak örnekleme yer almıştır. Türkiye’deki özel okulların oranına paralel olarak PISA 2009 örnekleminin %2’si özel okul %98’i de resmî okullardan oluşmaktadır.

Bu araştırmanın yürütüldüğü çalışma grubunu ise PISA 2009 uygulamasında 12. kitapçığı alan grup oluşturmaktadır. Çalışma grubu 338 kişiden oluşmaktadır ve analizin yürütüleceği cinsiyet değişkenine göre kişi sayısı Tablo 8’de sunulmuştur.

Tablo 8:Çalışma Grubuna Ait Frekans ve Yüzde Değerleri

Grup=Cinsiyet	Frekans (f)	Yüzde (%)
Kız	182	54,8
Erkek	156	45,2
Toplam	338	100

Tablo 8’de görüldüğü gibi toplamda 338 kişiden oluşan çalışma grubunun %54,8’i kızlardan, %45,2’si ise erkeklerden oluşmaktadır. Cinsiyete göre DMF analizi yapılırken çalışma grubu kızlar ve erkekler olmak üzere iki veri setine ayrılmış, ve analizler bu iki veri setinin 12. Kitapçık fen okuryazarlığı alt testi maddelerine verdikleri cevaplar dikkate alınarak yürütülmüştür.

2.3. PISA’DA KULLANILAN VERİ TOPLAMA ARAÇLARI

PISA 2009 uygulamasında temel olarak 13 adet kitapçık, 1 adet öğrenci anketi ve 1 adet okul anketi kullanılmıştır. Bunların yanı sıra, bazı ülkelerde isteğe bağlı anketler de kullanılmaktadır. 35 kişilik her bir grupta, 3 öğrenciden fazlasına aynı kitapçık verilmeyecek şekilde, her öğrencinin yanıtlayacağı kitapçık türü, bilgisayar ortamında seçkisiz yöntemle belirlenmektedir.

PISA 2009’da kullanılan üniteler, 13 madde demeti halinde gruplandırılmıştır. Her bir madde demetinin yanıtlanma süresi 30 dakikadır. Uygulamaya katılan her ülkede 7 okuma becerileri, 3 matematik ve 3 de fen madde demeti bulunmaktadır. Bu madde demetleri belirli bir döngü düzenine göre 13 kitapçığa yerleştirilmiştir. Bu kapsamda kitapçıklardaki madde sayıları ve uygulandıkları kişi sayıları farklılık göstermektedir. 13 adet kitapçığın her birinin Türkiye örnekleminde uygulandığı kişi sayısı yaklaşık 300-400 kişi arasında seyir edip birbirine yakın değerler göstermektedir. Her bir kitapçığın uygulandığı kişi sayısı ve bu kitapçıklarda bulunan toplam fen maddesi sayısı tablo)’da sunulmuştur.

Tablo 9:Kitapçıklara Ait Kişi ve Fen Okuryazarlığı Madde Sayısı

Kitapçık No	Kişi Sayısı			Madde Sayısı
	Kız	Erkek	Toplam	
1	202	186	388	0
2	175	211	386	17
3	179	199	378	35
4	198	184	382	18
5	182	191	373	0
6	188	197	385	0
7	175	205	380	18
8	196	196	392	17
9	188	202	390	20
10	183	202	385	35
11	183	206	389	0
12	209	174	383	36
13	187	198	385	18

Tablo 9’da sunulan kitapçıklara ait kişi ve madde sayıları incelendiğinde kitapçıklar arasında kişi sayısı bakımından büyük bir farklılık olmadığını ancak kitapçıklarda bulunan fen madde sayısı bakımından önemli farklılıklar bulunduğu (örneğin bazı kitapçıklarda hiç fen maddesi bulunmamaktadır) tespit edilmiştir. Bu kapsamda, araştırmada kullanılacak kitapçığın seçiminde kitapçıklarda bulunan fen maddesi sayısına özellikle önem verilmiş ve Tablo 9’da sunulan bilgiler ışığında 36 adet fen maddesinin bulunduğu ve 383 kişiye uygulanan 12. kitapçığın uygun bir seçim olabileceği düşünülmüştür.

PISA ’da farklı madde tipleri bir arada kullanılmaktadır. Bu maddelerin yaklaşık yarısı öğrencilerin 4 ya da 5 seçenekten birini işaretledikleri çoktan seçmeli ya da öğrencilerin

bir dizi önermeyi “evet/hayır”, veya “katılıyorum/katılmıyorum” gibi muhtemel iki yanıtta birini seçerek değerlendirdikleri maddelerden (*karmaşık* çoktan seçmeli) oluşmaktadır. Kalan maddeler ise, öğrencilerin kısa ya da uzun, kendi yanıtlarını oluşturmalarının istendiği açık uçlu maddelerdir. Bu araştırma kapsamında ele alınan kitapçıktaki 2 adet açık uçlu fen okuryazarlığı maddesi 1-0 şeklinde puanlanmadığı için analizden çıkarılmıştır.

2.4 VERİLERİN ÇÖZÜMLENMESİ

Verilerin çözümlenmesi aşağıdaki basamaklarla gerçekleşmiştir;

OECD PISA internet sitesinden elde edilen PISA 2009 uygulaması veri setinden Türkiye örnekleme seçilmiştir. Örnekleme uygulanan 13 adet kitapçık, her bir kitapçığın uygulandığı kişi sayısı ve bu kitapçıklarda bulunan maddeler tespit edilmiştir. Ardından uygulanan kişi ve madde sayısı değişkenleri göz önünde bulundurularak analizin yürütüleceği kitapçık seçilmiştir. Veri seti düzenlenirken 1-0 haricinde puanlanan tüm durumlar (2,7,8,9) analizden çıkarılmıştır.

Alt problemlere ilişkin istatistiksel çalışmalara başlamadan önce grubun dağılımının genel özelliklerinin belirlenebilmesi amacı ile grubun tamamına ve cinsiyet değişkenine göre elde edilen alt gruplara betimsel istatistik teknikleri uygulanmış, grubun tamamının ve alt grup dağılımlarının genel özellikleri tanımlanmıştır. Betimsel istatistiklerin elde edilmesinde SPSS paket programından faydalanılmıştır.

Birinci alt probleme ait bulguların elde edilmesi aşamasında, maddelerin cinsiyet alt değişkenine göre oluşturulan gruplar arasında DMF içerip içermediği MH yöntemi ile belirlenmeye çalışılmış, düzenlenen veri cinsiyet değişkenine göre iki gruba ayrılmıştır. Ardından MH yöntemine dayalı DMF analizi gerçekleştirilmiş, DMF gösteren maddeler DMF düzeyleri ve hangi grubun lehine işledikleri tespit edilmiş ve sonuçlar tablolastırılmıştır. MH yöntemi istatistiklerini elde etmek için EZDIFF (Waller, 1998) paket programından faydalanılmıştır.

İkinci alt probleme ait analizlerde, maddelerin cinsiyet alt değişkenine göre oluşturulan gruplar arasında DMF içerip içermediğinin LR yöntemi ile belirlenme aşamasında iki aşamalı bir analiz süreci uygulanmıştır. Öncelikle MH yönteminde kullanılmak üzere

hazırlanan iki alt gruba göre ayrılmış data seti benzer şekilde analiz edilmiş ve LR yöntemine ait β ve p değerleri elde edilmiştir. Hem grup hem de etkileşim değişkenlerinden elde edilen β değerlerine ait p değerleri manidarlık düzeyleri ile kıyaslanarak maddelerin DMF gösterip göstermedikleri ve DMF gösteriyorsa bu DMF'nin biçimi (TB ya da TBO) belirlenmiştir. Analizin ikinci aşamasında DMF gösterdiği belirlenen maddelerdeki DMF'lerin etki büyüklüklerini ve önemlilik düzeylerini belirleyebilmek için Zumbo (1999) tarafından yazılan özel bir scriptten faydalanılmıştır. Bu script sayesinde hem TB hem de TBO DMF'ler için ΔR^2 değerleri hesaplanmış ve Jodoin ve Gierl (2001) tarafından belirlenmiş kesme puanları kullanılarak DMF'lerin hangi düzeyde (A,B ya da C) olduğu belirlenmiştir. Elde edilen sonuçlara göre DMF gösteren maddeler, bu maddelerin DMF biçimleri, avantajlı gruplar ve ΔR^2 değerleri ile bu değerlerin düzeyleri tespit edilmiş ve tablolar halinde düzenlenmiştir.

LR analizlerinde hem EZDIFF paket programından hem de SPSS paket programından faydalanılmıştır. LR 'a ait β ve p değerleri için EZDIFF paket programı kullanılırken, Zumbo (1999) tarafından yazılan scriptin kullanılması ve ΔR^2 değerlerinin hesaplanması için de SPSS paket programı kullanılmıştır. SPSS paket programında kullanılan script Ek 1 'de sunulmuştur.

Üçüncü alt probleme ait analizlerde, kullanılan iki yöntemin gösterdiği uyum derecesini belirleyebilmek için her iki yöntemde de elde edilen DMF gösteren maddeler, tespit edilen DMF düzeyleri ve toplam DMF gösteren madde sayıları kıyaslanmıştır. DMF büyüklükleri haricinde, yöntemlerin maddelerdeki DMF büyüklük sıralamaları bazında uyum gösterip göstermediklerini incelemek içinse söz konusu iki yöntemden elde edilen istatistiklerin Spearman-Brown sıralama farkı korelasyon katsayısı değerlerine bakılmıştır. Korelasyon değerlerini elde etmek için de SPSS paket programı kullanılmıştır.

Çalışma Grubuna Ait Genel Özellikler

Fen okuryazarlığı alt testinde yer alan maddelerin DMF analiz sonuçlarına geçmeden önce, çalışma grubu ve çalışma grubunun cinsiyet değişkenine dayalı alt gruplarına ilişkin betimsel istatistik sonuçlarına yer verilerek, grup dağılımlarının genel özellikleri tanımlanmıştır.

PISA 2009 Türkiye örnekleme fen okuryazarlığı alt testine ait betimsel istatistikler Tablo 10’da sunulmuştur.

Tablo 10: Çalışma Grubuna Ait Betimsel İstatistikler

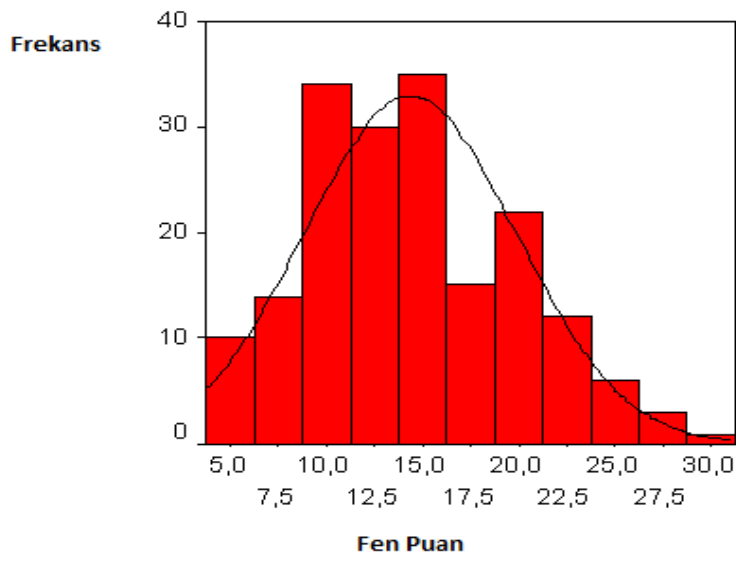
Madde Sayısı	34
Öğrenci Sayısı	338
Aritmetik Ortalama	13,74
Standart Sapma	5,96
Çarpıklık Katsayısı	0,423
Basıklık Katsayısı	-0,420
Testin Güvenirlik Katsayısı	0,82
Ortalama Güçlük	0,40

Tablo 10 incelendiğinde ele alınan testin puan dağılımının normal dağılıma göre biraz daha sağa çarpık ve basık bir dağılım olduğu görülmektedir. Ancak farkın küçüklüğünü dikkate alarak normal dağılıma yakın bir dağılımdır denmesi mümkündür. Öğrenciler 34 adet maddenin ortalama 13.74’üne cevap verebildiklerinden ortalama güçlük değeri 0,40 hesaplanmıştır. Yani test orta güçlük seviyesindedir. Ayrıca testin güvenirlilik katsayısı 0,82 bulunmuştur.

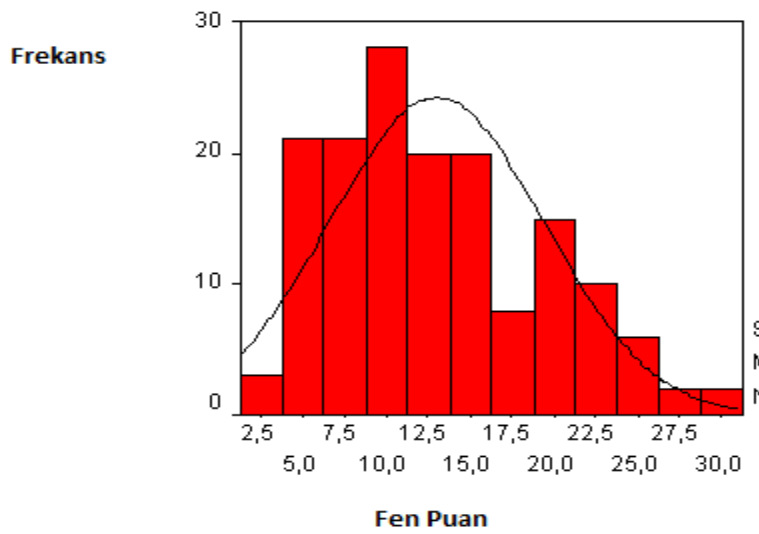
Araştırmada gruplama değişkeni olarak cinsiyet değişkeni kullanılması sebebi ile verinin kız ve erkek alt gruplarına göre ayrı ayrı betimsel istatistik değerleri hesaplanmış ve tablo 11’de sunulmuş, ayrıca iki alt gruba ait dağılımların grafikleri çizilmiş Şekil 2 ve Şekil 3’te sunulmuştur.

Tablo 11: Cinsiyete Göre Betimsel İstatistikler

	KIZ	ERKEK
Öğrenci Sayısı	182	156
Ortalama	14,36	13,01
Standart Sapma	5,49	6,41
Çarpıklık	0,40	0,54
Basıklık	-0,36	-0,42
Ortalama güçlük	0,42	0,38



Şekil2. Kızların Dağılım Grafiği



Şekil 3. Erkeklerin Dağılım Grafiği

Tablo 11, Şekli 2 ve Şekil 3 birlikte incelendiğinde, genel olarak cinsiyet alt gruplarının birbirine benzer özellikler gösterdiğini ve grup ortalamaları ve homojenlikleriyle ilgili benzer istatistiksel değerlere ulaşıldığını söylemek mümkündür. Alt grup ortalamalarına ve ortalama güçlük değerlerine birlikte bakıldığında değerlerin birbirine çok yakın olduğu görülmektedir. Çarpıklık ve basıklık değerleri incelendiğinde normal dağılıma göre küçük sapmalar gösterdiği ve her iki dağılımında normal dağılıma yakın dağılımlar gösterdiği görülmektedir. Ortalama güçlük değerleri kıyaslandığında testin erkek gruba biraz daha güç geldiğini ancak genel olarak birbirine yakın ve orta güçlükte değerler hesaplandığı görülmektedir.

BÖLÜM III

BULGULAR VE YORUMLAR

Bu bölümde PISA 2009 Türkiye örneklemini fen okuryazarlığı alt testini oluşturan maddelerin cinsiyete göre farklı fonksiyonlaşıp fonksiyonlaşmadığı MH ve LR yöntemleri ile incelenmiş, ardından bu iki yöntemin uyumluluk derecesi hem nicel hem de nitel verilerle belirlenmeye çalışılmıştır. Cinsiyet değişkenine göre yapılan analizlerde elde edilen MH ve LR sonuçları Ek 2 ve Ek 3'te sunulmuştur.

3.1 BİRİNCİ ALT PROBLEME İLİŞKİN BULGULAR VE YORUMLAR

Alt Problem 1, PISA 2009 yılı uygulamasında fen okuryazarlığı alt testindeki maddeler cinsiyete göre Mantel Haenszel yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

PISA 2009 fen okuryazarlığı testi kapsamında ele alınan maddelerin cinsiyete göre DMF içerip içermediğini belirlemek üzere MH tekniği kullanılmış ve analiz sonuçları Tablo 12'de sunulmuştur.

Tablo 12: Fen Okuryazarlığı Alt Testine Ait Cinsiyet Değişkenine Göre Maentel-Haenszel Analiz Sonuçları

Madde No	α	χ^2	p	Δ -MH	SE
1	1.341	0.786	0.375	-0.689	0.669
2	1.717	3.182	0.074	-1.271	0.665
3	0.936	0.002	0.969	0.156	0.773
4	1.056	0.001	0.976	-0.128	0.848
5	0.593	3.748	0.050*	1.227	0.603
6	0.505	6.183	0.013*	1.604	0.619
7	0.727	1.216	0.270	0.748	0.609
8	0.562	2.386	0.122	1.353	0.793
9	1.646	4.081	0.043*	-1.171	0.551
10	1.045	0.000	0.990	-0.104	0.668
11	0.819	0.456	0.500	0.469	0.588
12	1.142	0.163	0.686	-0.313	0.591
13	0.886	0.135	0.713	0.283	0.576
14	1.176	0.247	0.619	-0.381	0.608
15	0.948	0.001	0.973	0.126	0.693

Tablo 12'nin Devamı

Madde No	α	χ^2	p	Δ -MH	SE
16	1.416	1.564	0.211	-0.817	0.592
17	1.507	1.863	0.172	-0.964	0.649
18	1.011	0.015	0.901	-0.026	0.747
19	0.745	1.046	0.306	0.693	0.600
20	0.865	0.240	0.624	0.341	0.556
21	1.436	1.862	0.172	-0.850	0.569
22	0.673	2.186	0.139	0.931	0.582
23	0.734	1.199	0.273	0.725	0.591
24	0.974	0.000	0.983	0.062	0.591
25	0.635	2.863	0.091	1.068	0.588
26	0.900	0.100	0.752	0.249	0.571
27	0.902	0.102	0.750	0.243	0.555
28	0.876	0.173	0.677	0.310	0.573
29	1.222	0.217	0.641	-0.472	0.755
30	1.123	0.134	0.714	-0.274	0.563
31	0.556	3.752	0.050*	1.377	0.669
32	0.519	0.770	0.380	1.540	1.354
33	1.136	0.192	0.661	-0.300	0.541
34	1.348	0.972	0.324	-0.702	0.625

* $p \leq 0.05$ ** $p \leq 0.01$

Referans grup: erkekler (n=156) Odak grup: kızlar (n=182)

Tablo 12'de sunulan maddelerden sonuçları 0.05 düzeyinde manidar çıkan maddelere ait Δ MH değerleri değerlendirilmiş, MH yöntemi için kullanılan Δ MH eşik değerleriyle kıyaslanarak maddelerin hangi düzeyde DMF gösterdiği tespit edilmiş ve Δ MH değerinin negatif ya da pozitif olmasına göre de DMF gösteren maddelerin hangi gruba avantajlı davrandığı belirlenmiştir. Yapılan değerlendirme sonuçları Tablo 13'de sunulmuştur.

Tablo 13: Cinsiyet Değişkenine Göre Maentel-Haenszel Analiz Sonuçları

Avantajlı Grup	DMF Düzeyi		
	A	B	C
Kızlar		31, 5, 6	
Erkekler		9	

Tablo 12 ve Tablo 13 de yer alan bilgiler ışığında fen okuryazarlığı testinde bulunan ve analize dahil edilen 34 adet maddenin 3 tanesi kızlar lehine, 1 tanesi ise erkekler lehine B düzeyinde DMF göstermişlerdir.

Zenisky, Hambleton ve Robin (2003) geniş ölçekli bir sınavda fen bölümü için yaptıkları DMF analizi sonucunda çoktan seçmeli maddelerin, dünya uzay teknoloji konularıyla ilgili ya da görsel uzamsal zeka gerektiren maddelerin erkekler lehine çalıştığı belirlenmiştir. Bunu destekler nitelikte Liu ve Wilson (2009), PISA 2000 ve PISA 2003 Fen maddeleri üzerinden yaptıkları çalışmada da şekil uzay içeren maddelerde erkekler lehine DMF tespit etmişlerdir. Le (2009), bu araştırmaya benzer olarak, PISA 2000 verileri üzerinden yürüttüğü DMF analizlerinde hem kadın hem de erkekler lehine önemli düzeyde işleyen maddeler tespit etmiştir.

3.2 İKİNCİ ALT PROBLEME İLİŞKİN BULGULAR VE YORUMLAR

Alt Problem 2 PISA 2009 yılı uygulamasında fen okuryazarlığı alt testindeki maddeler cinsiyete göre Lojistik Regresyon yöntemiyle yapılan analizlerde değişen madde fonksiyonu göstermekte midir?

PISA 2009 fen okuryazarlığı testi kapsamında ele alınan maddelerin cinsiyete göre DMF gösterip göstermediğini belirlemek üzere lojistik regresyon tekniği kullanılmış ve analiz sonuçlarından elde edilen grup ve etkileşim (grup X yetenek) değişkenlerine ait β değerleri ile β değerine ait anlamlılık düzeyleri Tablo 14'de sunulmuştur.

Tablo 14: Fen Okuryazarlığı Alt Testine yönelik Cinsiyet Değişkenine Göre Lojistik Regresyon Analiz Sonuçları

Madde No	Grup		Grup X Yetenek (Etkileşim)		DMF Biçimi
	β	P	B	P	
1	0.974024	0.170333	-0.056650	0.323399	
2	-0.631624	0.519372	0.092530	0.212152	
3	-1.199667	0.275968	0.067275	0.259555	
4	-0.275882	0.799819	0.023464	0.691513	
5	-1.270309	0.072662	0.062829	0.254493	
6	-1.429648	0.092911	0.057245	0.341558	
7	-0.441000	0.587780	0.013334	0.810853	
8	-1.166868	0.414437	0.037361	0.632881	
9	0.213443	0.719901	0.016412	0.690030	

Tablo 14'ün Devamı

Madde No	Grup		Grup X Yetenek (Etkileşim)		DMF Biçimi
	β	P	β	P	
10	1.837956	0.035540*	-0.121701	0.017853*	TBO
11	-1.045046	0.134375	0.057826	0.188124	
12	-0.320679	0.616306	0.040029	0.328999	
13	-0.525097	0.438309	0.023457	0.591139	
14	0.687027	0.319081	-0.045003	0.406813	
15	-1.077247	0.328641	0.065591	0.328492	
16	1.162221	0.132406	0.050172	0.337695	
17	1.504444	0.091417	-0.063749	0.246561	
18	1.418189	0.182607	-0.090226	0.127596	
19	-0.111354	0.868816	-0.011373	0.828132	
20	-0.075930	0.902282	-0.007364	0.862409	TBO
21	0.377140	0.567445	0.000724	0.986691	
22	-2.108592	0.002857**	0.139238	0.008823**	
23	-0.865721	0.204671	0.043807	0.312500	
24	0.310519	0.683137	-0.016059	0.745040	
25	-0.488937	0.481519	0.009028	0.845461	TBO
26	0.179548	0.777561	-0.011487	0.776924	
27	1.031520	0.109949	-0.085194	0.047462*	
28	-0.487194	0.419753	0.029209	0.456850	
29	-0.426022	0.672429	0.040334	0.471870	
30	-0.827169	0.205827	0.069186	0.109860	TBO
31	0.359059	0.676558	-0.080393	0.300741	
32	-5.730612	0.044468*	0.264560	0.038824*	
33	0.694548	0.244817	-0.045993	0.254585	
34	1.112625	0.161665	-0.053698	0.279068	

* $p \leq 0.05$ ** $p \leq 0.01$

Tablo 14'te sunulan β anlamlılık düzeyleri 0.05 anlamlılık düzeyi ile kıyaslanmış ve manidar bulunma durumlarına göre DMF gösterip göstermedikleri ve bu DMF'nin biçimi belirlenmiştir. Buna göre 4 adet maddede hepsi TBO olmak üzere DMF tespit edilmiştir.

LR analizinin üç model kurulması ve bu modellerin analize hiyerarşik biçimde dahil edilmesi şeklinde gerçekleştiği bilinmektedir. Söz konusu her adımda β değerleri yanında R^2 değerleri de hesaplanmaktadır. Zumbo (1999), etki büyüklüğünün TB DMF'de İkinci modelde elde edilen R^2 değeri ile birinci modelde elde edilen R^2 değeri farkından, TBO DMF 'de ise üçüncü modelde elde edilen R^2 ile birinci modelden elde edilen R^2 değerleri farkından elde edileceğini belirtmiştir.

Tablo 3.3 deki bilgiler kullanılarak, β değerleri manidar çıkan maddeler için belirlenen modellerde elde edilen R^2 değerleri farkından ΔR^2 değerleri elde edilmiş, ve Jodoin ve Gierl (2001) tarafından belirlenen sınır değerler kullanılarak maddelerin gösterdiği DMF'lerin düzeyleri belirlenmiş ve Tablo 15'te sunulmuştur.

Tablo 15: Cinsiyet Değişkenine Göre Etki Büyüklüğü Değerleri

Madde No	(md2-md1) ΔR^2	(md3-md2) ΔR^2	DMF Düzeyi
10	0.023	0.001	A
22	0.031	0.007	A
27	0.015	0.001	A
32	0.037	0.000	B

Tablo 15'teelde edilen bulgulara göre testte incelemeye alınan maddelerin 3 tanesinin A düzeyinde 1 tanesinin ise B düzeyinde DMF gösterdiği görülmüştür. Yani DMF gösterdiği belirlenen 3 madde ihmal edilebilir düzeydeyken bir madde de orta düzeyde DMF tespit edilmiştir.

Bu araştırmaya benzer olarak Liu ve Wilson (2009), PISA 2000 ve PISA 2003 verisi üzerinden yürüttükleri çalışmalarında da DMF tespit edilen maddelerdeki etki büyüklükleri genel olarak küçük ve önemsiz bulunmuştur. Diğer yandan Yıldırım ve Berberoğlu (2009) 'nın PISA 2003 sınavı üzerinden yürüttükleri DMF analizlerinde ise önemli düzeyde DMF gösteren maddeye rastlanmış ve bunun olası sebepleri olarak maddelerin ölçtüğü bilişsel yetenekler, nicel kelime kullanımları ve çeviri hataları sayılmıştır.

3.3 ÜÇÜNCÜ ALT PROBLEME İLİŞKİN BULGULAR VE YORUMLAR

Alt Problem 3 Lojistik Regresyon ve Mantel Haenszel yöntemleriyle yapılan analizlerde DMF ile ilgili elde edilen sonuçlar uyum göstermekte midir?

PISA 2009 uygulaması, Türkiye örnekleme, fen okuryazarlığı alt testine ait maddelerde cinsiyet değişkenine ait DMF analizleri LR ve MH teknikleri kullanarak elde edilmiştir. Her iki yöntemden de elde edilen nihai sonuçlar Tablo 16'da sunulmaktadır.

Tablo 16 : LR ve MH yöntemleriyle elde edilen DMF sonuçları

	A	B	C	Toplam
MH		31,5,6,9		4
LR	10,22,27	32		4

Tablo 16’da yer alan bilgiler incelendiğinde MH yöntemi ile de LR yöntemi ile de 4 adet DMF gösteren madde tespit edilmiştir. MH yöntemi bu dört adet maddenin dördünü de orta düzeyde belirlerken LR ile belirlenen dört maddeden 3’ü A bir tanesi de B düzeyindedir. DMF gösterdiği tespit edilen maddelere ayrıntılı olarak bakıldığında iki yöntemle tespit edilen hiç ortak madde olmadığı görülmüş ve bu iki yöntemin belirledikleri DMF düzeyi bakımından düşük düzeyde uyum gösterirken DMF gösterdiği belirlenen maddeler bakımından uyumsuz olduğu yorumuna gidilmiştir. Elde edilen bu bulgu Pang, Tian ve Boss (1994) çalışmaları ile çelişirken, Gök ve ark. (2010) ve Doğan ve Öğretmen (2008) çalışmaları ile belirlenen maddelerdeki düşük uyum noktasında tutarlı bulunurken, yöntemlerle belirlenen DMF düzeylerinde uyum görülmesi noktasında çeliştiği tespit edilmiştir.

Diğer yandan LR ile tespit edilen maddelerin hepsinin TBO DMF gösterdiği ve MH’nin 4 adet orta düzeyde DMF gösteren madde belirlemesine rağmen LR’nin hiç TB DMF gösterememesinden TB DMF belirleme de MH daha duyarlı sonuçlar üretmektedir yorumuna gidilmiştir. Elde edilen bu bulgu literatürdeki kuramsal çalışmaların birçoğu (Gierl, Khaliq ve Boughton (1999), Hidalgo ve Lopez-Pina (2004), Doğan ve Öğretmen (2008)) ile desteklenmekte diğer yandan log lineer model ile MH ve LR tekniklerini kıyaslayan Wiberg (2009) çalışmasında elde edilen MH yönteminin daha kapalı sonuçlar üretmesi sonucu ile ayrıca TB DMF’yi belirlemede LR’nin de MH kadar güçlü sonuçlar verdiği bulgusuna ulaşan Swaminathan ve Rogers (1990), ve Gök ve ark. (2010) çalışmaları ile çelişmektedir. Diğer yandan MH yöntemi ile TBO DMF belirlenememesi bilgisi ışığında LR ile belirlenen 4 adet orta düzeyde TBO DMF gösteren maddenin MH ile belirlenememesi durumu beklenen bir durum olarak yorumlanmaktadır.

Yöntemlerin maddelerdeki DMF miktarlarının büyüklük sıralaması bakımından benzerliklerinin belirlenebilmesi için her iki yöntem kullanıldığında elde edilen DMF büyüklüğü istatistikleri arasında sıralama farkı korelasyon katsayıları elde edilmiştir. İki

yöntem arası DMF büyüklük sıralaması arasında 0,73 büyüklüğünde $p=0.01$ düzeyinde manidar bir ilişki bulunduğu görülmektedir. Bu sonuca göre iki yöntemin DMF büyüklüğü sıralamasına göre uyum gösterdiği yorumuna ulaşılmaktadır. Wiberg (2009), LR ve MH yöntemleri arasında elde edilen korelasyon değeri 0.749 olarak bulduğu ve modeller arasında elde edilen korelasyon değerlerinin yüksek olduğunu belirttiği çalışmasının bulguları ile bu çalışmada elde edilen korelasyon değerleri arasında tutarlılık tespit edilmiştir.

Özet olarak LR ve MH teknikleri DMF büyüklükleri sıralaması bakımından benzer, DMF belirledikleri maddeler bakımından farklı sonuçlar ürettiği belirlenmiştir. Alan yazında bu iki teknikle elde edilen farklı sayıdaki DMF gösteren maddenin sebebi olarak maddelerin güçlük ve ayırt edicilik değerlerindeki, grup ortalamalarındaki ve örneklem büyüklüklerindeki farklılıklar gösterilmektedir (Hidalgo ve Pina, 2004; Narayanan ve Swaminathan, 1996; Penfield ve Camilli, 2007). Pang ve ark. (1994) yaptığı simülasyon çalışmasında ise LR ve MH tekniklerinin her ikisi için de tüm düzeydeki DMF'leri hassasiyetle belirleyebilmek için grupların her birinin 500 ya da 500 'den büyük olması gerektiği belirtilmiştir. Diğer yandan yöntemler arasında DMF büyüklük sıralamalarındaki paralelliğe rağmen DMF düzeylerine ilişkin bulgular ve DMF gösterdiği belirlenen madde sayılarındaki farklılıkların ise teknikten tekniğe değişen DMF'li maddeleri belirlemede kullanılan ölçütlerden kaynaklandığı düşünülmektedir. Bununla ilgili olarak Doğan ve Öğretmen (2008) çalışmalarında Özellikle DMF için Ki-kare değeri üreten ve kullanan bu tekniklerin hepsinde kullanılacak ortak ölçütler geliştirilmenin faydalı olacağını belirtmekte ve bu sayede karşılaştırmaların daha sağlam temellere dayanmasının sağlanabileceğini belirtmektedir.

IV. BÖLÜM

SONUÇ VE ÖNERİLER

Araştırmanın bu bölümünde yapılan analizler sonucunda her bir alt problem için elde edilen sonuçlar ve benzer konularda yapılması düşünülen yeni çalışmalara yönelik önerilere yer verilmiştir.

4.1.SONUÇLAR

LR ve MH Yöntemleri İle DMF Analizi Sonuçları:

Cinsiyete göre MH yöntemi ile yapılan DMF analizi sonuçlarına göre incelenen 34 adet maddenin 4 tanesi B düzeyinde DMF içermektedir. PISA 2009 Fen okuryazarlığı alt testinde 3 tanesi kızlar 1 tanesi ise erkekler lehine çalışmak üzere toplamda 4 adet madde orta düzeyde DMF göstermektedir sonucuna ulaşılmıştır.

Cinsiyete dayalı LR yöntemi ile yapılan incelemede ise yine 4 adet TBO DMF gösterdiği tespit edilen madde belirlenmiştir. Bu maddelere ait etki büyüklükleri hesaplandığında 3 adet maddenin A düzeyinde 1 adet maddenin ise B düzeyinde DMF içerdiği belirlenmiştir. Buna bağlı olarak PISA 2009 fen okuryazarlığı alt testinde yer alan maddelerin 1 tanesinin orta düzeyde TBO DMF gösterdiği diğerlerinin ihmal edilebilir düzeyde olduğu sonucuna ulaşılmıştır.

İki Yöntemin Karşılaştırılmasına Yönelik Analiz Sonuçları:

LR yöntemi ile yapılan DMF analizinde 1 adet maddede orta düzeyde DMF tespit edilirken MH yöntemi ile yapılan analizlerde 4 adet orta düzeyde DMF tespit edilmiştir. Her iki yöntemle DMF gösterdiği tespit edilen maddelerde bir uyum görülmediği buna ek olarak belirlenen DMF düzeyleri bakımından da düşük düzeyde uyum olduğu belirlenmiştir. Diğer yandan LR ile tespit edilen TBO DMF'lerin hiçbirinin MH ile tespit edilemediği, buna ek olarak da MH ile tespit edilen maddelerin hiçbirinin de LR

ile tespit edilemediği, dolayısıyla MH yöntemi ile TBO DMF belirlemezken diğer yandan TB DMF tespitinde MH daha duyarlıdır sonucuna ulaşılmıştır.

Yöntemlerle maddelerde tespit edilen DMF büyüklüklerinin sıralamaları arasında bir ilişki olup olmadığını incelemek için korelasyon katsayısı hesaplanmış, 0.73 büyüklüğünde manidar bir ilişki bulunmuş ve yöntemler arası DMF büyüklüğü sıralamalarında bir benzerlik olduğu sonucuna ulaşılmıştır.

4.2 ÖNERİLER

Aşağıda yer alan önerilerin DMF konusundaki yeni çalışmalara katkı sağlayacağı düşünülmektedir.

1. Bu çalışmada PISA 2009 fen okuryazarlığı alt testindeki maddelere sadece DMF analizi yapılmıştır. DMF gösteren maddeler ayıca incelemeye alınarak DMF gösterme nedenleri, yanlılık içerip içermedikleri, ve DMF içeren maddelerin varlığının testin geçerliğine etkileri de incelenebilir.
2. Maddelerin DMF göstermesinin bir sebebinin de kız ve erkek öğrencilerin farklı ilgi alanları ve yeteneklere sahip olmasından kaynaklandığı literatürdeki bir çok araştırma tarafından kanıtlanmıştır. Test hazırlanırken bu gerçek göz önüne alınarak DMF gösteren madde sayıları gruplar arasında eşitlenmesi yoluna gidilebilir.
3. Bu çalışma Türkiye verisi ve cinsiyet değişkeni ile sınırlı tutulmuştur. Araştırmaya başka değişkenler de eklenebilir veya diğer ülke verileri de dikkate alınarak kültürel ya da dil grupları ile ilgili çalışmalar da yapılabilir.
4. PISA uygulamasının diğer yıllarda yapılan çalışmalarında da DMF analizleri yapılarak genel sonuçlar çıkarılabilir, DMF'ye neden olan noktalar belirlenerek sonraki çalışmalarda benzer hatalar yapılmaması noktasında test geliştiricilere tavsiyelerde bulunulabilir.
5. Son olarak DMF çalışmalarının testlerin geçerlik çalışmaları kapsamında tüm sınavlar için rutin olarak yapılması sağlanmalıdır.

KAYNAKÇA

- Ateşok Deveci, N. (2008). *Üniversitelerarası Kurul Yabancı Dil Sınavının Madde Yanlılığı Bakımından İncelenmesi*. Doktora Tezi. Ankara Üniversitesi, Ankara.
- Bakan Kalaycıoğlu, D. (2008). *Öğrenci Seçme Sınavı'nın Madde Yanlılığı Açısından İncelenmesi*. Doktora Tezi. Hacettepe Üniversitesi, Ankara.
- Bakan Kalaycıoğlu, D. ve Berberoğlu, G. (2010). Differential Item Functioning Analysis of the Science and Mathematics Items in the University Entrance Examinations in Turkey. *Journal of Psychoeducational Assessment*, 20, 1-12.
- Bekçi, B. (2007). *Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavının Değişen Madde Fonksiyonlarının Cinsiyete ve Okul Türüne Göre İncelenmesi*. Yüksek Lisans Tezi. Hacettepe Üniversitesi, Ankara.
- Camili, G. ve Shepard, L. A. (1994). *Methods for identifying biased test items*. London: Sage publications.
- Clauser, B. E. ve Mazor, K. M. (1998). Using statistical procedures to identify differential item functioning test items. *Educational Measurement: Issues and Practice*, 17, 31-44.
- Crocker, L. ve Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Holt, Rinehart and Winston, Inc.
- Çepni, Z. (2011). *Değişen Madde Fonksiyonlarının SIBTEST, Mantel Haenszel, Lojistik Regresyon ve Madde Tepki Kuramı Yöntemleriyle İncelenmesi*. Doktora Tezi. Hacettepe Üniversitesi, Ankara.

- 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.
- Doolittle A. & Welch C. (1989). Gender Differences in Performance on a College Level Achievement Test. (ACT Research Report Series, 89-9, İngiltere).
- Finch, W. H. & French F. B. (2007). Detection of Crossing Differential Item Functioning: A Comparison of Four Methods. *Educational and Psychological Measurement*, 71 (6), SAYFA NUMARALARI YOK.
- Gierl, M. H., Khaliq, S. N. & Boughton, K. (1999, June). *Gender differential item functioning in mathematics and science: prevalence and policy implications*. Paper presented at the Annual Meeting of the Canadian Society for the Study of Education, Canada. Zenisky, A.L.
- Gök, B., Kelecioğlu, H. ve Doğan, N. (2010). Değişen madde fonksiyonunu belirlemede Mantel-Haenszel ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 35, 3-16.
- Hambleton, R.K. ve Robin, F. (2003). DIF Detection and Interpretation in Large-Scale Science Assessments: Informing Item Writing Practices. *Educational Assessment*, 9 (1&2), 61-78.
- Hamilton Laura S. & Snow R. E. (1998). Exploring Differential Item Functioning on Science Achievement Tests. (CSE Technical Report 483, CRESST/Stanford University, Los Angeles).
- Hauger, B.J. ve Sireci, G. S. (2008). Detecting Differential Item Functioning Across Examinees Tested in Their Dominant Language and Examinees Tested in a Second Language. *International Journal of Testing*, 8, 237–250.

- Henderson D. L. (2001). Prevalence of Gender DIF in Mixed Format High School Exit Examinations. (Reports-Research (143)- Speeches-Meeting Papers (150), American Educational Research Association, Seattle, Washington).
- Hidalgo, M. D. & Lopez-Pina, J. (2004). Differential item functioning detection and effect size: A comparison between LR and MH procedures. *Educational and Psychological Measurement*, 64(6), 903–915.
- Holland P. W. & Wainer H. (1993). *Differential Item Functioning*, United States of American Lawrence Erlbaum Associates, Inc.
- Jodoin M. G. & Gierl, M. J. (2001). Evaluating Type I Error and Power Using an Effect Size Measure with the Logistic Regression Procedure for DIF Detection. *Applied Measurement in Education*, 14(4), 329-349.
- Korkmaz, M. (2005). *Madde Cevap Kuramına Dayalı Olarak Çok Kategorili Maddelerde Madde ve Test Yanlılığının (İşlevsel Farklılığın) İncelenmesi*. Doktora tezi. Ege üniversitesi, İzmir.
- Liu, L. O. & Wilson, M. (2009). Gender Differences in Large-Scale Math Assessments. PISA Trend 2000 and 2003, *Applied Measurement in Education*, 22:2, 164-184.
- Le, L.T. (2009). Investigation gender differential item functioning across countries and Test languages for PISA science items. *International Journal of Testing*, 9(2), 122- 133.
- MEB (2010). PISA 2009 Uluslararası Öğrenci Değerlendirme Programı Ulusal Ön Rapor, MEB, Ankara.

Osterlind, J. S. (1983). *Test item bias*. London: Sage Publications

Osterlind J. S. & Everson T. H. (2009). *Differential Item Functioning*, London: Sage Publications, Inc.

Öğretmen, T. (1995). *Differential Item Functioning Analysis of the Verbal Ability Section of the First Stage of the University Entrance Examination in Turkey*. Yüksek Lisans Tezi. ODTÜ, Ankara.

Öğretmen, T. ve Doğan, N. (2004). OKÖSYS matematik alt testine ait maddelerin yanlılık analizi, *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 5(8). SAYFA NUMARALARI YOK.

Özçelik, D. A. (1981). *Ölçme ve Değerlendirme*, Pegem Akademi, ANKARA.

Pang X. L. , Tian F. ve Boss M. W. (1994), Performance of Mantel-Haenszel and Logistic Regression DIF Procedures Over Replications Using Real Data. (Evaluative/Feasibility (142)-Speeches/Conference Papers (150) Reports, University of Ottawa, USA).

Stoneberg B. D. (2004). A Study of Gender-Based and Ethnic-Based Differential Item Functioning (DIF) in the Spring 2003 Idaho Standards Achievement Tests Applying the Simultaneous.

Waller, N. G. (2005) *EZDIF: A computer Program for Detecting Uniform and Non uniform Differential Item Functioning with the Mantel-Haenszel and Logistic Regression Procedures*.

Wiberg, M. (2009). Differential item functioning in mastery tests: A comparison of three methods using real data. *International Journal of Testing*, 9, 41–59.

Yıldırım, H.H. ve Berberoğlu, G. (2006). Judgementel and statistical analyses of the PISA 2003 mathematics literacy items. *International Journal of Testing*, 9(2), 108-121.

Yıldırım, H. H. ve Berberoğlu, G. (2009). Judgmental and Statistical DIF Analyses of the PISA-2003 Mathematics Literacy Items, *International Journal of Testing*, 9 : 108-121.

Yurdugöl, H. (2003). *Ortaöğretim kurumları seçme ve yerleştirme sınavının madde yanlılığı açısından incelenmesi*. Doktora tezi. Hacettepe Üniversitesi, Ankara.

Zumbo, B. D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. Ottawa, ON: Directorate of HumanResources Research and Evaluation, Department of National Defense.

Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, Where it is now, and Where It Is Going. *Language Assessment Quarterly*, 4(2), 223-233.

Zumbo, B. D., & Thomas, D. R. (1996). *A measure of DIF effect size using logistic regression procedures*. Paper presented at the National Board of Medical Examiners, Philadelphia.

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Cansu AYAN

Doğum Yeri ve Tarihi : Çorum 14.06.1987

Eğitim Durumu

Lisans : Hacettepe Üniversitesi - 2009

Eğitim Fakültesi

İlköğretim Matematik Öğretmenliği

İş Deneyimi

2010- ... : Ankara Üniversitesi Eğitim Bilimleri Fakültesi

Araştırma Görevlisi

İletişim

Adres : Ankara Üniversitesi Eğitim Bilimleri Fakültesi/

ANKARA

Telefon : 312 433 6994

Elektronik posta : can_su_ayan@hotmail.com

Tarih : 27.12.2011

EKLER

Ek 1: LR Etki Büyüklüğünü Hesaplamada Kullanılan Syntax

GET

FILE= ' C:\FEN\fen.sav' .

EXECUTE .

compute item=s1 .

compute total=toplam .

compute grp=cinsiyet .

* Bu bolum DMF test etmek için kurulan model 2

*Cozumunu uretir. Model anlamlı ise TBO-DMF var .

*model de cinsiyet ve etkilesim var .

LOGISTIC REGRESSION VAR=item

/METHOD=ENTER total

/METHOD=ENTER grp grp*total

/CONTRAST (cinsiyet)=Indicator

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

execute.

* bu bolum DMF'nu test etmek için kurulan Model 1 .

* Cozumunu uretir . Model anlamlı ise TB-DMF var .

* Modelde yalnızca cinsiyet var .

LOGISTIC REGRESSION VAR=item

/METHOD=ENTER total

/METHOD=ENTER grp

/CONTRAST (grp)=Indicator

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

execute.

Ek 2:MH Analiz Sonuçları

Mantel-Haenszel and Logistic Regression Analysis of

DIFFERENTIAL ITEM FUNCTIONING

Programmed by Niels G. Waller

EZDIF example input file

Reference Group: C:\DIF\erkek

Focal Group: C:\DIF\kiz

Number of Cases in Reference Group: 156

Number of Cases in Focal Group: 182

Conditioning Levels

0 7 14 21 28

6 13 20 27 34

Note:

Alpha > 1.00 favors Reference Group; Alpha < 1.00 favors Focal Group

D-DIF < 0.00 favors Reference Group, D-DIF > 0.00 favors Focal Group

Results for Pass Number: 1

WARNING: Insufficient Data Found in Level: 28 - 34

	ITEM	Alpha	X ²	P-Value	MH D-DIF	SE (MH D-DIF)
A	1	1.341	0.786	0.375	-0.689	0.669
A	2	1.717	3.182	0.074	-1.271	0.665
A	3	0.936	0.002	0.969	0.156	0.773
A	4	1.056	0.001	0.976	-0.128	0.848
B	5	0.593	3.748	0.053	1.227	0.603
B	6*	0.505	6.183	0.013	1.604	0.619
A	7	0.727	1.216	0.270	0.748	0.609
A	8	0.562	2.386	0.122	1.353	0.793
B	9*	1.646	4.081	0.043	-1.171	0.551
A	10	1.045	0.000	0.990	-0.104	0.668
A	11	0.819	0.456	0.500	0.469	0.588
A	12	1.142	0.163	0.686	-0.313	0.591
A	13	0.886	0.135	0.713	0.283	0.576
A	14	1.176	0.247	0.619	-0.381	0.608
A	15	0.948	0.001	0.973	0.126	0.693
A	16	1.416	1.564	0.211	-0.817	0.592
A	17	1.507	1.863	0.172	-0.964	0.649
A	18	1.011	0.015	0.901	-0.026	0.747
A	19	0.745	1.046	0.306	0.693	0.600
A	20	0.865	0.240	0.624	0.341	0.556
A	21	1.436	1.862	0.172	-0.850	0.569
A	22	0.673	2.186	0.139	0.931	0.582
A	23	0.734	1.199	0.273	0.725	0.591
A	24	0.974	0.000	0.983	0.062	0.591
A	25	0.635	2.863	0.091	1.068	0.588
A	26	0.900	0.100	0.752	0.249	0.571
A	27	0.902	0.102	0.750	0.243	0.555
A	28	0.876	0.173	0.677	0.310	0.573
A	29	1.222	0.217	0.641	-0.472	0.755
A	30	1.123	0.134	0.714	-0.274	0.563
B	31	0.556	3.752	0.053	1.377	0.669
A	32	0.519	0.770	0.380	1.540	1.354
A	33	1.136	0.192	0.661	-0.300	0.541
A	34	1.348	0.972	0.324	-0.702	0.625

Number of Items Purged in Pass 1: 0

Item Numbers:

Number of Items Purged in Pass 1: 0
Item Numbers:

Results for Pass Number: 2
WARNING: Insufficient Data Found in Level: 28 - 34

	ITEM	Alpha	X ²	P-value	MH D-DIF	SE (MH D-DIF)
A	1	1.341	0.786	0.375	-0.689	0.669
A	2	1.717	3.182	0.074	-1.271	0.665
A	3	0.936	0.002	0.969	0.156	0.773
A	4	1.056	0.001	0.976	-0.128	0.848
B	5	0.593	3.748	0.053	1.227	0.603
B	6*	0.505	6.183	0.013	1.604	0.619
A	7	0.727	1.216	0.270	0.748	0.609
A	8	0.562	2.386	0.122	1.353	0.793
B	9*	1.646	4.081	0.043	-1.171	0.551
A	10	1.045	0.000	0.990	-0.104	0.668
A	11	0.819	0.456	0.500	0.469	0.588
A	12	1.142	0.163	0.686	-0.313	0.591
A	13	0.886	0.135	0.713	0.283	0.576
A	14	1.176	0.247	0.619	-0.381	0.608
A	15	0.948	0.001	0.973	0.126	0.693
A	16	1.416	1.564	0.211	-0.817	0.592
A	17	1.507	1.863	0.172	-0.964	0.649
A	18	1.011	0.015	0.901	-0.026	0.747
A	19	0.745	1.046	0.306	0.693	0.600
A	20	0.865	0.240	0.624	0.341	0.556
A	21	1.436	1.862	0.172	-0.850	0.569
A	22	0.673	2.186	0.139	0.931	0.582
A	23	0.734	1.199	0.273	0.725	0.591
A	24	0.974	0.000	0.983	0.062	0.591
A	25	0.635	2.863	0.091	1.068	0.588
A	26	0.900	0.100	0.752	0.249	0.571
A	27	0.902	0.102	0.750	0.243	0.555
A	28	0.876	0.173	0.677	0.310	0.573
A	29	1.222	0.217	0.641	-0.472	0.755
A	30	1.123	0.134	0.714	-0.274	0.563
B	31	0.556	3.752	0.053	1.377	0.669
A	32	0.519	0.770	0.380	1.540	1.354
A	33	1.136	0.192	0.661	-0.300	0.541
A	34	1.348	0.972	0.324	-0.702	0.625

Ek 3: LR Analiz Sonuçları

Logistic Regression Procedure. Note: Items purged from previous step

Item Number: 1	Estimate	SE Estimate	Z	p-value
Intercept	-0.438071	0.391584	-1.118715	0.263530
Trait	0.139411	0.028670	4.862658	0.000001
Group	0.974024	0.709867	1.372122	0.170333
Trait x Group	-0.056650	0.057339	-0.987984	0.323399

Item Number: 2	Estimate	SE Estimate	Z	p-value
Intercept	-4.317811	0.504921	-8.551461	0.000000
Trait	0.327229	0.037058	8.830268	0.000000
Group	-0.631624	0.979958	-0.644542	0.519372
Trait x Group	0.092530	0.074115	1.248467	0.212152

Item Number: 3	Estimate	SE Estimate	Z	p-value
Intercept	-4.570109	0.505684	-9.037486	0.000000
Trait	0.182337	0.029818	6.114937	0.000000
Group	-1.199667	1.100593	-1.090019	0.275968
Trait x Group	0.067275	0.059636	1.128084	0.259555

Item Number: 4	Estimate	SE Estimate	Z	p-value
Intercept	-4.319183	0.501586	-8.611058	0.000000
Trait	0.152934	0.029557	5.174114	0.000000
Group	-0.275882	1.087651	-0.253649	0.799819
Trait x Group	0.023464	0.059115	0.396917	0.691513

Item Number: 5	Estimate	SE Estimate	Z	p-value
Intercept	-1.738871	0.378299	-4.596554	0.000005
Trait	0.186670	0.027553	6.775008	0.000000
Group	-1.270309	0.706966	-1.796847	0.072662
Trait x Group	0.062829	0.055105	1.140158	0.254493

Item Number: 6	Estimate	SE Estimate	Z	p-value
Intercept	-3.518110	0.429357	-8.193896	0.000000
Trait	0.254813	0.030080	8.471151	0.000000
Group	-1.429648	0.850046	-1.681847	0.092911
Trait x Group	0.057245	0.060160	0.951543	0.341558

Item Number: 7	Estimate	SE Estimate	Z	p-value
Intercept	-3.483680	0.405429	-8.592566	0.000000
Trait	0.234704	0.027849	8.427627	0.000000
Group	-0.441000	0.813305	-0.542232	0.587780
Trait x Group	0.013334	0.055699	0.239389	0.810853
Item Number: 8	Estimate	SE Estimate	Z	p-value
Intercept	-6.499644	0.653334	-9.948421	0.000000
Trait	0.325754	0.039095	8.332288	0.000000
Group	-1.166868	1.429196	-0.816450	0.414437
Trait x Group	0.037361	0.078191	0.477823	0.632881
Item Number: 9	Estimate	SE Estimate	Z	p-value
Intercept	-1.250273	0.300582	-4.159506	0.000035
Trait	0.098080	0.020570	4.768139	0.000002
Group	0.213443	0.595058	0.358692	0.719901
Trait x Group	0.016412	0.041140	0.398930	0.690030
Item Number: 10	Estimate	SE Estimate	Z	p-value
Intercept	-3.669463	0.412018	-8.906083	0.000000
Trait	0.162282	0.025648	6.327396	0.000000
Group	1.837956	0.873136	2.105006	0.035540
Trait x Group	-0.121701	0.051295	-2.372574	0.017853
Item Number: 11	Estimate	SE Estimate	Z	p-value
Intercept	-2.488281	0.337678	-7.368793	0.000000
Trait	0.125936	0.021953	5.736647	0.000000
Group	-1.045046	0.697494	-1.498286	0.134375
Trait x Group	0.057826	0.043906	1.317049	0.188124
Item Number: 12	Estimate	SE Estimate	Z	p-value
Intercept	-1.618187	0.314016	-5.153193	0.000000
Trait	0.051522	0.020494	2.514019	0.012092
Group	-0.320679	0.639760	-0.501249	0.616306
Trait x Group	0.040029	0.040988	0.976608	0.328999
Item Number: 13	Estimate	SE Estimate	Z	p-value
Intercept	-2.393223	0.331169	-7.226595	0.000000
Trait	0.133337	0.021826	6.109186	0.000000
Group	-0.525097	0.677228	-0.775362	0.438309
Trait x Group	0.023457	0.043651	0.537360	0.591139

Item Number: 14	Estimate	SE Estimate	Z	p-value
Intercept	-1.553976	0.372915	-4.167105	0.000033
Trait	0.179297	0.027114	6.612596	0.000000
Group	0.687027	0.689202	0.996844	0.319081
Trait x Group	-0.045003	0.054229	-0.829865	0.406813
Item Number: 15	Estimate	SE Estimate	Z	p-value
Intercept	-5.246535	0.520360	-10.082506	0.000000
Trait	0.294474	0.033546	8.778195	0.000000
Group	-1.077247	1.102232	-0.977333	0.328641
Trait x Group	0.065591	0.067092	0.977633	0.328492
Item Number: 16	Estimate	SE Estimate	Z	p-value
Intercept	-3.188836	0.383923	-8.305919	0.000000
Trait	0.211406	0.026153	8.083323	0.000000
Group	1.162221	0.771774	1.505909	0.132406
Trait x Group	-0.050172	0.052307	-0.959189	0.337695
Item Number: 17	Estimate	SE Estimate	Z	p-value
Intercept	-4.046453	0.425799	-9.503191	0.000000
Trait	0.224295	0.027492	8.158589	0.000000
Group	1.504444	0.890417	1.689595	0.091417
Trait x Group	-0.063749	0.054984	-1.159421	0.246561
Item Number: 18	Estimate	SE Estimate	Z	p-value
Intercept	-4.753870	0.491855	-9.665196	0.000000
Trait	0.210039	0.029584	7.099792	0.000000
Group	1.418189	1.063351	1.333698	0.182607
Trait x Group	-0.090226	0.059168	-1.524919	0.127596
Item Number: 19	Estimate	SE Estimate	Z	p-value
Intercept	-1.366158	0.362616	-3.767510	0.000175
Trait	0.163702	0.026185	6.251678	0.000000
Group	-0.111354	0.674033	-0.165205	0.868816
Trait x Group	-0.011373	0.052371	-0.217154	0.828132
Item Number: 20	Estimate	SE Estimate	Z	p-value
Intercept	-1.623093	0.311884	-5.204153	0.000000
Trait	0.119135	0.021240	5.609015	0.000000
Group	-0.075930	0.618276	-0.122810	0.902282
Trait x Group	-0.007364	0.042480	-0.173353	0.862409

Item Number: 21	Estimate	SE Estimate	Z	p-value
Intercept	-2.230954	0.325226	-6.859704	0.000000
Trait	0.132820	0.021698	6.121261	0.000000
Group	0.377140	0.659327	0.572008	0.567445
Trait x Group	0.000724	0.043396	0.016685	0.986691

Item Number: 22	Estimate	SE Estimate	Z	p-value
Intercept	-1.910253	0.366868	-5.206926	0.000000
Trait	0.177340	0.026532	6.684099	0.000000
Group	-2.108592	0.705200	-2.990063	0.002857
Trait x Group	0.139238	0.053063	2.623993	0.008823

Item Number: 23	Estimate	SE Estimate	Z	p-value
Intercept	-2.382753	0.331673	-7.184052	0.000000
Trait	0.124570	0.021676	5.746888	0.000000
Group	-0.865721	0.682109	-1.269183	0.204671
Trait x Group	0.043807	0.043352	1.010501	0.312500

Item Number: 24	Estimate	SE Estimate	Z	p-value
Intercept	-3.182452	0.371697	-8.561944	0.000000
Trait	0.189376	0.024684	7.671853	0.000000
Group	0.310519	0.760506	0.408305	0.683137
Trait x Group	-0.016059	0.049369	-0.325276	0.745040

Item Number: 25	Estimate	SE Estimate	Z	p-value
Intercept	-2.534025	0.344223	-7.361584	0.000000
Trait	0.161504	0.023153	6.975469	0.000000
Group	-0.488937	0.694390	-0.704125	0.481519
Trait x Group	0.009028	0.046306	0.194964	0.845461

Item Number: 26	Estimate	SE Estimate	Z	p-value
Intercept	-1.871881	0.310638	-6.025921	0.000000
Trait	0.084236	0.020266	4.156533	0.000035
Group	0.179548	0.635401	0.282575	0.777561
Trait x Group	-0.011487	0.040532	-0.283407	0.776924

Item Number: 27	Estimate	SE Estimate	Z	p-value
Intercept	-1.987632	0.321071	-6.190638	0.000000
Trait	0.122132	0.021464	5.690162	0.000000
Group	1.031520	0.644763	1.599845	0.109949
Trait x Group	-0.085194	0.042927	-1.984599	0.047462

Item Number: 28	Estimate	SE Estimate	Z	p-value
Intercept	-1.226671	0.298065	-4.115443	0.000042
Trait	0.036911	0.019621	1.881227	0.060231
Group	-0.487194	0.603573	-0.807184	0.419753
Trait x Group	0.029209	0.039242	0.744333	0.456850
Item Number: 29	Estimate	SE Estimate	Z	p-value
Intercept	-4.261717	0.466642	-9.132736	0.000000
Trait	0.173459	0.028021	6.190404	0.000000
Group	-0.426022	1.007283	-0.422941	0.672429
Trait x Group	0.040334	0.056041	0.719712	0.471870
Item Number: 30	Estimate	SE Estimate	Z	p-value
Intercept	-2.105627	0.322465	-6.529776	0.000000
Trait	0.124784	0.021617	5.772364	0.000000
Group	-0.827169	0.653399	-1.265947	0.205827
Trait x Group	0.069186	0.043235	1.600244	0.109860
Item Number: 31	Estimate	SE Estimate	Z	p-value
Intercept	-2.105036	0.512550	-4.106983	0.000043
Trait	0.274851	0.038823	7.079589	0.000000
Group	0.359059	0.860460	0.417288	0.676558
Trait x Group	-0.080393	0.077646	-1.035381	0.300741
Item Number: 32	Estimate	SE Estimate	Z	p-value
Intercept	-8.304235	1.254867	-6.617621	0.000000
Trait	0.300354	0.063941	4.697345	0.000003
Group	-5.730612	2.847956	-2.012184	0.044468
Trait x Group	0.264560	0.127883	2.068768	0.038824
Item Number: 33	Estimate	SE Estimate	Z	p-value
Intercept	-1.333112	0.299459	-4.451741	0.000009
Trait	0.092839	0.020173	4.602019	0.000005
Group	0.694548	0.596837	1.163716	0.244817
Trait x Group	-0.045993	0.040347	-1.139938	0.254585
Item Number: 34	Estimate	SE Estimate	Z	p-value
Intercept	-3.396484	0.382550	-8.878536	0.000000
Trait	0.184873	0.024791	7.457211	0.000000
Group	1.112625	0.794431	1.400531	0.161665
Trait x Group	-0.053698	0.049582	-1.083003	0.279068

Ek 4 : LR Yöntemine Ait R^2 Değerleri

No	Hiyerarşik Nagelkerke R2 Değerleri				
	M1	M2	M3	M3-M1	M2-M1
1	0,120	0,130	0,126	0,006	0,010
2	0,472	0,485	0,481	0,009	0,013
3	0,222	0,228	0,222	0,000	0,006
4	0,162	0,164	0,163	0,001	0,002
5	0,256	0,274	0,270	0,014	0,018
6	0,395	0,415	0,412	0,017	0,020
7	0,367	0,370	0,370	0,003	0,003
8	0,483	0,491	0,490	0,007	0,008
9	0,089	0,103	0,102	0,013	0,014
10	0,194	0,217	0,195	0,001	0,023
11	0,150	0,158	0,152	0,002	0,008
12	0,027	0,036	0,032	0,005	0,009
13	0,169	0,172	0,171	0,002	0,003
14	0,218	0,221	0,219	0,001	0,003
15	0,457	0,460	0,457	0,000	0,003
16	0,306	0,319	0,316	0,010	0,013
17	0,327	0,343	0,339	0,012	0,016
18	0,284	0,294	0,285	0,001	0,010
19	0,203	0,207	0,207	0,004	0,004
20	0,140	0,142	0,142	0,002	0,002
21	0,157	0,167	0,167	0,010	0,010
22	0,235	0,266	0,242	0,007	0,031
23	0,150	0,156	0,153	0,003	0,006
24	0,278	0,279	0,279	0,001	0,001
25	0,229	0,236	0,236	0,007	0,007
26	0,073	0,073	0,073	0,000	0,000
27	0,137	0,152	0,138	0,001	0,015
28	0,016	0,019	0,016	0,000	0,003
29	0,215	0,220	0,218	0,003	0,005
30	0,146	0,156	0,147	0,001	0,010
31	0,343	0,355	0,352	0,009	0,012
32	0,285	0,322	0,285	0,000	0,037
33	0,085	0,090	0,085	0,000	0,005
34	0,258	0,266	0,262	0,004	0,008