

T.C
ABANT İZZET BAYSAL ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ESTİTÜSÜ

PISA 2009 MATEMATİK OKURYAZARLIĞI ALT TESTİNDE BULUNAN
MADDELERİNİN MANTEL-HAENSZEL, SIBTEST VE LOJİSTİK
REGRESYON YÖNTEMLERİ İLE DEĞİŞEN MADDE FONKSİYONUNUN
İNCELENMESİ

Süleyman DEMİR

Haziran, 2013

ABANT İZZET BAYSAL ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ESTİTÜSÜ
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI

PISA 2009 MATEMATİK OKURYAZARLIĞI ALT TESTİNDE BULUNAN
MADDELERİNİN MANTEL-HAENZEL, SIBTEST VE LOJİSTİK
REGRESYON YÖNTEMLERİ İLE DEĞİŞEN MADDE FONKSİYONUNUN
İNCELENMESİ

Yüksek Lisans Tezi

Hazırlayan
Süleyman DEMİR

Danışman
Yrd. Doç. Dr. İ. Alper KÖSE

Bolu-2013

BİLDİRİM

Yüksek lisans tezi olarak sunduğum “PISA 2009 Matematik Okuryazarlığı Alttestinde Bulunan Maddelerin Mantel-Haenzsel, SIBTEST ve Lojistik Regresyon Yöntemleri ile Değişen Madde Fonksiyonunun İncelenmesi” isimli araştırmamın yazılmasında, bilimsel ve etik kurallarına uyulduğu başkalarının eserlerinden yararlanılması durumunda atıfta bulunulduğunu, kullanılan verilerde herhangi bir tahribat yapılmadığını, tezin tamamının ya da bir kısmının bu üniversite veya başka üniversitede tez çalışması olarak sunulmadığını beyan ederim.

13.06.2013

Süleyman DEMİR



EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE,

Süleyman DEMİR'e ait PISA 2009 Matematik Okuryazarlığı Alt Testinde Bulunan Maddelerinin Mantel-Haenszel, SIBTEST ve Lojistik Regresyon Yöntemleri ile Değişen Madde Fonksiyonunun İncelenmesi adlı çalışma, jürimiz tarafından Eğitimde Ölçme ve Değerlendirme Anabilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir. (05/06/2018)

Akademik Unvan ve Adı Soyadı

İmza

Üye (Tez Danışmanı) :Yrd. Doç. Dr. İ. Alper KÖSE

Üye :Doç. Dr. Zekeriya NARTGÜN

Üye :Yrd. Doç. Dr. Neşe GÜLER



Eğitim Bilimleri Enstitüsünün Onayı

Prof. Dr. Soner DURMUŞ

Enstitü Müdürü

Abstract

AN ANALYSIS OF THE DIFFERENTIAL ITEM FUNCTION FOR THE ITEMS AVAILABLE IN THE PISA 2009 MATHEMATICS LITERACY SUB-TEST THROUGH MANTEL-HAENZSEL, SIBTEST AND LOGISTIC REGRESSION METHODS

Measurement and Assessment in Education Programme

Thesis Advisor: Assist. Prof. Dr. İ. Alper KÖSE

June, 2013 xii + 64 Pages

This study performs a Differential Item Function (DIF) analysis in terms of gender and culture on the items available in the PISA 2009 mathematics literacy sub-test. The DIF analyses were done through the Mantel Haenszel, Logistic Regression and the SIBTEST methods.

The data for the gender variable were collected from the responses given by 332 students to the items in the mathematics literacy sub-test during the administration of the 5th booklet in the PISA 2009 application whereas the data for the culture variable were collected through the application of the 5th booklet in Turkey, Germany, Finland and the United States in the PISA 2009 application. Besides, the descriptive statistics and the reliability analyses were conducted for the mathematics literacy sub-test, and general information concerning the group distribution prior to the DIF analysis was obtained.

Following the DIF analysis, which had been performed for the gender variable, 2 DIFs were specified in the MH technique while 3 were found in the LR technique and 4 were found in the SIBTEST technique. It may be said that 4 of the 5 items with DIF were in favour of male students, and only one was in favour of female students.

Based on the MH analysis results which had been performed in relation to the data obtained from the students in Turkey and Germany, the DIF was found in 10 items; and 5 of them were to the advantage of students in Germany and the other items were to the advantage of those in Turkey. With the LR analysis the DIF was found in 10 items, and 7 of them were to the advantage of German students and 3 were to the advantage of Turkish students. In consequence of the SIBTEST, the DIF was found in 12 items; 7 of

them were found to be to the advantage of German students and 5 were found to be to the advantage of Turkish students.

According to the results of the MH analyses conducted for the data obtained from students in Turkey and Finland, the DIF was found in 12 items; and 6 of the items were to the advantage of Finnish students whereas the other 6 were to the advantage of Turkish students. In consequence of the LR analysis, the DIF was found in 10 items; and all of the items were to the advantage of Finnish students. The SIBTEST results, on the other hand, demonstrated that the DIF was available in 13 items; and that 7 of them were to the advantage of Finnish students while 6 were to the advantage of Turkish students.

Based on the MH analysis results for the data obtained from the students in Turkey and in the United States, the DIF was found for 11 items; and 5 of them were to the advantage of American Students whereas 6 were to the advantage of Turkish students. According to the LR analysis results, the DIF was found in 9 items; and 4 of them were to the advantage of American students while 5 were to the advantage of Turkish students. According to the results of the SIBTEST, however, the DIF was found in 13 items; and 8 of them were to the advantage of American students while 5 were to the advantage of Turkish students.

Özet

Bu çalışmada, PISA 2009 uygulaması Matematik okuryazarlığı alttestinde, cinsiyete ve kültüre göre değişen madde fonksiyonu (DMF) içeren madde bulunup bulunmadığı incelenmiştir. DMF analizleri Mantel Haenszel, Lojistik Regresyon ve SIBTEST yöntemleriyle gerçekleştirilmiştir.

Çalışmanın veri setini, cinsiyet değişkeni için PISA 2009 uygulamasında Türkiye’de 5. kitapçığın uygulandığı 332 öğrencinin matematik okuryazarlığı alt testindeki maddelere verdikleri cevaplar, kültür değişkeni için ise PISA 2009 uygulamasında Türkiye, Almanya, Finlandiya ve ABD’deki öğrencilerin 5. kitaptaki matematik testine verdikleri cevaplar oluşturmaktadır. Ayrıca Matematik 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 MH tekniğinde 2 adet DMF belirlenirken, LR tekniğinde 3 adet DMF, SIBTEST tekniğinde 4 adet DMF gösteren madde belirlenmiştir. Toplamda DMF’li olduğu belirlenen 5 maddenin 4’ü erkekler lehine, sadece bir tane maddenin kızlar lehine avantajlı olduğu söylenebilir.

Türkiye ile Almanya’dan katılan öğrencilerden elde edilen veri setlerine göre yapılan MH analizi sonuçlarına göre 10 maddede DMF bulunmuş, bu maddelerin 5’i Almanya için avantajlı diğer maddeler ise Türkiye için avantajlı olduğu görülmüştür. LR analizi sonucunda 10 maddede DMF bulunmuş, bu maddelerin 7 tanesi Almanya lehine 3 tanesi Türkiye lehine avantaj sağlamaktadır. SIBTEST analizi sonucunda 12 maddede DMF bulunmuş ve bu maddelerin 7 tanesi Almanya lehine 5 tanesi ise Türkiye lehine avantaj sağlamaktadır.

Türkiye ile Finlandiya’dan katılan öğrencilerden elde edilen veri setlerine göre yapılan MH analizi sonuçlarına göre 12 maddede DMF bulunmuş, bu maddelerin 6 tanesi Finlandiya için avantajlı 6 tanesi ise Türkiye için avantajlı olduğu görülmüştür. LR analizi sonucunda 10 maddede DMF bulunmuş, bu maddelerin tamamı Finlandiya

lehine avantaj sağlamaktadır. SIBTEST analizi sonucunda 13 maddede DMF bulunmuş ve bu maddelerin 7 tanesi Finlandiya lehine 6 tanesi ise Türkiye lehine avantaj sağlamaktadır.

Türkiye ile ABD'dan katılan öğrencilerden elde edilen veri setlerine göre yapılan MH analizi sonuçlarına göre 11 maddede DMF bulunmuş, bu maddelerin 5 tanesi ABD için avantajlı 6 tanesi ise Türkiye için avantajlı olduğu görülmüştür. LR analizi sonucunda 9 maddede DMF bulunmuş, bu maddelerin 4 tanesi ABD lehine, 5 tanesi ise Türkiye avantaj sağlamaktadır. SIBTEST analizi sonucunda 13 maddede DMF bulunmuş ve bu maddelerin 8 tanesi ABD lehine 5 tanesi ise Türkiye lehine avantaj sağlamaktadır.

Teşekkür

Tez çalışmam ve yüksek lisans eğitimim boyunca değerli yorumlarını benden esirgemeyen başta tez danışmanım ve değerli hocam Yrd. Doç. Dr. İ. Alper KÖSE'ye, gelişimimde büyük katkısı bulunan Doç. Dr. Zekeriya NARTGÜN'e, Doç. Dr. Zeki ARSAL'a ve Yrd. Doç. Dr. Sevilay KİLMEN'e sonsuz teşekkürlerimi bir borç bilirim.

Akademik hayatı tüm olumsuzluklarına rağmen bana sevdiren ve kendimi hiçbir zaman ayrı düşünmek istemediğim başta Yrd. Doç. Dr. Neşe GÜLER olmak üzere Öğr. Gör. Gülşen TAŞDELEN TEKER'e ve Arş. Gör. Gülden KAYA UYANIK'a,

Tez çalışmam boyunca her çıkmaza düştüğümde bana yardımcı olan ve bana katlanan mesai ve oda arkadaşlarım Arş. Gör. Tuğba Seda ÇOLAK'a Arş. Gör. Betül DÜŞÜNCELİ'ye, ve Arş. Gör. Ahmet FİDAN'a,

Sevgimi, bağlılığımı kelimelerle ifade edemeyeceğim, hayatımda aldığım kararların hepsine onay veren AİLEME,

Son olarak hayatımda bulunmalarından son derece memnuniyet duyduğum değerli arkadaşlarıma TEŞEKKÜR EDERİM.

İçindekiler

Abstract	v
Özet	vii
Teşekkür	ix
İçindekiler	x
Tablolar listesi	xii
BÖLÜM I	1
1 Giriş	1
1.1 Problem Durumu	1
1.2 Problem Cümlesi	7
1.3 Alt problemler	7
1.4 Araştırmanın Amacı	7
1.5 Araştırmanın Önemi	8
1.6 Sınırlılıklar	8
1.7 Sayıtlar	9
1.8 Tanımlar	9
BÖLÜM II	10
2 Kuramsal Temeller ve İlgili Literatür	10
2.1 Madde Yanlılığı	11
2.2 DMF Belirleme Yöntemleri	14
2.3 İlgili Araştırmalar	23
BÖLÜM III	29
3 Yöntem	29
3.1 Araştırma Modeli	29
3.2 Çalışma Grubu ve Veri Toplama Araçları	29
3.3 Verilerin Çözümlemesi	32

3.4	Çalışma Grubuna ait İstatistikler	33
BÖLÜM IV		38
4	Bulgular ve Yorum	38
4.1	Birinci Alt Probleme İlişkin Bulgular	38
4.2	İkinci Alt Probleme İlişkin Bulgular	40
4.3	Üçüncü Alt Probleme İlişkin Bulgular	42
4.4	Dördüncü Alt Probleme İlişkin Bulgular	44
4.5	Beşinci Alt Probleme İlişkin Bulgular	46
4.6	Altıncı Alt Probleme İlişkin Bulgular	48
BÖLÜM V		51
5	Sonuçlar ve Öneriler	51
5.1	Sonuçlar	51
5.2	Öneriler	56
KAYNAKÇA		59

Tablolar listesi

Tablo 1: PISA sınavlarında değerlendirmede temele alınan alt testler.....	4
Tablo 2: Matematik Maddelerinin Ölçtüğü Alt Boyutlar	5
Tablo 3: Mantel-Haenszel yöntemi olasılık tablosunda kullanılan veri düzeni.....	17
Tablo 4: Gruplara göre düzenlenmiş madde güçlük indeksleri veri düzeni	18
Tablo 5: Zieky (1993) tarafından düzenlenen sınıflama sistemi	19
Tablo 6: Zumbo ve Thomas (1996) ve Jodoin ve Gierl (2001) tarafından düzenlenen sınıflama sistemleri.....	21
Tablo 7: Bakan Kalaycıoğlu ve Kelecioğlu (2011) tarafından düzenlenen sınıflama sistemi.....	21
Tablo 8: Roussos ve Stout (1996) tarafından düzenlenen sınıflama sistemi	23
Tablo 9: Matematik Okuryazarlığına Ait Madde Sayıları	31
Tablo 10: Cinsiyete Göre Betimsel İstatistikler.....	33
Tablo 11: Öğrencilerin Katıldıkları Ülkelere Göre Betimsel İstatistikler	34
Tablo 12: Maddelere ait güçlük indeksleri	35
Tablo 13: Doğrulayıcı Faktör Analizi Model Veri Uyumu İstatistikleri	37
Tablo 14: Matematik testine ait Mantel-Haenszel yöntemine göre yapılan analiz sonuçları.....	39
Tablo 15: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler.	40
Tablo 16: Matematik testine ait Lojistik Regresyon yöntemine göre yapılan analiz sonuçları.....	41
Tablo 17: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler.	42
Tablo 18: Matematik testine ait SIBTEST yöntemine göre yapılan analiz sonuçları	43
Tablo 19:DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler..	44
Tablo 20: Matematik testine ait Mantel-Haenszel yöntemine göre yapılan analiz sonuçları.....	45

Tablo 21: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler.	46
Tablo 22: Matematik testine ait Lojistik Regresyon yöntemine göre yapılan analiz sonuçları.....	47
Tablo 23: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler.	48
Tablo 24: Matematik testine ait SIBTEST yöntemine göre yapılan analiz sonuçları	49
Tablo 25: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler.	50
Tablo 26: Cinsiyet Değişkenine göre Elde Edilen DMF’li Maddeler ve Avantaj sağladıkları gruplar	51
Tablo 27: Ülke Değişkenine göre Elde Edilen DMF’li Maddeler ve Avantaj sağladıkları gruplar.....	53
Tablo 28: Ülke Değişkenine göre Elde Edilen DMF’li Maddeler ve Avantaj sağladıkları gruplar.....	54
Tablo 29: Ülke Değişkenine göre Elde Edilen DMF’li Maddeler ve Avantaj sağladıkları gruplar.....	55

BÖLÜM I

1 Giriş

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

1.1 Problem Durumu

Gelişmiş ve gelişmekte olan dünya ülkelerinin çoğunda olduğu gibi Türkiye’de de son yıllarda hem kendi ekonomik, sosyal ve kültürel değerlerini hem de dünyanın evrensel değerlerini dikkate alarak, eğitim alanında reform niteliği taşıyan çalışmalar yapılmaktadır. Milli Eğitim bakanlığı eğitim-öğretim programlarındaki yapmış olduğu bu reformların ihtiyacı karşılama derecesini ulusal testlerin yanı sıra uluslararası testler ile de ölçmektedir. Ulusal boyutta; Seviye Belirleme-SBS, Yükseköğretime Geçiş-YGS, Lisans Yerleştirme-LYS sınavları yapılırken, uluslararası boyutta yapılan çalışmalar arasında ise Uluslararası Öğrenci Değerlendirme Programı-PISA (The Programme for International Student Assessment), Üçüncü Uluslararası Fen ve Matematik Çalışması-TIMSS, Okuma Becerilerinde Gelişim-PIRLS sınavları bulunmaktadır.

Millî Eğitim Bakanlığı eğitim sistemiyle ilgili sağlıklı ve karşılaştırmalı veriler almak amacıyla ulusal ve uluslararası düzlemde periyodik olarak, ölçme-değerlendirme çalışmaları gerçekleştirmektedir. Eğitimi Araştırma ve Geliştirme Daire Başkanlığı tarafından yürütülen ulusal düzeydeki “Öğrenci Başarılarını Belirleme Sınavı (ÖBBS)” çalışması ile uluslararası ölçme-değerlendirme çalışmaları olan TIMSS, PIRLS ve PISA projeleri eğitim sistemiyle ilgili veriler sunmaktadır. Bu çalışmalarda öğrencilere

yönelik standart başarı testleri ile öğrenci, öğretmen ve okul anketleri uygulanmakta; buradan elde edilen bilgilerle temel eğitimi tamamlayan öğrencilerin neleri bildikleri, hangi becerileri ne derecede kazandıkları ölçülmeye çalışılmaktadır. Çalışmalardan elde edilen sonuçlarla eğitim sisteminin mevcut durumu tespit edilmekte; özellikle temel eğitimde eğitim politikaları, öğretim programları, öğretim yöntem ve teknikleri, öğretmenlerin yeterlilikleri, ders araç-gereç ve materyalleri gibi unsurların gözden geçirilmesine yönelik bilimsel veriler elde edilmekte, diğer ülkelerin eğitim sistemleriyle de karşılaştırmalar yapılmaktadır. Elde edilen bilgiler, eğitimde politika belirleyicilerin, karar vericilerin, akademisyenlerin ve üniversitelerin, eğitim programcılarının, eğitim yöntem ve teknikleri ile eğitim materyali hazırlayanların, eğitimle ilgilenen sivil toplum kuruluşlarının ve araştırmacıların hizmetine sunmaktadır (EARGED, 2011).

Ekonomik İşbirliği ve Kalkınma Teşkilatı-OECD (Organisation for Economic Co-Operation and Development) tarafından düzenlenen PISA; Matematik, Fen ve Okuma becerilerini ölçmeyi amaçlayan okuryazarlık alt testlerinden oluşmaktadır. Üç yılda bir yapılan bu araştırmalarla, OECD üyesi ülkeler ve diğer katılımcı ülkelerdeki (dünya ekonomisinin yaklaşık olarak %90'ı) 15 yaş grubu öğrencilerin modern toplumda yerlerini alabilmeleri için gereken temel bilgi ve becerilere ne ölçüde sahip oldukları değerlendirilmektedir (OECD, 2009).

PISA'nın değerlendirme çerçevesi ve kavramsal temelleri, projeye katılan ülkelerdeki uzmanlar tarafından belirlenmekte, yapılan görüşmeler sonrasında katılımcı ülkelerin görüş 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 (EARGED, 2010).

PISA projesinin geliştirilmesinde öne çıkan temel özellikler şunlardır (EARGED, 2010):

- *Politika yönlendirici özelliği:* Başarı düzeyleri arasındaki farklılıklara dikkat çekmek ve yüksek performans standardı olan öğrenciler, okullar ve eğitim sistemlerinin özelliklerini belirlemek amacıyla; öğrenme çıktılarıyla ilgili veriler, öğrenci özellikleri ve okul içinde ve dışında öğrenmeyi şekillendiren etkenlerle ilgili veriler arasında bir ilişki kurar.

- *Yeni bir “okuryazarlık” kavramı:* PISA’da kullanılan okuryazarlık kavramı, öğrencilerin temel konu alanlarındaki çeşitli durumlarda karşılarına çıkan problemleri yorumlarken ve çözerken, bilgi ve becerilerini kullanma, analiz etme, mantıksal çıkarımlar yapma ve etkili iletişim kurma yeterlikleri ile ilgilidir.

- *Yaşam boyu öğrenmeyle ilgili olması:* Öğrencilerin belirli konu alanlarındaki yeterliklerinin değerlendirilmesinin yanı sıra, bu projede öğrencilerin öğrenmeye yönelik motivasyonları, kendileri hakkındaki düşünceleri ve öğrenme stratejileri hakkındaki veriler de toplanmaktadır.

- *Düzenli olarak yapılması:* Değerlendirmenin belirli aralıklarla yapılması ülkelerin eğitim ile ilgili temel hedeflerinin ne kadarına ulaştıklarını izlemelerine imkân tanır.

- *Geniş coğrafi kapsamı ve iş birliğine dayalı yapısı:* PISA 2009 araştırmasına, OECD üyesi 33 ülke ile üye olmayan 32 ülke katılmıştır.

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 (OECD, 2009).

Türkiye’ deki PISA projesinde her dönem Okuma Becerileri, Matematik ve Fen Okuryazarlığı alanlarından sadece birine temel alan olarak ağırlık verilmektedir. Bununla birlikte, diğer iki alan da yapılan değerlendirme kapsamına dâhil edilmektedir. Dokuz yıllık bir döngüde bu alanlardan her biri, bir kez temel alan olmaktadır. İlk defa 2000 yılında gerçekleştirilen PISA uygulamasında temel alan Okuma Becerileri, 2003’te Matematik okuryazarlığı, 2006’da ise Fen okuryazarlığı olmuştur. 2009 yılında

ise dokuz yıllık yeni bir döngü başlamıştır. 2009 yılında ağırlıklı alan Okuma Becerileridir.

Tablo 1: PISA sınavlarında değerlendirmede temele alınan alt testler

Uygulama Yılı	2000	2003	2006	2009	2012	2015
Değerlendirmede Temele Alınan Alttest	Okuma Becerileri	Matematik	Fen Bil.	Okuma Becerileri	Matematik	Fen Bil.

Öğrencilerin üç temel konu alanındaki bilgi ve becerilerini değerlendirmenin yanında PISA, 2000 yılındaki uygulamasında öğrencilere motivasyonları ve öğrenmeye yönelik tutumları, bilgisayar kullanımına yatkınlıkları ile öz düzenleyici öğrenme (self-regulated learning) başlığı altında kendi öğrenme süreçlerini yönetmeleri ve izlemeleriyle ilgili sorular sorulmuştur. PISA 2003'te, bu öğeler geliştirilmiş ve değerlendirmeye problem çözme bilgi ve becerileri de dâhil edilmiştir. Fen okuryazarlığının değerlendirildiği PISA 2006'da öğrencilerin motivasyonlarının ve tutumlarının değerlendirilmesine devam edilmiştir. Okuma becerilerinin ağırlıklı alan olarak ele alındığı PISA 2009'da ise, öğrencilerin okuma etkinliklerine katılımına, kendi okuma ve öğrenme stratejileri hakkındaki düşüncelerine odaklanılmıştır (OECD, 2009).

PISA uygulamasına katılan öğrenciler, uygulama kapsamında yer alan bütün maddeleri cevaplamazlar. PISA 2009'da kullanılan üniteler, 13 madde demeti halinde gruplandırılmıştır. Her bir madde demetinin yanıtlanma süresi 30 dakikadır. PISA 2009'da yedi Okuma Becerileri, üç Matematik ve üç de Fen madde demeti bulunmaktadır. Bu madde demetleri belirli bir döngü düzenine göre 13 kitapçığa yerleştirilmiştir. Her bir kitapçıkta dört madde demeti yer alır ve her öğrenci seçkisiz yöntemle belirlenen 13 kitapçıktan birini cevaplar. Ayrıca her bir kitapçıkta en az bir okuma becerileri madde demeti yer almaktadır. Değerlendirme kapsamındaki tüm maddelerin toplam yanıtlanma süresi 390 dakikadır. PISA 2009 başarı testinde kullanılan bütün maddeler dikkate alındığında, bütün kitapçıklarda okuma becerilerinin değerlendirilmesine ayrılan toplam süre 210 dakika (Toplam test süresinin %54'ü), Matematiğe ve Fen bilimlerine ayrılan süre ise doksanar dakikadır (Toplam test

süresinin %23'ü) ve her bir öğrencinin değerlendirilme süresi 120 dakikadır. Matematik testinde bulunan maddelerin ölçtüğü alt boyutlar ile ilgili bilgi Tablo 2'de verilmiştir.

Tablo 2: Matematik Maddelerinin Öldtüğü Alt Boyutlar

Matematik	
Tanımlama ve Ayırt Edici Özellikleri	Bireyin, dünyada matematiğın oynadığı rolü fark etme ve anlama, sağlam temellere dayanan yargılara ulaşma, yapıcı, ilgili, duyarlı bir vatandaş olarak kendi ihtiyaçlarını karşılayabilecek bir şekilde matematiğı kullanma birikimi ile ilgilidir. <i>Matematik Okuryazarlığı</i> matematiğın daha geniş ve işlevsel kullanımını dikkate almakta ve çeşitli durumlarda matematiksel problemleri tanıma ve formülleştirme becerisini kapsamaktadır.
Bilgi Alanı	İlgili matematiksel alan ve kavram grupları: <i>Nicelik</i> <i>Uzay ve Şekil</i> <i>Değişme ve ilişkiler</i> <i>Belirsizlik</i>
İlgili Beceriler ve Düşünme süreçleri	Matematik için gerekli beceri kümeleri: <i>Yeniden oluşturma</i> (basit matematiksel işlemler) <i>İlişkilendirme</i> (bir problemi çözmek için farklı düşünce ve yöntemleri bir araya getirme) <i>Derinlemesine düşünme</i> (daha kapsamlı matematiksel düşünme)
Bağlam ve Konular	Matematiğın: Kişisel Eğitimle ilgili ve meslekî Kamusal Bilimsel olmak üzere kişisel, sosyal ve küresel ortamlarla ilişkili kullanımları üzerinde yoğunlaşan uygulama alanı.

Ölçme araçlarının tüm ülkelerde geçerli ve güvenilir olmasını sağlamak, kültürel ve dilsel farklılıkları en aza indirmek için oldukça yoğun bir çaba ve kaynak sarf edilmektedir. Çeviri, örneklem oluşturma ve veri toplama işlemlerinde uyulması zorunlu olan kalite güvence mekanizmaları bulunmaktadır. Sonuç olarak PISA'dan elde edilen bulguların yüksek derecede geçerlik ve güvenilirliğe sahip olduğu düşünülmektedir. Bu bulgular ile dünyanın ekonomik olarak en gelişmiş ülkelerinden ve gelişmekte olan ülkelerinden elde edilen eğitim çıktıları daha iyi anlaşılabilen ve karşılaştırılabilir (OECD, 2009). Bu kadar yoğun bir şekilde yapılan geçerlik ve

güvenirlik çalışmalarına rağmen PISA sınavlarında bazı eksikliklerin olduğu görülmektedir.

Ölçme aracında bulunması gereken nitelikler güvenilirlik, geçerlik ve kullanılabilirlik. Geçerlik kavramıyla bir ölçme aracının, onunla ölçülmek istenen değişkeni ölçüp ölçmediği; ölçebiliyorsa onu başka değişkenlerden ne derece arınık olarak ölçtüğü ifade edilir (Turgut, 1997). Ölçme aracından kaynaklanan ve geçerliği etkileyen önemli tehditlerden birisi de madde yanlılığıdır (Clauser ve Mazor, 1998).

Genel anlamdaki tanımıyla madde yanlılığı; testin şartlarından ya da testteki maddelerin bazı özelliklerinden dolayı, maddeyi belirli bir grubun doğru yanıt olma olasılığının diğer gruba göre daha az ya da çok olmasıdır (Mantel ve Haenszel, 1959; Cleary, 1968; Camilli ve Shepard, 1994; Zumbo, 1999; Penfield ve Camilli, 2007). Madde yanlılığının belirlenmesi, test geçerliğini ve güvenilirliğini artırmak için yapılan önemli çalışmalardan biridir. Madde yanlılığı çalışmaları madde fonksiyonunun aynı yetenek düzeyinde ve farklı gruplarda yer alan bireyler için farklılaşıp farklılaşmadığını belirleyen istatistiksel bir süreçle başlar. Aynı yetenek düzeyindeki alt gruplarda fonksiyonu farklılaşan maddeler, değişen madde fonksiyonu (DMF) gösteren maddeler olarak adlandırılır (Dorans ve Holland, 1993).

DMF içeren maddelerin belirli bir gruba veya gruplara avantaj sağladığı göz önüne alındığında DMF'li maddeler ölçme sonucunda sistematik hataya sebep olmaktadır. Bu bağlamda bir testte DMF içeren maddelerin bulunması o testin geçerli olmasını negatif olarak etkileyecektir. Ölçme sonuçları, sınava giren farklı gruplara karşı yansızlık özelliğini sağlamalıdır. Madde yansızlığı sadece teknik olarak geçerlikle yakından ilgili değil, aynı zamanda politik, felsefi, ekonomik, sosyal ve yasal boyutları da olan bir sorundur (Clauser ve Mazor, 1998). Yapılan çalışmalara göre PISA sınavlarındaki bazı maddelerin yanlı olduğu tespit edilmiştir (Ayan, 2011; Yıldırım 2006; Çet, 2006; Le, 2009).

Bu çalışmada, PISA 2009 Matematik okuryazarlığı alttestindeki maddeler için, Değişen Madde Fonksiyonu analizi yapılmıştır. Bunun için araştırmanın problem cümlesi ve alt problemleri şu şekilde belirlenmiştir:

1.2 Problem Cümlesi

1. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde cinsiyete göre DMF gösteren madde var mıdır?
2. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde öğrencilerin katıldığı ülkelere göre DMF gösteren madde var mıdır?

1.3 Alt problemler

1. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde cinsiyet açısından Mantel-Haenszel yöntemine göre DMF gösteren madde var mıdır?
2. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde cinsiyet açısından SIBTEST yöntemine göre DMF gösteren madde var mıdır?
3. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde cinsiyet açısından Lojistik Regresyon yöntemine göre DMF gösteren madde var mıdır?
4. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde uygulandığı ülke açısından Mantel-Haenszel yöntemine göre DMF gösteren madde var mıdır?
5. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde uygulandığı ülke açısından Lojistik Regresyon analizi yöntemine göre DMF gösteren madde var mıdır?
6. PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde uygulandığı ülke açısından SIBTEST yöntemine göre DMF gösteren madde var mıdır?

1.4 Araştırmanın Amacı

PISA, TIMSS, PIRLS vb. uluslararası yapılan uygulamalarda bulunan maddelerin tüm öğrencilere aynı şekilde hitap etmesi gerekmektedir. Ancak yapılan sınavların büyük kitlelere uygulanması maddelerin işlevsel özelliklerini

sınırlandırmaktadır. PISA 2009'a önce 33'ü OECD üyesi olmak üzere toplam 65 ülke katılmış, daha sonra 9 ülkenin katılımı ile bir ek çalışma yapılmıştır (OECD, 2009). Testler veya anketler bir dilden başka bir dile uyarlandığında farklı dillerdeki formlarının eşdeğerliğinin sağlandığından emin olunmalıdır. Uyarlanan veya çevirisi yapılan her bir form farklı bir yapıyı ölçebilir, zorluk ya da aşinalık dereceleri bakımından farklılık gösterebilir (Asil ve Gelbal, 2012). Ayrıca uygulanacak ölçeklerdeki veya anketlerdeki maddeler öğrencilerin bireysel özelliklerini de dikkate alarak hazırlanmalıdır. Bu çalışmada PISA 2009 uygulaması Matematik okuryazarlığı alttestinden bulunan maddelerin DMF analizleri yapılmıştır. Bu çalışma ile teset uygulayıcıların testlerin geçerliğini artırmak için kullanabilecekleri bilgiler sağlanması ve uygulanan yöntemlerin avantajlı ve dezavantajlı yönlerinin belirlenmesi amaçlanmıştır.

1.5 Araştırmanın Önemi

Yapılmış olan çalışmalar incelendiğinde tüm Değişen Madde Fonksiyonu (DMF) belirleme yöntemlerinin eksik ya da zayıf yönleri bulunmaktadır. Uygun bir DMF belirleme yöntemi kullanılarak DMF'nin belirlenmesi yansız ve geçerli bir test oluşturmada önemli bir role sahiptir. Bu nedenle büyük ölçekli sınavlarda yanlış maddelerden kaçınmak ve sınavlardan elde edilen test puanlarına dayalı olarak doğru kararlar almak için hangi koşullarda hangi yöntemin kullanılmasının doğru sonuçlar vereceği önemlidir. Madde yanlılığı geçerliği tehdit eden en büyük unsurlardan biridir. Dünya çapında öğrencileri ölçmeyi amaçlayan uluslararası sınavlarda öğrencilerin gerçek düzeylerinin belirlenebilmesi için bu ölçme araçlarının geçerli ve güvenilir olması gerekmektedir. Bu doğrultuda yapılan bu çalışmayla madde yanlılığı belirlemenin ilk aşaması olan farklı türde DMF belirleme yöntemleri ile PISA 2009 Matematik okuryazarlığı alttestindeki DMF içeren maddeler belirlenecektir.

1.6 Sınırlılıklar

Bu araştırmada DMF analizi, Klasik Test Kuramına dayalı SIBTEST, Mantel-Haenszel ve Lojistik Regresyon yöntemleri ile sınırlı tutulmuştur.

Arařtırmada, PISA uygulamasına katılan her ğrencinin aynı soruyu görmemesi durumundan kaynaklı, DMF analizine alınan maddeler bir kitapçıkla sınırlı tutulmuřtur.

1.7 Sayıtlar

Öğrencilerin matematik okuryazarlığı alt testindeki maddelere verdikleri cevapların gerçek durumlarını yansıttığı varsayılmıřtır.

1.8 Tanımlar

Çalıřmanın bu bölümünde çalıřma sırasına sıklıkla kullanılan bazı kavramların ne anlama geldiđi üzerinde durulmuřtur.

Deđişen Madde Fonksiyonu (DMF): Bir maddenin farklı gruplar için farklı işlev göstermesidir.

Yanlılık: Testte bulunan DMF içerdiği tespit edilen maddelerin belirli bir gruba avantaj sağlayıp sağlamadığının uzmanlar tarafından verilmiş karardır.

Geçerlik: Bir ölçme aracının, onunla ölçülmek istenen deđişkeni ölçüp ölçmediđi; ölçbiliyorsa onu başka deđişkenlerden ne derece arınık olarak ölçtüğüdür.

BÖLÜM II

2 Kuramsal Temeller ve İlgili Literatür

Son yıllarda ölçme ve değerlendirme alanındaki geçerlik ile ilgili tanımlamada yeniden bir oluşum söz konusudur. Geçerlik, klasik tanımı ile bir testin sadece o testle ölçülmek istenilen değişkeni ölçmesi, başka değişkenlerle karıştırmamasıdır. Zumbo (1999) geleneksel geçerlik kavramının;

- ölçülmek istenilen değişkeni ölçmesi,
- güvenilirliğin geçerlik için gerekli olduğunun ancak yeterli olmadığı,
- geçerliğin ölçme aracında bulunması gereken nitelik gibi konuların üzerine yoğunlaştığını belirtmiştir.

Geleneksel anlamda dört geçerlik türünden bahsedilmektedir. Hubley ve Zumbo (1996) araştırmacıların ve test uygulayıcılarının yaptıkları çalışmalarda dört geçerlik türünden sadece birinin gösterilmesinin yeterli olacağını düşündüklerini belirtmiştir.

Geçerlik kavramındaki güncel tanımlamalar ise;

- yapı geçerliliği, geçerlik türlerinden en merkezi geçerlik türü,
- geçerlik bir süreçtir, geçerli veya geçerli olmayan olarak sınıflandırılmaması,
- testin kullanımı ve amacı geçerliğin gerekliliği üzerine yoğunlaştığı belirtilmiştir (Zumbo, 1999).

Geçerlik için bir ölçüt de yansızlıktır (Mellor, 1995). Yansızlık, bir anlamda testle ölçülen özelliklere sahip bir cevaplayıcının testten yüksek puan alması olarak düşünülebilir. Bazı durumlarda bireylerin ölçmek istediğimiz özelliklerine başka değişkenler karışabilir. Bunlar arasında; cinsiyet, okul türü, sosyo-ekonomik düzey, etnik köken v.b. yer alır. Bu değişkenlerin test puanlarına etkisi geçerlik üzerinde bir tehdittir ve test puanlarının yanlı olmasına yol açar. Yanlılık, ölçme sonuçları üzerinde belirli bir gruba karşı sistematik hata anlamına gelir (Atalay, 2010; Camilli, 2006; Camilli ve Shepard, 1994).

Yanlılık genel anlamda test yanlılığı ve madde yanlılığı olmak üzere iki kategori de incelenir. Test yanlılığı testteki maddelerin yanlı olması durumunda ortaya çıkmaktadır. Farklı kültür veya dilden bireyleri karşılaştırmak amacıyla yapılan her bir çalışmada, karşılaştırmanın anlamlı olabilmesi için testlerin ölçtüğü yapıların eşdeğer olması gerekir. Bu tür çalışmalarda ölçme eşdeğerliğinin diğer bir ifadeyle ölçme değişmezliğinin sağlanması temel sayılıdır (Gierl, 2000; Asil ve Gelbal, 2012). Ölçme değişmezliği madde bazında sağlanamıyorsa ‘Değişen Madde Fonksiyonu’ (DMF), test bazında sağlanamıyorsa ‘Değişen Test Fonksiyonu’ (DTF) olarak isimlendirilir. Farklı gruplara ait test puanlarının karşılaştırılabilirliği, yani eşdeğerliği genellikle DMF analizleriyle değerlendirilmektedir (Ercikan, Gierl, McCreith, Puhan, ve Koh, 2004; Asil ve Gelbal, 2012).

2.1 Madde Yanlılığı

Gruplar arasında ortaya çıkan farklılık sistematik yanlılıktan mı kaynaklanmaktadır? Yoksa test sonuçlarındaki farklılıklar bireylerin gerçek tutum ve başarılarını mı yansıtmaktadır? Araştırmacıları düşündüren bu sorular ölçeklerin/testlerin geçerliklerinin belirlenmesinde önemli rol oynamaktadır. Madde yanlılığı belirleme çalışmalarındaki amaç gruplar arasındaki farklılığın gerçek (testten alınan puanların doğru performans değerlerini göstermesi) veya gerçek olmayan (ölçme işleminin oluşumundan kaynaklanan) sonuçların hangisinden kaynaklandığını belirlemektir (Camilli ve Shepard, 1994).

DMF maddelerin iki farklı grup için nasıl işleyiş gösterdiğini belirlemek için yol göstermektedir. İki grup arasındaki farklılığın grupların özelliklerinden mi yoksa maddelerin işleyişinden mi kaynaklandığını belirlemek ölçme ve değerlendirme için ayırt edilmesi gereken bir problemdir (Dorans ve Holland, 1993). DMF' nin ortaya çıkmasının iki temel nedeni olabilir. Bunlar alt gruplar arasındaki gerçek farklılığı gösteren maddenin gerçek etkisi ve madde yanlılığıdır (Camilli ve Shepard, 1994).

Madde yanlılığı araştırmaları Alfred Binet'in 1910 yılında düşük sosyoekonomik düzeydeki çocuklar üzerinde yaptığı çalışma ile başladığı söylenebilir. Bu çalışmada Binet bazı maddelerin duyuşsal özelliklerin yerine kültürel özelliklerin etkilerini ölçtüğünü fark etmiştir. Bu araştırma çalışmasını daha geçerli kılabilmek için Binet bu tür maddeleri testten çıkartmıştır (akt. Camilli ve Shepard, 1994).

1912 yılında William Stern, Binet'e benzer olarak gruplar arasındaki farklılıklar üzerine bir çalışma yapmıştır. Stern aynı yetenek düzeyindeki bu gruplar arasındaki farklılığı anlamak için çaba sarf etmiş ve bu amaç için testlerin bazı gruplar için avantajlı olabileceği yorumunu yapmıştır (akt. Camilli ve Shepard, 1994; Angoff 1993).

Binet (1910) ve Stern (1914)' den sonraki araştırmacılar daha çok IQ testi sonuçları ile sosyal sınıflar arasındaki ilişkiye odaklanmalarına rağmen, Eells ve arkadaşları (1951) davranışlar arasında gözlenen farklılıkların sebebi olarak genetik ve çevresel farklılıkları göz önünde bulundurmışlardır. Eells ve arkadaşları ilk olarak gruplar arasında gözlenen farklılıkların davranışların gerçek yansımalarından kaynaklanmayabileceğini ifade etmişlerdir. 1951 yılında yapılan bu araştırma ile Eells ve arkadaşları, Binet (1910) ve Stern (1914)' in yapmış olduğu çalışmaları ileriye taşımış ve modern anlamdaki madde yanlılık çalışmalarını başlatmışlardır (akt. Camilli ve Shepard, 1994).

Yapılan bu araştırmalar madde yanlılığı üzerine çalışma yapan araştırmacılara dayanak noktası olmuş ve maddelerin yanlılıklarını belirleyecek istatistiksel analiz yöntemleri ortaya çıkmaya başlamıştır. 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; Angoff 1993).

Bir testteki maddelerin yanlılığının var olduğunun ve derecesini belirlemek için yapılan ilk istatistiksel analiz çalışmaları arasında Cardall ve Coffman'ın 1964 yılında farklı ırklar üzerine yaptığı çalışmada, gruplar ile testteki maddelerin etkileşimini incelemek amacıyla varyans analizi yöntemini, 1963 yılında uygulanan SAT (Scholastic Aptitude Test) sınavına katılan bireylerin verileri için uygulamıştır. Aynı yöntem Cleary ve Hilton, (1968) ve Angoff ve Sharon, (1974) tarafından da devam ettirilmiştir. Ancak bu çalışma alanda çok fazla geçerli kabul edilmemiştir (akt. Angoff, 1993).

1972 yılında Angoff, Delta-plot adıyla bilinen yöntemi kültürel farklılıkları belirlemek için kullanmıştır. Bu teknik DMF belirlemede madde ile grup arasındaki etkileşimi temele almaktadır. Kısa süre içinde popüler olan bu yöntemin kullanılmasındaki avantajlar içinde mantıksal yapısının ve uygulanabilirliğinin kolaylığı olması bulunmaktadır. Bu yöntem gruplardaki tüm maddelerin eşit olduğu sürece yanlılığın yapay bir kanıtını oluşturmaktadır. Bu problemi çözmek amacıyla Angoff 1982 yılında hatanın kaynağını doğrulamak için değişiklikler yapmıştır (Baghi ve Ferrara, 1989).

Lord' un 1952 yılında yazdığı "Madde Tepki Kuramı" isimli tezinde açıklanan test kuramının kısa süre içerisinde DMF belirleme çalışmalarında uygulanabilir olduğu açıkça anlaşılmıştır. Şu anda da iyi bilindiği gibi kuramın temeli farklı gruplara ait aynı yetenek düzeyindeki bireyler için çizilen madde tepki fonksiyonuna dayanmaktadır. 1960 yılında Rasch tarafından önerilen ve kendi adıyla anılan Rasch modeli, bir başka DMF belirleme yöntemidir. MTK' nın tek parametreliliği olan Rasch modeli tüm maddelerin eşit ayırıcılık gücüne sahip olduğu ve şans eseri doğru cevaplanmadığı varsayımları üzerine kuruludur. Bu tekniğin maddelerdeki yanlılığın tespiti için kullanımı madde güçlüklerindeki farkın değerlendirilmesiyle sınırlıdır (akt. Angoff, 1993).

1979'da Scheuneman DMF belirlemek amacıyla ki-kareye benzeyen bir yöntem yayınlamış ancak bu yöntem daha sonra Baker (1981) tarafından örneklem büyüklüğünden ve örneklem dağılımının bilinmemesinden etkilenmesi nedeniyle eleştirilmiştir. Daha sonrasında tam ki-kare yöntemi olarak bilinen ve düzeltilmiş istatistiği olan model Bishop, Fienberg ve Holland (1975) tarafından bulunmuştur (akt. Camilli ve Shepard, 1994).

Holland (1985), ve Holland ve Thayer (1988), Mantel ve Haenszel (1959) tarafından geliştirilen bir tekniğin DMF araştırmaları için de kullanılabilir şekilde uyarlamasını geliştirmişlerdir (Angoff, 1993). Madde tepki kuramına alternatif olarak birçok olasılık tablosu analizi yöntemi bulunmaktadır. Bu yöntemler MTK' ya dayalı yöntemlerden farklılıklar göstermektedir. Olasılık tablosuna dayalı yöntemlerde bireyler öncelikli olarak gözlenebilen puanlar açısından karşılaştırılmakta daha sonrasında ise farklı gruplarda bulunan bireyler maddeyi doğru/yanlış cevaplandırma durumlarına göre karşılaştırılmaktadır (Clauser ve Mazor, 1998).

Yanlılık çalışmalarında atılması gereken ilk adım, alt grupların maddeleri cevaplama yapıları arasında farklılık olup olmadığının bir başka deyişle DMF içerip içermediğinin belirlenmesine yönelik istatistiksel analizlerinin yapılmasıdır. Bir maddenin yanlı olduğunu söyleyebilmek için o maddenin DMF içermesi gerekmektedir. Ancak DMF içeren her madde için madde yanlıdır denilemez. İkinci adım olarak ise DMF analizi sonucunda elde edilen bulguların belirli bir gruba avantaj sağlayıp sağlamadığı ile ilgili uzman görüşü almaktır (Camilli ve Shepard, 1994; Zumbo, 1999).

2.2 DMF Belirleme Yöntemleri

DMF maddelerin iki farklı grup için nasıl işleyiş gösterdiğini belirlemek için yol göstermektedir. İki grup arasındaki farklılığın grupların özelliklerinden mi yoksa maddelerin işleyişinden mi kaynaklandığını belirlemek ölçme ve değerlendirme için ayırt edilmesi gereken bir problemdir (Dorans ve Holland, 1993).

DMF'nin tek biçimli (uniform) ve tek biçimli olmayan (nonuniform) olmak üzere iki çeşidi vardır. Tek biçimli DMF yetenek düzeyleri ile grup arasında hiçbir etkileşim olmadığı durumlarda, tek biçimli olmayan DMF ise yetenek düzeyleri ile gruplar arasında bir etkileşim olduğu durumlarda ortaya çıkan DMF için kullanılmaktadır (Swaminathan ve Rogers, 1990). Diğer bir deyişle bir madde bütün yetenek düzeylerinde bir gruba avantaj sağlıyorsa tek biçimli DMF, eğer belirli bir yetenek düzeyine kadar bir gruba diğer yetenek düzeylerinde farklı gruba avantaj sağlıyorsa tek biçimli olmayan DMF olarak isimlendirilir.

2.2.1 Madde tepki kuramına Dayalı Yöntemler

Son yıllarda DMF analizi yapmamızı mümkün kılan çok sayıda Madde Tepki Kuramı (MTK)'na dayalı yöntem geliştirilmiştir. MTK'na dayalı yöntemler bir veya birden fazla maddelerin belirlenmesinde ve değerlendirilmesinde daha kullanışlı yöntemlerdir. Hem teorik açıdan hem de uygulama açısından MTK temelli DMF belirleme yöntemleri Klasik Test Kuramına (KTK) dayalı yöntemlere göre daha kapsamlı araştırma fırsatı vermektedir.

Thissen, Steinberg ve Wainer (1993) tarafından geliştirilen Madde Tepki Kuramı Olabilirlik Oranı (MTK-OO) yöntemi Madde Tepki Kuramı (MTK) temelli bir yöntemdir. Dolayısıyla bireylerin eşleştirilmesinde gözlenen puanları değil MTK temelli yetenek kestirimlerini kullanır. MTK-OO analizleri IRTLRDIF yazılımı ile yapılabilmektedir (Thissen, 2001). MTK-OO ile yapılan DMF analizinde ilk olarak iki grup için madde parametrelerinin serbest bırakıldığı bir genelleştirilmiş model kurulur. Ardından madde parametrelerinin iki grup için aynı olarak kısıtlanmasını sağlayan sınırlandırılmış model kurulur. Her iki modelin model veri uyumu için kullanılan $-2\log(\text{olabilirlik oranı})$ değerleri karşılaştırılır. İki modelin uyumları arasındaki fark G^2 istatistiği olarak rapor edilir (Çepni, 2011).

DMF analizleri için BILOG-MG (Zimowski, Muraki, Mislevy ve Bock, 1996) yazılımı MTK temelli bir algoritma sağlamaktadır. Bu algoritmada, denk olmayan gruplar test eşitleme desenine benzer bir biçimde iki ayrı grup için parametre kestirimi

yapılmakta ve aynı ölçek üzerine getirilmektedir. Aynı ölçek üzerine getirilen güçlük parametrelerinin farkı ve bu farka ait standart hata değerleri rapor edilmektedir (du Toit, 2003). Ayarlanmış güçlük farkı (Δb) değerleri standart hatalarına bölünerek bir hipotez testi gerçekleştirilmektedir. Δb değerleri DMF miktarının büyüklüğüne dair bir etki büyüklüğü belirtmektedir. Ancak bu değerlere dair yaygın olarak kullanılan bir sınıflama yoktur. BILOG-MG algoritması, madde ayırıcılıklarının maddeler arasında değişmesine izin vermekte ancak gruplar arasında farklılaşmasına izin vermemektedir. Dolayısıyla, yalnızca tek biçimli DMF belirlemede ve yorumlanmasında kullanılması uygundur (Smith ve Reise, 1998, akt., Çepni 2011).

2.2.2 Klasik test kuramına dayalı yöntemler

2.2.2.1 Çok kategorili testler için kullanılan yöntemler

Ordinal Lojistik Regresyon: Ordinal lojistik regresyon (OLR), Miller ve Spray (1993) tarafından çok kategorili veriler için önerilmiştir. OLR yöntemi iki kategorili ve çok kategorili verilerde DMF belirlemede kullanılabilir. Zumbo (1999) OLR yönetiminde madde cevaplarının bağımlı değişken, grup değişkeninin her birey için toplam puanla beraber grup toplam puan etkileşiminin bağımsız değişken olarak kullanıldığını belirtmiştir. Ölçeğin tek boyutlu olması OLR'nin varsayımdır. Zumbo (1999) tarafından önerilen OLR yaklaşımı, grup etkisi ve alt-üst ölçek puanı etkileşimi için DMF testi sağlar. Bu yöntemde değişkenler modele hiyerarşik olarak sırayla eklenir. Modele ilk eklenen değişken toplam puan, ikinci değişken grup değişkeni (cinsiyet v.b.), son değişken ise toplam puan ile grup etkileşimidir.

Poly-SIBTEST: SIBTEST parametrik olmayan örtük değişken modelidir. Chang, Mazzeo ve Roussos (1996, akt., Atalay, 2010), SIBTEST yönteminin çok kategorili veriler için farklı formunu geliştirmişlerdir. Poly-SIBTEST, SIBTEST yönteminin iki ve çok kategorili veriler için kullanılabilen genel halidir. Poly-SIBTEST örtük değişkenler için tanımlanmasına rağmen, bu yöntem MTK yetenek ve madde parametrelerini kullanmaz. Bu yöntem, tek biçimli DMF'yi belirlemek için tasarlanmıştır (Atalay, 2010).

2.2.2.2 İki kategorili testler için kullanılan yöntemler

2.2.2.2.1 Mantel-Haenszel yöntemi

Çalışma kapsamına alınan bireyler değişken açısından referans ve odak gruplarına ayrılırlar. Referans ve odak grupları yetenek veya yeterlik düzeylerine göre (genel olarak 4 veya 5 olmak üzere) tabakalara ayrılırlar (Osterlind ve Everson, 2009). Mantel ve Haenszel tarafından 1959 yılında oluşturulan olasılık tablosu tabakalara ayrılmış grupların karşılaştırılmasını kolaylaştırmıştır. Holland (1985) ve Holland ve Thayer (1988) tarafından yapılan çalışmalar sonucunda ise Mantel ve Haenszel tarafından oluşturulan bu olasılık tablosunun DMF belirleme çalışmalarında da kullanılabilirliğini göstermişlerdir.

Analizin yapılabilmesi için farklı gruplardan bireylerin maddeyi doğru ve yanlış cevaplamalarına göre, oluşturulan yetenek grubu sayısı N olmak üzere, $2 \times 2 \times N$ büyüklüğünde üç boyutlu bir matris oluşturulur. Tablo 3'deki veri düzeni i . yetenek düzeyi için oluşturulmuştur.

Tablo 3: Mantel-Haenszel yöntemi olasılık tablosunda kullanılan veri düzeni

		Geçerli Madde İçin Veriler		
		Doğru	Yanlış	Toplam
GRUP	R	A_i	B_i	N_{Ri}
	O	C_i	D_i	N_{Oi}
	Toplam	N_{Di}	N_{Yi}	T_i

Tablo 3'de T_i olarak gösterilen değer referans (R) ve odak (O) gruplarında maddeye cevap veren bireylerin toplamı; N_{Ri} ve N_{Oi} olarak gösterilen değerler ise ayrı ayrı referans ve odak gruplarında maddeye cevap veren öğrencilerin toplamı; N_{Di} ve N_{Yi} olarak gösterilen değerler referans ve odak gruplarında maddeye doğru ve yanlış cevap veren öğrencilerin toplamını göstermektedir. Burada i , herhangi bir yetenek düzeyini/aralığını göstermektedir (Holland ve Thayer, 1988).

Holland ve Thayer (1988) istatistiksel durumu daha net ifade edebilmek için Tablo 3'teki A_i değeri N_{Ri} 'ye ve C_i değeri N_{Oi} 'ye bölünerek maddenin doğru cevaplanma olasılıklarını; B_i değeri N_{Ri} 'ye ve D_i değeri N_{Oi} 'ye bölünerek maddenin doğru cevaplanmama olasılıklarını hesaplamıştır.

Tablo 4: Gruplara göre düzenlenmiş madde güçlük indeksleri veri düzeni
Geçerli Madde İçin Veriler

	Grup	Madde Güçlük İndeksi	Maddeyi Doğru Yanıtlamama Oranı	Toplam
GRUP	R	p_{Ri}	q_{Oi}	1
	O	p_{Oi}	q_{Oi}	1

Holland ve Thayer (1988)'in kurmuş olduğu hipoteze göre; bütün yetenek grupları için eğer referans grubundan elde edilen madde güçlük indeksi (p_{Ri}), odak grubundan elde edilen madde güçlük indeksine (p_{Oi}) eşit ise ($p_{Ri}=p_{Oi}$ ise) o madde için DMF olmadığı söylenebilir (Holland ve Thayer, 1988).

Odds-Oranı bir olayın bir grup için olasılıklar oranının başka bir grup için olasılıklar oranına olan oranı veya bunun bir örnekleme dayalı bir tahmini olarak tanımlanmaktadır. Mantel-Haenszel yöntemi referans ve odak grupları için Odds-Oranının bir kestiricisi olarak α kestirim katsayısını hesaplamaktadır. i yetenek düzeylerini göstermek üzere, α kestirim katsayısı referans ve odak grubundaki her bir yetenek düzeylerindeki doğru ve yanlış yapan öğrenci sayılarına göre hesaplanır. MH istatistiği sonucunda Tablo 3'teki bilgiler kullanılarak hesaplanan α kestirim katsayısı aşağıdaki gibidir.

$$\alpha_{MH} = \frac{\sum A_i D_i / T_i}{\sum B_i C_i / T_i}$$

MH katsayısının daha kolay yorumlanabilmesi için logaritmik bir dönüşüm ile Δ_{MH} katsayısı elde edilmiştir. Bu katsayı ise; $\Delta_{MH} = -2,35 \ln(\alpha_{MH})$ şeklinde ifade edilmektedir. Oluşturulan bu yeni denklem için Δ_{MH} değerinin 0'a eşit olması DMF'nin olmadığını göstermektedir. Eğer Δ_{MH} değeri pozitif değer almışsa odak grubu için, negatif değer almışsa referans grubuna avantaj sağlayacak DMF bulunduğu tespit edilmiş olur (Holland ve Thayer, 1988; Holland ve Wainer, 1993; Osterlind ve Everson, 2009). Zieky (1993) DMF içeren maddelerin düzeylerini belirlemek için bir sınıflama sistemi düzenlemiştir. Bu sınıflama sistemi ise Tablo 5'te gösterilmiştir.

Tablo 5: Zieky (1993) tarafından düzenlenen sınıflama sistemi

DMF düzeyi	Açıklama	Değer Aralığı
A	DMF yok veya ihmal edilebilir.	$ \Delta_{MH} < 1,0$
B	Orta Düzeyde DMF vardır.	$1,0 \leq \Delta_{MH} < 1,5$
C	Yüksek Düzeyde DMF vardır.	$ \Delta_{MH} \geq 1,5$

Elde edilen Δ_{MH} değeri 1,0'dan küçük ise maddede DMF'nin olmadığı veya ihmal edilebileceği, 1,0 ile 1,5 arasında ise maddede B düzeyinde DMF bulunduğu ve 1,5'den büyük ise C düzeyinde DMF bulunduğunu belirtmişlerdir.

Lojistik regresyon

Mantel-Haenszel yöntemi pratik ve kullanımının kolay olmasına rağmen Tek Biçimli ve Tek Biçimli olmayan DMF'leri belirleme konusunda yetersiz kalmaktadır. Swaminathan ve Rogers 1990 yılında yaptıkları çalışmada Mantel-Haenszel ve MTK'ya dayalı DMF belirleme yöntemlerine alternatif ve daha etkili olacak bir yöntem geliştirmişlerdir. Bu yöntem Lojistik Regresyon modeli üzerinden analiz yapmayı gerektirmektedir. Bu yöntem ile analizleri yapılan maddelerin hangi tür DMF içerdiklerine ilişkin bilgi de elde edilmiş olur.

İki kategorili bağımsız değişken için oluşturulan lojistik regresyon modelinde toplam puan (X), ait olduğu grup (G) ve ait olduğu grup ile toplam puanın (X ile G) etkileşimi incelenmektedir. Bireyin maddeyi doğru veya yanlış cevaplamış olmasının asıl sebebinin bu değişkenlerden hangisinden kaynaklandığı belirlenmektedir (Penfield ve Camilli, 2007). Lojistik regresyon modeli;

$$Y = \beta_0 + \beta_1 X_i + \beta_2 G_i + \beta_3 G_i X_i \text{ olmak üzere;}$$

$$P(Y = 1|X, G) = \frac{\exp(Y)}{1 + \exp(Y)}$$

ile ifade edilmektedir. β_0 değeri kesim noktası, β_1 değeri toplam puanın maddenin işleyişi üzerine etkisi, β_2 değeri ait olduğu grubun madde işleyişi üzerine etkisi ve β_3 değeri ait olduğu grup ile toplam puanın etkileşiminin madde işleyişi üzerine etkisini ifade etmektedir. Eğer β_2 ve β_3 0'a eşit ise analizi yapılan madde için DMF olmadığı, sadece β_3 değeri 0'a eşit ise analizi yapılan madde için tek biçimli DMF olduğu, β değerlerinden hiçbiri 0'a eşit değilse analizi yapılan madde için tek biçimli olmayan DMF olduğunu gösterir (Penfield ve Camilli, 2007).

Lojistik regresyon analizinde madde bağımlı değişken, gruplar ve toplam puan bağımsız değişkenlerdir. DMF belirleme çalışmalarında değişkenler modele hiyerarşik olarak ilave edilirler. SPSS programında lojistik regresyon analizi işlemi yapılırken Model 1'de toplam puan, Model 2'de ise grup değişkeni dâhil edilir. Model 3'te ise toplam puan ile grup değişkenleri birlikte ilave edilmektedir. SPSS paket programı ile elde edilen Nagelkerke R^2 değerleri maddelerde araştırması yapılan DMF için yol gösterici olmaktadır. Model 1 ile Model 2 arasındaki R^2 ve Model 2 ile Model 3 arasındaki değerlerinin farkı (ΔR^2) DMF miktarının yorumlanmasında kullanılan değerdir. DMF düzeylerini belirlemek için Zumbo ve Thomas (1996) ve Jodoin ve Gierl (2001) tarafından yapılmış olan sınıflamalar şu şekildedir:

Tablo 6: Zumbo ve Thomas (1996) ve Jodoin ve Gierl (2001) tarafından düzenlenen sınıflama sistemleri

DMF Düzeyi	Açıklama	Zumbo ve Thomas (1996)'a göre değer aralığı	Jodoin ve Gierl (2001)'e göre değer aralığı
A	DMF yok veya ihmal edilebilir.	$\Delta R^2 < 0,13$	$\Delta R^2 < 0,035$
B	Orta Düzeyde DMF vardır.	$0,13 \leq \Delta R^2 < 0,26$	$0,035 \leq \Delta R^2 < 0,070$
C	Yüksek Düzeyde DMF vardır.	$0,26 \leq \Delta R^2$	$0,070 \leq \Delta R^2$

Model 1 ile Model 2 den elde edilen değerlerin farkının anlamlı olması DMF türünün tek biçimli (uniform) olduğunu, Model 2 ile Model 3 arasındaki farkın anlamlı olması ise DMF türünün tek biçimli olmayan (nonuniform) olduğunu göstermektedir. Bakan Kalaycıoğlu ve Kelecioğlu (2011)'nin yapmış oldukları çalışmada yöntemlerin örneklem büyüklüklerinden etkileniyor olması ve farklı sınıflama sistemlerinde ortak bir görüş sağlanamamasından dolayı daha hassas bir sınıflama sistemi geliştirmişlerdir. Bu sınıflama sistemi Tablo 7'de verilmiştir.

Tablo 7: Bakan Kalaycıoğlu ve Kelecioğlu (2011) tarafından düzenlenen sınıflama sistemi

DMF Düzeyi	Açıklama	Bakan Kalaycıoğlu ve Kelecioğlu (2011)'e göre değer aralığı
A	DMF yok veya ihmal edilebilir.	$\Delta R^2 < 0,010$
B	Orta Düzeyde DMF vardır.	$0,010 \leq \Delta R^2 < 0,020$
C	Yüksek Düzeyde DMF vardır.	$0,020 \leq \Delta R^2$

Elde edilen ΔR^2 değeri 0,010'dan küçük ise maddede DMF' nin olmadığı veya ihmal edilebileceği, 0,010 ile 0,020 arasında ise maddede B düzeyinde DMF bulunduğu ve 0,020'den büyük ise C düzeyinde DMF bulunduğunu belirtmişlerdir.

2.2.2.2.2 SIBTEST (Simultaneous Item Bias Test)

Shealy ve Stout tarafından 1993 yılında geliştirilen SIBTEST yöntemi Mantel-Haenszel ve lojistik regresyon yöntemlerinden farklı olarak ile madde yanlılığı ve test yanlılığı aynı anda belirlenebilmektedir. Lojistik regresyon yönteminde olduğu gibi SIBTEST yönteminde de tek biçimli ve tek biçimli olmayan DMF belirlemede etkilidir. Ayrıca bu yönteme göre DMF belirlemesi maddelerin etkilerinden bağımsız olarak yapılmaktadır. Bunun için öncelikli olarak tüm maddeler kullanılarak karşılaştırılma yapılır. Bu karşılaştırma sonucunda maddeler DMF içerdiği belirlenen maddeler ve diğerleri olmak üzere iki ayrı gruba ayrılırlar. Maddelerin son olarak karşılaştırılması DMF içermeyen maddelerden elde edilen toplam puanlar üzerinden yapılmaktadır. SIBTEST yöntemi I. tip hatayı kontrol etmede kullanışlı regresyon temeline dayanmaktadır (Clauser ve Mazor, 1998). SIBTEST yöntemi Madde Tepki Kuramına dayalı olmasına rağmen madde kalibrasyonunun ve eşleştirme değişkeninin gizil olduğu bir yöntemdir (Atalay, Gök, Kelecioğlu ve Arslan, 2012).

Shealy ve Stout tarafından geliştirilen SIBTEST yönteminin temelleri Dorans ve Kulick (1986) tarafından oluşturulan ve “Standartlaşma” diye bilinen yönteme dayanmaktadır. $B(D)$ referans ve odak grupları arasında maddeyi doğru cevaplama durumları açısından farklılığı, $P_R(D)$ ve $P_F(D)$ sırasıyla referans ve odak gruplarındaki bireyin maddeyi doğru cevaplama olasılıkları olmak üzere;

$$B(D) = P_R(D) - P_F(D)$$

Toplam DMF, β , $B(D)$ ' nin beklenen değeri olmak üzere SIBTEST'e ait yokluk hipotezi, diğer bir deyişle maddede DMF olmadığı gösterimi;

$$\text{Exp}[B(D)] \equiv \beta = 0$$

ile tanımlanmaktadır. SIBTEST analizi sonuçları, elde edilen β 'nin aldığı değerlere göre yorumlanmaktadır. Roussos ve Stout (1996) tarafından hazırlanan sınıflama Tablo 8' de verilmiştir.

Tablo 8: Roussos ve Stout (1996) tarafından düzenlenen sınıflama sistemi

DMF düzeyi	Açıklama	Değer Aralığı
A	DMF yok veya ihmal edilebilir.	$\beta < 0,059$
B	Orta Düzeyde DMF vardır.	$0,059 \leq \beta < 0,088$
C	Yüksek Düzeyde DMF vardır.	$0,088 \leq \beta$

Elde edilen β değeri 0,059'dan küçük ise maddede DMF' nin olmadığı veya ihmal edilebileceği, 0,059 ile 0,088 arasında ise maddede B düzeyinde DMF bulunduğu ve 0,088'den büyük ise C düzeyinde DMF bulunduğunu belirtmişlerdir.

2.3 İlgili Araştırmalar

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.

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 incelenen 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 Bilimleri 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.

Çet (2006) tarafından yapılan çalışmada PISA 2003 Matematik sorularının Türkçe ve İngilizce formları arasındaki madde yanlılığını araştırılmıştır. Bunun için yayınlanmış madde sayısının çoğunlukta olduğu iki kitapçık seçilmiştir. Bu çalışmada testlerin faktör yapıları faktör çözümlemesi yöntemleri ile tespit edildikten sonra seçilen maddeler analiz edilmiştir. Yanlı çalıştığı tespit edilen maddelerde ölçtükleri matematiksel beceriler ve bilişsel yeterlilikler, madde türü, madde kökü ve diğer görsel ve uzamsal unsurlar dikkate alınarak madde yanlılığının kaynağının tespit edilmesi için niteliksel analizler yapılmıştır. Bu analizler müfredat farklılıkları, kültürel farklılıklar ve çeviriden kaynaklanan farklılıklar olarak üç ana başlık altında yapılmıştır. Türkiye ve ABD'daki öğrencilerden eşit yeteneklerde olanların niye bazı maddelere doğru cevap verme olasılıklarının farklı olduğu araştırıldığında bunun matematik programlarının farklılığından kaynaklanabileceği ya da İngilizce 'den Türkçe 'ye çeviri yapılırken matematik maddelerindeki bazı nicelik bildiren kelimelerin anlamlarının değişebileceği görülmüştür.

Yıldırım (2006) TIMSS-1999 ve PISA 2003 uluslararası sınavlarının matematik başarı testlerinin kültürler arası denkliliğini, ve matematik başarısı kavramının altında kültürler özgün yapılar olup olmadığını araştırmıştır. Bu amaçla, TIMSS-1999 ve PISA 2003 Türkçe ve İngilizce versiyonlarındaki matematik başarı testi maddeleri, (a) farklı dildeki test maddelerinin psikometrik özellikleri, ve (b) maddelerin farklı dil testler arasında yanlı çalışmasının muhtemel sebepleri açısından değerlendirilmiştir. Yanlı çalışan maddelerin tespitinde, Sınırlandırılmış Faktör Çözümlemeleri, Mantel-Haenszel Yöntemi ve Madde Tepki Kuramı En Çok Olabilirlik Oran Analizi yöntemleri kullanılmıştır. PISA çalışmasında, madde yanlılığı tespitinde kullanılan farklı yöntemlerin sonuçları arasında yüksek derecede uyum tespit edilmiştir. Ancak, TIMSS çalışmasında, yanlı çalışan madde sayısının fazla olması, maddelerin tahmin ve ayırt edicilik indislerindeki farklılıklar bu uyumu kayda değer ölçüde bozmuştur.

Yıldırım (2008) tarafından yapılan çalışmada farklı işleyen maddelerin belirlenmesinde, Mantel-Haenszel (M-H), Olabilirlik Oranı Analizi (OOA), ve Sınırlandırılmış Faktör Çözümlemesi (SFC) yöntemlerini karşılaştırmıştır. Söz konusu çalışmada PISA 2003 matematik maddeleri kullanılmış, İngilizce ve Türkçe formları arasında yanlı çalışma potansiyeli olan sorular tespit edilmiştir. SFC yönteminin karşılaştırılan grup ortalamaları farklı ya da eşit olduğu durumlarda M-H ve OOA yöntemlerine göre, farklı işleyen maddeleri belirleme oranına göre daha doğru sonuçlar verdiği görülmüştür.

Doğan ve Öğretmen (2008) tarafından yapılan çalışmada DMF belirleme tekniklerinden Ki-Kare, Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırarak uygulamada ortaya çıkan benzerlik veya farklılıklarını belirlemişlerdir. Çalışma, 2003 yılında Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı'na (OKÖSYS) katılan yaklaşık 600.000 öğrenci arasından yansız olarak seçilen 3345 öğrenciden oluşan bir örneklem üzerinde yürütülmüştür. Çalışmanın verileri, OKÖSYS fen bilgisi alt testine öğrencilerin verdiği yanıtlardan oluşturulmuştur. DMF analizi sadece cinsiyet grupları üzerinde yürütülmüştür. Araştırmanın sonuçları, söz konusu tekniklerin bazı açılardan birbirlerine göre benzerlik ve farklılıklar sağladığını göstermiştir.

Le (2009), PISA 2006 Fen Bilgisi okuryazarlığı 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 ABD örnekleminde madde tipi ve konu alanları sınıflarında cinsiyet bakımından farklılaşmasını araştırmışlardır. Sonuç olarak Matematik testindeki maddelerin genelde erkekler yönüne avantaj sağladığı görülmüştür. 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 logaritmik doğrusal model MH ve LR yöntemlerini kıyaslamıştır. Araştırma sonucunda MH yönteminden elde edilen sonuçların diğer yöntemlere göre elde edilenlerden daha karmaşık sonuçlar ürettiği gözlemlenmiş, ayrıca üç model arasında elde edilen korelasyon değerlerinin de yüksek olduğu belirtilmiştir. LR ve MH yöntemleri arasında elde edilen korelasyon değeri 0.75 olarak bulunmuştur.

Atalay (2010) çalışmasında PISA 2006 öğrenci anketinde yer alan Fen Bilimlerine karşı tutum maddeleri, ABD ve Türkiye örneklemelerinde cinsiyete karşı ve bu iki ülke arasında kültüre karşı DMF gösterip göstermediğini incelemiştir. Türkiye verisinde cinsiyete göre yapılan analizlerde Ordinal Lojistik Regresyon yöntemi ile 12 maddede, poly-SIBTEST yöntemi ile 13 maddede DMF gözlenmiştir. ABD verisinde cinsiyete göre yapılan analizlerde ise her iki yöntemde 15 maddede DMF gözlenmiştir. Her iki yöntemde ortak ve önemli düzeyde DMF gösteren madde sayısı Türkiye 'de 7, ABD 'de 13'tür. İki ülke arasında kültüre göre her iki yöntemde yapılan analizlerde 30 maddede önemli düzeyde DMF görülmüştür.

Gök, Kelecioğlu ve Doğan (2010) DMF belirlemek için Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) tekniklerini kullanmışlardır. Bu amaçla araştırmada, 2005 yılında yapılan Ortaöğretim Kurumları Sınavı (OKS)'nin matematik ve fen bilgisi alt testleri kullanılmıştır. Test maddelerinin cinsiyet ve okul türüne göre fonksiyonlarının farklılaşıp farklılaşmadığı incelenmiştir. Elde edilen bulgular doğrultusunda, MH ve LR teknikleri arasında matematik ve fen bilgisi alt testlerinde cinsiyet ve okul türü değişkenlerine göre genel olarak düşük düzeyde bir uyum görülmüştür.

Ayan (2011) yapmış olduğu çalışmada, Uluslararası Öğrenci Değerlendirme Programı (PISA) 2009 fen okuryazarlığı alt testinde yer alan maddelerle, cinsiyet değişkenine göre, DMF analizi yürütülmüştür. DMF analizinde lojistik regresyon (LR)

ve Mantel Haenszel (MH) yöntemleri kullanılmış ve ardından bu iki teknikten elde edilen sonuçlar karşılaştırılmıştır. Cinsiyet değişkenine göre yürütülen DMF analizleri sonucunda LR tekniğinde 1 adet orta düzeyde Tek Biçimli Olmayan 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. Farklı yöntemlerle tespit edilen DMF düzeyleri bakımından düşük düzeyde uyum belirlenirken, belirledikleri maddeler bakımından ise tamamen farklı sonuçlar ürettikleri tespit edilmiştir.

Çepni (2011) Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı Sayısal Yetenek testlerinde cinsiyete ve adayların mezun olduğu/olacağı bölüme göre DMF ve değişen madde grubu fonksiyonu (DMGF) analizleri yürütülmüştür. DMF analizleri Mantel Haenszel, lojistik regresyon, SIBTEST, Madde Tepki Kuramı-Olabilirlik Oranı ve BILOG-MG DMF Algoritması yöntemleriyle gerçekleştirilmiştir. DMGF analizleri için SIBTEST yöntemi kullanılmıştır. Çalışmanın veri setleri, sınavın 2008 Sonbahar uygulamasından elde edilmiştir. Cinsiyete göre DMF analizlerinde Sayısal 1 Testinde üç maddenin erkek öğrenciler lehine, dört maddenin ise kız öğrenciler lehine islediği görülmüştür. Sayısal 2 Testinde ise bir madde erkek öğrenciler lehine islerken, üç madde kız öğrenciler lehine işlemiştir. Bölüme göre DMF analizleri için öğrencilerin mezun oldukları bölümler sayısal, sözel ve eşit ağırlık olmak üzere üç gruba toplanmıştır. Bölüme göre DMF analizlerinde Sayısal 1 Testinde en belirgin farkların Sayısal-Sözel bölümler kıyaslamasında ortaya çıktığı görülmüştür. Beş madde sayısal bölüm öğrencileri lehine islerken, dört madde sözel bölüm öğrencileri lehine işlemiştir. Sayısal 2 Testinde üç maddenin sayısal bölüm öğrencileri lehine DMF gösterdiği görülmüştür.

Bakan Kalaycıoğlu ve Kelecioğlu (2011) tarafından yapılan çalışmada 2005 yılı Öğrenci Seçme Sınavı'nda yer alan maddeler cinsiyete göre DMF içerip içermediği incelenmişlerdir. DMF içerdiği belirlenen maddelerin yanlı olup olmadığına karar vermek için uzman görüşüne başvurulmuştur. DMF belirleme yöntemlerinden Mantel-Haenszel ve Lojistik Regresyon kullanılmıştır. Analizler 2005 yılında ÖSS'ye giren 599330 lise son sınıf öğrencisinin her biri 45 maddeden oluşan Türkçe, Sosyal bilimler, Matematik ve Fen Bilimleri alt testlerine verdikleri cevaplar kullanılarak yapılmıştır.

Araştırma sonucunda Türkçe alttestinde sıfır, Sosyal Bilimler alttestinde yedi, matematik ve Fen bilimleri alttestlerinde ise üçer maddede DMF olduğu tespit edilmiştir. Uzman görüşü sonucunda ise bu maddelerden sadece Fen Bilimleri alttestindeki bir maddede yanlışlık olduğu bulunmuştur.

Asil ve Gelbal (2012) tarafından yapılan araştırmada Uluslararası Öğrenci Değerlendirme Programı (PISA) 2006 kapsamında uygulanan öğrenci anketinin kültürler ve diller arası eşdeğerliği, Avustralya, Yeni Zelanda, ABD ve Türkiye örneklemi üzerinde karşılaştırmalı olarak incelenmiştir. Araştırmanın birinci aşamasında, öğrenci anketinin faktör yapısı Doğrulayıcı Faktör Analizi (DFA) yöntemiyle incelenmiştir. İkinci aşamada ise, Ortalama ve Kovaryans Yapıları (OKY) metoduna dayanan Çoklu Grup Doğrulayıcı Faktör Analizi (ÇG-DFA) kullanılarak anket maddelerinin farklı kültürler ve diller arasında DMF gösterip göstermediği araştırılmıştır. Son aşamada, uzman görüşlerine başvurularak, maddelerin DMF göstermesinin nedenleri belirlenmeye çalışılmıştır. DFA sonuçları ölçme modelinin tüm örneklerde aynı faktör yapısına sahip olduğunu göstermiştir. ÇG-DFA bulguları ülkeler arasında değişen fonksiyon gösteren maddelerin olduğunu ortaya koymuştur. Ülkeler arasında dilsel ve kültürel farklılıklar arttıkça DMF gösteren maddelerin sayısının da arttığı gözlenmiştir. Maddelerin DMF göstermesinin asıl nedenlerinin çeviri problemleri ve kültürel farklılıklar olduğu sonucuna varılmıştır.

BÖLÜM III

3 Yöntem

3.1 Araştırma Modeli

Araştırma, PISA 2009 sınavının matematik okuryazarlığı alt testinde yer alan maddelerinin cinsiyete ve kültüre göre DMF içerip içermediğini incelemeye yöneliktir. Araştırma, var olan durumu ortaya koymaya yönelik ve yöntemler arasında kıyaslama yapılacağı için betimsel bir araştırma özelliği taşımaktadır.

3.2 Çalışma Grubu ve Veri Toplama Araçları

PISA 2009 uygulaması Türkiye’de 2009 yılının Nisan ayı içerisinde yapılmıştır. PISA 2009 uygulamasında, 12 istatistiki bölge biriminden 56 il ve okul türlerine göre PISA uluslararası merkez tarafından seçkisiz yöntemle belirlenen toplam 170 okuldan 4996 öğrenci yer almıştır. 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 örnekleme 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 uygulamasına katılan öğrenciler, uygulama kapsamında yer alan bütün maddeleri yanıtlamazlar. 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. Her bir kitapçıkta 4 madde demeti yer alır ve her öğrenci seçkisiz yöntemle belirlenen 13 kitapçıktan birini yanıtlar. Ayrıca her bir kitapçıkta en az 1 okuma becerileri madde demeti yer almaktadır.

Cinsiyete göre yapılan DMF analizlerinde ulusal ve uluslararası testlerde DMF içeren maddelerin bulunduğu görülmüştür. PISA, TIMSS ve PIRLS gibi uluslararası düzeyde yapılan sınavların farklı kültürlerle sahip öğrencilere uygun olacak şekilde hazırlanması gerekmektedir. Aksi takdirde elde edilecek sonuçların öğrencilerin gerçek başarı düzeylerinden mi yoksa öğrencilerin sahip oldukları kültürel özelliklerden mi kaynaklandığı belirlenememektedir. Bu sebeple DMF analizi için bireylerin cinsiyetleri ve kültürleri değişken olarak kullanılmıştır. Kitapçıklarda bulunan matematik okuryazarlığına ait madde sayıları Tablo 9’da verilmiştir.

Tablo 9: Matematik Okuryazarlığına Ait Madde Sayıları
Öğrenci Sayıları

Kitapçık Kodu	Öğrenci Sayıları				Madde Sayısı
	Türkiye	Almanya	Finlandiya	ABD	
1	388	382	454	406	23
2	386	369	446	400	0
3	378	362	453	409	12
4	382	362	449	400	0
5	373	379	446	402	24
6	385	371	447	396	0
7	380	360	438	398	11
8	392	370	440	402	12
9	390	367	438	413	12
10	385	379	455	406	11
11	389	366	444	407	23
12	383	363	454	398	12
13	385	370	446	396	0

Tablo 9’da sunulan kitapçıklara ait kişi ve madde sayıları incelendiğinde kitapçıklar arasında kişi sayılarının büyüklük bakımından birbirlerine yakın olduğu ancak matematik okuryazarlığı alttestinde bulunan madde sayısı bakımından farklılıklar bulunduğu tespit edilmiştir. Bu kapsamda, araştırmada kullanılacak kitapçığın seçiminde kitapçıklarda bulunan matematik maddesi sayısına önem verilmiş ve Tablo 9’da sunulan bilgiler ışığında 24 adet matematik maddesinin bulunduğu 5 nolu kitapçık için analizler yapılmıştır.

PISA’da kullanılan madde türleri ile ilgili olarak; maddelerin yaklaşık yarısı 4 ya da 5 seçenekten oluşan çoktan seçmeli (basit çoktan seçmeli) ya da öğrencilerin bir dizi önermeyi “evet/hayır”, veya “katılıyorum/katılmıyorum” gibi muhtemel iki yanıtın birini seçerek değerlendirdikleri maddelerden (karmaşık çoktan seçmeli) oluşmaktadır. Sadece bir kısmı doğru olan ya da istenenden daha basit bir açıklama yapılan yanıtlara kısmî puan verilmektedir. Ele alınan 5 numaralı kitapçık için 3 adet madde doğru/yanlış diye kodlanmadığı için analiz için kullanılmamıştır.

3.3 Verilerin Çözülmesi

DMF analizleri öncesinde, testlerinden elde edilen verilerinin tek boyutluluğunun sınanması amacıyla, matematik okuryazarlığı alttesti için, doğrulayıcı faktör analizi ile test edilmiştir.

Birinci ve dördüncü alt probleme ilişkin bulguların elde edilmesi amacıyla 5 numaralı kitapçığın matematik okuryazarlığı alttestindeki maddeler için MH yöntemi ile DMF analizi yapılmıştır. Bu analiz için gerekli olan referans ve odak grupları cinsiyet değişkenine göre (sırasıyla kız ve erkek olacak şekilde) oluşturulmuştur. İki gruba ayrılan veri demeti Waller (1998) tarafından geliştirilen EZDIF paket programı analiz edilmiştir. DMF gösteren maddeler için DMF düzeyleri ve hangi grup için avantaj sağladıkları tespit edilmiş ve sonuçlar tablolar halinde verilmiştir. DMF içerdiği tespit edilen maddeler için DMF düzeyleri Zieky (1993) tarafından düzenlenen sınıflama sistemine göre belirlenmiştir.

İkinci ve beşinci alt probleme ilişkin bulguların elde edilmesi amacıyla matematik okuryazarlığı alttestindeki maddeler için Lojistik Regresyon yöntemi ile DMF analizi yapılmıştır. Lojistik regresyonun varsayımlarından olan tek boyutluluk gösterimi için yapılan doğrulayıcı faktör analizi yapılmıştır. Doğrulayıcı faktör analizi için LISREL paket programından faydalanılmıştır. Lojistik regresyon analizinde madde bağımlı değişken, gruplar ve toplam puan bağımsız değişkenler olmak üzere, DMF belirleme çalışmalarında değişkenler modele hiyerarşik olarak (toplam puan, grup ve toplam puan ile grup değişkenleri birlikte) ilave edilmiştir. Lojistik Regresyon analizi yapmak ve bireylere ait toplam puanları hesaplamak için SPSS paket programından yararlanılmıştır. Analiz sonucunda DMF içeren maddelerin düzeyleri Bakan Kalaycıoğlu ve Kelecioğlu (2011) tarafından belirlenen sınıflama sistemine göre belirlenmiştir.

Üçüncü ve altıncı alt probleme ilişkin bulguların elde edilmesi amacıyla testteki maddeler için SIBTEST analizi DMF içeren maddeleri belirlemek için yapılmıştır. SIBTEST analizi SIBTEST (Simultaneous Item Bias TEST) paket programı

ile yapılmaktadır. SIBTEST analizinin yapılabilmesi için verilerin referans (kız) ve odak (erkek) olmak üzere iki gruba ayrılmışlardır. SIBTEST analizi sonuçları elde edilen β 'nin aldığı değerlere ve Roussos ve Stout (1996) tarafından hazırlanan sınıflama sistemine göre yorumlanmaktadır.

3.4 Çalışma Grubuna ait İstatistikler

3.4.1 Betimsel istatistikler

Öğrencilere ait betimsel istatistiklerin analizi için SPSS paket programı kullanılmıştır. Matematik okuryazarlığı alttestine ait betimsel istatistik sonuçları Tablo 10 ve Tablo 11'de verilmiştir.

Tablo 10: Cinsiyete Göre Betimsel İstatistikler

	Kız	Erkek
Öğrenci Sayısı	161	160
Ortalama	8,25	8,44
Standart Sapma	4,31	4,27
Çarpıklık	,419	,242
Basıklık	-,404	-,759
Ortalama Güçlük	0,39	0,40
Cronbach's Alpha	0,820	0,817

Analiz sonuçlarına göre 21 maddeden oluşan testin 161 kız öğrenciden oluşan referans grubu için ortalaması 8,25, 160 öğrenciden oluşan odak grubunun aritmetik ortalaması 8,44'tür. Ortalama güçlük katsayıları referans grubu için 0,39 ve odak grubu için ise 0,40 olarak bulunmuştur.

Tablo 11: Öğrencilerin Katıldıkları Ünelere Göre Betimsel İstatistikler

	Türkiye	Almanya	Finlandiya	ABD
Öğrenci Sayısı	321	351	424	383
Ortalama	8,35	11,68	12,58	9,98
Standart Sapma	4,29	4,77	4,32	4,55
Çarpıklık	0,33	-0,24	-0,36	0,08
Basıklık	-0,60	-0,87	-0,39	-0,75
Ortalama Güçlük	0,40	0,56	0,60	0,48
Cronbach's Alpha	0,82	0,84	0,81	0,83

Ünelere göre yapılan analizlere göre betimsel istatistiksel sonuçlarına bakılacak olursa 0-21 arasında puan alınabilecek matematik testinden 321 öğrenciden oluşan Türkiye örneklemini için ortalama 8,35, standart sapma 4,29; 351 öğrenciden oluşan Almanya örneklemini için ortalama 11,68, standart sapma 4,77; 424 öğrenciden oluşan Finlandiya örneklemini için ortalama 12,58, standart sapma 4,32; 383 öğrenciden oluşan ABD örneklemini için ortalama 9,98, standart sapma 4,55 olduğu görülmektedir. Çarpıklık katsayılarına bakıldığında dağılımın normale yakın olduğu bulunmuştur.

3.4.2 Madde istatistikleri

Referans ve odak gruplarına ait madde güçlük indeksleri ve standart sapma değerleri hesaplanmıştır. Tablo 12'de verilen maddelere ait güçlük indeksleri incelendiğinde referans grubu için 0,07 ile 0,93 aralığında, odak grubu için ise 0,06 ile 0,98 aralığında değişmektedir. Maddelere ait standart sapma değerlerinin birbirine yakın olduğu söylenebilir.

Tablo 12: Maddelere ait güçlük indeksleri

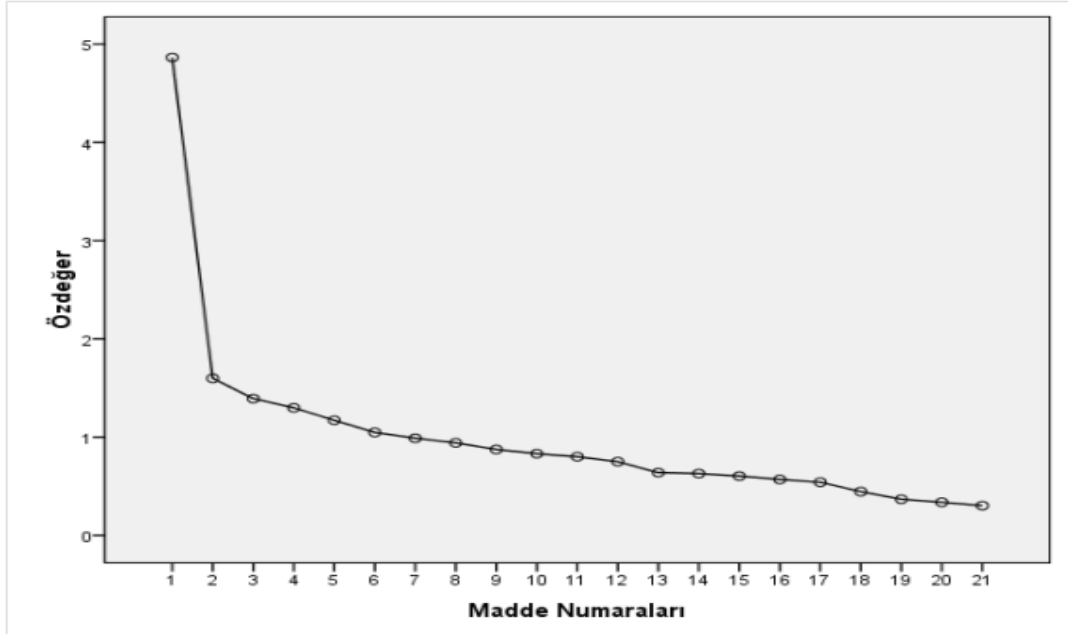
Madde	Ülkeler				Cinsiyet	
	Türkiye	Almanya	Finlandiya	ABD	Kız	Erkek
1	0,56	0,84	0,85	0,74	0,51	0,61
2	0,20	0,48	0,53	0,30	0,19	0,21
3	0,49	0,66	0,72	0,65	0,5	0,47
4	0,39	0,57	0,70	0,52	0,39	0,39
5	0,47	0,68	0,68	0,51	0,43	0,51
6	0,38	0,52	0,56	0,34	0,35	0,41
7	0,19	0,59	0,59	0,48	0,19	0,19
8	0,29	0,52	0,52	0,44	0,3	0,28
9	0,36	0,59	0,52	0,61	0,35	0,38
10	0,16	0,39	0,44	0,30	0,17	0,14
11	0,72	0,66	0,78	0,78	0,76	0,67
12	0,07	0,07	0,09	0,05	0,07	0,06
13	0,65	0,76	0,77	0,63	0,66	0,65
14	0,21	0,38	0,37	0,16	0,23	0,19
15	0,54	0,77	0,84	0,69	0,53	0,55
16	0,51	0,67	0,78	0,58	0,48	0,53
17	0,95	0,88	0,97	0,80	0,93	0,98
18	0,14	0,35	0,36	0,26	0,12	0,16
19	0,34	0,31	0,47	0,30	0,35	0,33
20	0,55	0,58	0,63	0,59	0,53	0,57
21	0,17	0,43	0,42	0,27	0,17	0,17

3.4.3 Tek boyutluluk

Verilerin Tek Boyutluluk varsayımlarını karşılayıp karşılamadıklarını belirlemek amacıyla faktör analizi çalışmaları yapılmıştır. Faktör analizi, birbiriyle ilişkili çok sayıda değişkeni bir araya getirerek az sayıda kavramsal olarak anlamlı yeni değişkenler (faktörler, boyutlar) bulmayı, keşfetmeyi amaçlayan çok değişkenli bir istatistik olarak tanımlanabilir (Büyüköztürk 2002). Testlerden elde edilen verilerin tek boyutluluğunun sınanması amacıyla, matematik okuryazarlığı testi için, tek boyutlu bir

ölçme modeli doğrulayıcı faktör analizi ve açımlayıcı faktör analizi ile test edilmiştir. Doğrulayıcı faktör analizine ilişkin uyum indeksi verileri Tablo 13’de ve tek boyutluluk değerlendirmeleri için özdeğerlere ilişkin çizgi grafiği sonuçlarından (scree plot) yararlanıldığında alt testlere ilişkin çizgi grafikleri Şekil 3’de gösterilmiştir.

Alt testlerin tek boyutluluk değerlendirmeleri için öz değerlere ilişkin çizgi grafiği sonuçlarından (scree plot) yararlanıldığında alt testlere ilişkin çizgi grafikleri Şekil 1’de gösterilmiştir.



Şekil 1: Faktör Analizi Sonuçlarına ait Scree-Plot Grafikği

Matematik alt testinin 1. faktörü açıklama varyansı %22,47 ve 2. faktörü açıklama varyansı %6,61’dir. Tek boyutluluk değerlendirmelerinde bir diğer yöntem çizgi grafiklerindeki 1. ile 2. faktörlerdeki öz değerlerin büyüklüklerinin karşılaştırılmasıdır. 1. faktördeki öz değer miktarının, 2. faktördeki öz değer miktarından üç katı büyük olması ve 1. faktörün toplam varyansın %20’sini açıkladığı durumlarda ölçek tek boyutlu olarak kabul edilmektedir (Lee, 1995; Büyüköztürk, 2011). Yapılan faktör analizi sonucunda 1. faktör toplam varyansın %22,47’sini

açıkladığı görülmüş ve 5 numaralı Matematik Okuryazarlığı testinin tek boyutlu olduğu kabul edilmiştir

Tablo 13: Doğrulayıcı Faktör Analizi Model Veri Uyumu İstatistikleri

Ki-Kare/sd	1,49
RMSEA	0,041
SRMR	0,051
NFI	0,89
NNFI	0,96
CFI	0,96
GFI	0,92
AGFI	0,90

Tablo 13'te verilen doğrulayıcı faktör analizi sonuçlarına göre doğrulayıcı faktör analizi sonucunda modelin birçok uyum iyiliği kriterlerini karşıladığı ancak birkaçında da yeterli düzeyde uyumu karşılamadığı gözlenmiştir [Ki-Kare/sd = 1,49, GFI= 0.90, NFI= 0.89, NNFI= 0.96, CFI= 0.96, AGFI= 0.90, RMSEA= 0,041, SRMR=0,051].

BÖLÜM IV

4 Bulgular ve Yorum

Bu bölümde, araştırmanın alt problemlerine cevap aramak amacıyla toplanan verilerin analizleri sonucunda elde edilen bulgular yer almaktadır.

4.1 Birinci Alt Probleme İlişkin Bulgular

Birinci Alt Problem: PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde cinsiyet açısından Mantel-Haenszel yöntemine göre DMF gösteren madde var mıdır?

PISA 2009 matematik okuryazarlığı alt testinde bulunan maddeler Mantel-Haenszel yöntemi ile analiz edilmiştir. Elde edilen sonuçların tablolaştırılmış hali tablo 14'de verilmiştir.

Tablo 14: Matematik testine ait Mantel-Haenszel yöntemine göre yapılan analiz sonuçları

Madde	Alfa	X^2	p	Δ -MH	SE (MH D-DIF)
1	0.609	3.466	0.063	1.167	0.588
2	0.769	0.338	0.561	0.618	0.817
3	1.176	0.218	0.641	-0.381	0.634
4	0.996	0.012	0.913	0.009	0.581
5	0.691	1.910	0.167	0.867	0.580
6	0.690	1.456	0.228	0.871	0.647
7	0.871	0.058	0.810	0.324	0.792
8	1.177	0.219	0.639	-0.383	0.637
9	0.874	0.157	0.692	0.315	0.603
10	1.361	0.416	0.519	-0.724	0.871
11	2.263	5.797	0.016	-1.919	0.758
12	1.658	0.368	0.544	-1.188	1.345
13	1.052	0.002	0.960	-0.119	0.639
14	1.323	0.409	0.523	-0.658	0.812
15	0.977	0.001	0.975	0.054	0.583
16	0.828	0.389	0.533	0.442	0.589
17	0.344	2.382	0.123	2.507	1.427
18	0.670	0.731	0.392	0.939	0.899
19	1.180	0.196	0.658	-0.390	0.666
20	0.835	0.341	0.559	0.423	0.596
21	1.159	0.070	0.791	-0.346	0.796

* $p < 0,5$

Maddelerin incelemesi yapılırken öncelikli olarak %95 anlamlılık düzeyine göre p değerlerine bakılmıştır. Daha sonrasında ise sayfa 19'da verilen sınıflama sistemine göre DMF düzeyleri belirlenmiştir. Maddelerin Δ -MH katsayısının almış olduğu değer negatif ise referans grubu için pozitif ise odak grubu için bir avantaj sağlamaktadır. Bu durumda DMF içeren maddeler, düzeyleri ve hangi gruba avantaj sağladığı ile ilgili bilgiler Tablo 15'de verilmiştir.

Tablo 15: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler

Madde	DMF düzeyi	Avantaj sağladığı grup
1	B	Odak (Erkek)
11	C	Referans (Kız)

Tablo 15 incelendiğinde MH yöntemine göre kız ve erkeklere farklı işleyen birer tane madde bulunmuştur. Kızlara avantajlı olacak şekilde işleyen madde C düzeyinde, erkeklere avantajlı olacak şekilde işleyen madde ise B düzeyindedir.

4.2 İkinci Alt Probleme İlişkin Bulgular

İkinci alt problem: PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde cinsiyet açısından Lojistik Regresyon yöntemine göre DMF gösteren madde var mıdır?

PISA 2009 matematik okuryazarlığı alt testinde bulunan maddelerin analizi SPSS paket programı ile yapılmıştır. Lojistik Regresyon ile DMF belirlenirken Zumbo (1999) tarafından yazılmış olan kod sistemi (syntax) kullanılmıştır. Elde edilen sonuçların tablolaştırılmış hali Tablo 16'da verilmiştir.

Tablo 16: Matematik testine ait Lojistik Regresyon yöntemine göre yapılan analiz sonuçları

Madde	β	p	R_1^2	R_2^2	R_3^2	ΔR^2
1	,244	0,030	0,256	0,267	0,267	0,011
2	-1,371	0,000	0,413	0,413	0,413	0,000
3	-0,560	0,615	0,411	0,414	0,414	0,003
4	-0,437	0,000	0,223	0,223	0,223	0,000
5	-0,119	0,289	0,254	0,259	0,259	0,005
6	-0,503	0,000	0,396	0,399	0,399	0,003
7	-1,430	0,000	0,352	0,352	0,352	0,000
8	-0,882	0,000	0,244	0,245	0,245	0,002
9	-0,569	0,000	0,239	0,239	0,239	0,000
10	-1,690	0,000	0,383	0,390	0,390	0,007
11	-0,927	0,000	0,515	0,538	0,538	0,023
12	-2,659	0,000	0,364	0,368	0,368	0,004
13	0,638	0,000	0,374	0,375	0,375	0,001
14	-1,314	0,000	0,462	0,468	0,468	0,006
15	0,156	0,163	0,202	0,202	0,202	0,000
16	0,031	0,780	0,292	0,294	0,294	0,002
17	3,016	0,000	0,218	0,250	0,250	0,032
18	-1,788	0,000	0,365	0,369	0,369	0,004
19	-0,651	0,000	0,419	0,421	0,421	0,002
20	0,194	0,084	0,264	0,265	0,265	0,001
21	-1,576	0,000	0,275	0,275	0,275	0,000

Analiz sonuçlarının değerlendirilmesi yapılırken, analize Model 1’de ilave edilen toplam puan, Model 2’de ilave edilen grup ve model 3’de ilave edilen grup ile toplam puanın etkileşimi değişkenlerinden elde edilen R^2 (Nagelkerke R Square) değerleri arasındaki farkı ($R_2^2 - R_1^2$ ile $R_3^2 - R_2^2$) temsil eden ΔR^2 ’nin almış olduğu değere göre yapılmaktadır. Ancak yapılan çalışmada R_3^2 ile R_2^2 değerlerinin aynı olmasından dolayı değerlendirmeler sadece $R_2^2 - R_1^2$ farkından elde edilen ΔR^2 değeri ile yapılmıştır.

DMF içerdiği belirlenen maddelerin hangi gruba avantaj sağladıklarını belirleyebilmek için ise İlişkisiz Örneklem t Testi uygulanmıştır. DMF içerdiği tespit edilen maddelerin düzeylerini belirlemek için Bakan Kalaycıođlu ve Keleciođlu (2011) tarafından düzenlenen sınıflama sistemi kullanılmıştır. Bu durumda DMF içeren maddeler, düzeyleri ve hangi gruba avantaj sağladığı ile ilgili bilgiler Tablo 17’de verilmiştir.

Tablo 17: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler

Madde	DMF düzeyi	Avantaj sağladığı grup
1	B	Odak (Erkek)
11	C	Referans (Kız)
17	C	Odak (Erkek)

Tablo 17 incelendiğinde Lojistik Regresyon yöntemine göre kız öğrencilere farklı işleyen 1 madde ve erkek öğrencilere farklı işleyen 2 tane madde bulunmuştur. Kızlara avantajlı olacak şekilde işleyen madde C düzeyinde, erkeklere avantajlı olacak şekilde işleyen maddeler ise sırasıyla B ve C düzeyindedir.

4.3 Üçüncü Alt Probleme İlişkin Bulgular

Üçüncü alt problem: PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde cinsiyet açısından SIBTEST yöntemine göre DMF gösteren madde var mıdır?

PISA 2009 matematik okuryazarlığı alt testinde bulunan maddeler SIBTEST yöntemi ile analiz edilmiştir. Elde edilen sonuçların tablolaştırılmış hali Tablo 18’de verilmiştir.

Tablo 18: Matematik testine ait SIBTEST yöntemine göre yapılan analiz sonuçları

Madde	β	Standart Hata	p
1	-0.117	0.055	0.032
2	-0.026	0.040	0.512
3	0.038	0.050	0.448
4	-0.017	0.053	0.750
5	-0.100	0.055	0.066
6	-0.066	0.049	0.183
7	-0.008	0.042	0.851
8	0.024	0.046	0.612
9	-0.022	0.052	0.667
10	0.039	0.040	0.332
11	0.130	0.044	0.003
12	0.031	0.028	0.260
13	0.006	0.051	0.913
14	0.040	0.040	0.323
15	-0.009	0.056	0.867
16	-0.048	0.054	0.377
17	-0.050	0.023	0.028
18	-0.052	0.036	0.154
19	0.016	0.048	0.748
20	-0.050	0.054	0.358
21	0.015	0.045	0.743

Verilerin değerlendirilmesi yapılırken β değerine bakılır. β değerinin pozitif olduğu maddeler referans grubu için, negatif olduğu maddeler ise odak grubu için avantaj sağladığını göstermektedir. DMF içeren maddelerin DMF düzeylerinin belirlenmesinde Roussos ve Stout (1995) tarafından oluşturulan sınıflama sistemi kullanılmıştır. Bu sınıflama sistemine göre DMF içeren maddeler, DMF düzeyleri ve hangi gruba avantaj sağladığı ile ilgili bilgiler Tablo 19’de verilmiştir.

Tablo 19:DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler

Madde	DMF düzeyi	Avantaj sağladığı grup
1	C	Odak (Erkek)
5	C	Odak (Erkek)
6	B	Odak (Erkek)
11	C	Referans (Kız)

Tablo 19 incelendiğinde SIBTEST yöntemine göre kız öğrencilere farklı işleyen 1 madde ve erkek öğrencilere farklı işleyen 3 tane madde bulunmuştur. Kızlara avantajlı olacak şekilde işleyen madde C düzeyinde, erkeklere avantajlı olacak şekilde işleyen maddeler ise sırasıyla C, C ve B düzeyindedir.

4.4 Dördüncü Alt Probleme İlişkin Bulgular

Dördüncü alt problem: PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde uygulandığı ülke açısından Mantel-Haenszel yöntemine göre DMF gösteren madde var mıdır?

PISA 2009 matematik okuryazarlığı alt testinde bulunan maddeler Mantel-Haenszel yöntemi ile analiz edilmiştir. Elde edilen sonuçların tablolaştırılmış hali tablo 20’de verilmiştir.

Tablo 20: Matematik testine ait Mantel-Haenszel yöntemine göre yapılan analiz sonuçları

	Almanya Türkiye			Finlandiya Türkiye			ABD Türkiye		
	p	Δ -MH	Düzye	p	Δ -MH	Düzye	p	Δ -MH	Düzye
1	0.000	2.272	C	0.000	2.122	C	0.002	1.352	B
2	0.003	1.443	B	0.001	1.574	C	0.859	0.137	
3	0.911	0.097		0.934	-0.084		0.049	0.891	
4	0.613	0.257		0.015	1.061	B	0.127	0.646	
5	0.117	0.710		0.723	0.191		0.184	-0.575	
6	0.195	-0.631		0.743	-0.177		0.000	-1.882	C
7	0.000	3.127	C	0.000	2.425	C	0.000	2.928	C
8	0.053	0.899		0.931	-0.004		0.067	0.806	
9	0.062	0.828		0.637	-0.237		0.000	1.844	C
10	0.128	0.891		0.029	1.150	B	0.040	1.111	B
11	0.000	-3.730	C	0.000	-2.552	C	0.547	-0.355	
12	0.003	-2.804	C	0.002	-2.615	C	0.010	-2.299	C
13	0.196	-0.687		0.015	-1.284	B	0.001	-1.424	B
14	0.529	-0.381		0.005	-1.495	B	0.000	-2.820	C
15	0.000	1.619	C	0.000	2.155	C	0.008	1.114	B
16	0.891	0.102		0.056	0.914		0.734	-0.171	
17	0.000	-3.809	C	0.082	-2.031		0.000	-4.853	C
18	0.157	0.836		0.710	0.257		0.437	0.495	
19	0.000	-3.407	C	0.000	-1.831	C	0.000	-2.054	C
20	0.000	-1.741	C	0.001	-1.462	B	0.357	-0.411	
21	0.001	1.586	C	0.059	0.941		0.434	0.438	

Maddelerin incelemesi yapılırken öncelikli olarak %95 anlamlılık düzeyine göre p değerlerine bakılmıştır. Daha sonrasında ise Zieky (1993) tarafından düzenlenen sınıflama sistemine göre DMF düzeyleri belirlenmiştir. Maddelerin Δ -MH katsayısının almış olduğu değer negatif ise referans grubu için pozitif ise odak grubu için bir avantaj sağlamaktadır. Bu durumda DMF içeren maddeler, düzeyleri ve hangi gruba avantaj sağladığı ile ilgili bilgiler Tablo 21’de verilmiştir.

Tablo 21: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler

Almanya Türkiye		Finlandiya Türkiye		ABD Türkiye	
Madde	Av. sağ. grup	Madde	Av. sağ. grup	Madde	Av. sağ. grup
1	Almanya	1	Finlandiya	1	ABD
2	Almanya	2	Finlandiya	6	Türkiye
7	Almanya	4	Finlandiya	7	ABD
11	Türkiye	7	Finlandiya	9	ABD
12	Türkiye	10	Finlandiya	10	ABD
15	Almanya	11	Türkiye	12	Türkiye
17	Türkiye	12	Türkiye	13	Türkiye
19	Türkiye	13	Türkiye	14	Türkiye
20	Türkiye	14	Türkiye	15	ABD
21	Almanya	15	Finlandiya	17	Türkiye
		19	Türkiye	19	Türkiye
		20	Türkiye		

Tablo 21'e göre Türkiye ile diğer üç ülke arasındaki DMF analizi sonuçlarında ülkeler arasındaki DMF'li madde sayıları Türkiye-Finlandiya arasında 12 madde, Türkiye-ABD 11 madde ve Türkiye-Almanya arasında ise 10 maddedir. Testteki 1., 7. ve 15. maddeler Türkiye ile karşılaştırılan ülkelerdeki öğrencilere avantaj sağlamış; 12. ve 19. Maddelerin ise Türkiye'deki öğrencilere avantaj sağladıkları tespit edilmiştir.

4.5 Beşinci Alt Probleme İlişkin Bulgular

Beşinci alt problem: PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde uygulandığı ülke açısından Lojistik Regresyon analizi yöntemine göre DMF gösteren madde var mıdır?

PISA 2009 matematik okuryazarlığı alt testinde bulunan maddeler Lojistik Regresyon yöntemi ile analiz edilmiştir. Elde edilen sonuçların tablolaştırılmış hali tablo 22'de verilmiştir.

Tablo 22: Matematik testine ait Lojistik Regresyon yöntemine göre yapılan analiz sonuçları

Madde	Almanya Türkiye		Finlandiya Türkiye		ABD Türkiye	
	ΔR^2	Düzeyi	ΔR^2	Düzeyi	ΔR^2	Düzeyi
1	0,037	C	0,026	C	0,017	B
2	0,011	B	0,014	B	0	
3	0		0		0,005	
4	0		0,006		0,004	
5	0,004		0		0,003	
6	0,004		0		0,026	C
7	0,051	C	0,032	C	0,056	C
8	0,004		0		0,006	
9	0,004		0,002		0,029	C
10	0,002		0,004		0,008	
11	0,088	C	0,046	C	0,001	
12	0,041	C	0,05	C	0,03	
13	0,005		0,014	B	0,018	B
14	0,002		0,014	B	0,049	C
15	0,017	B	0,028	C	0,011	B
16	0		0,004		0	
17	0,078	C	0,02		0,135	C
18	0,001		0		0,002	
19	0,077	C	0,025	C	0,028	C
20	0,028	C	0,022	C	0,001	
21	0,014	B	0,004		0,001	

Analiz sonuçlarının değerlendirilmesi yapılırken, analize Model 1’de ilave edilen toplam puan, Model 2’de ilave edilen grup ve model 3’de ilave edilen grup ile toplam puanın etkileşimi değişkenlerinden elde edilen R^2 (Nagelkerke R Square) değerleri arasındaki farkı ($R_2^2 - R_1^2$ ile $R_3^2 - R_2^2$) temsil eden ΔR^2 ’nin almış olduğu değere göre yapılmaktadır. Ancak yapılan çalışmada R_3^2 ile R_2^2 değerlerinin aynı olmasından dolayı değerlendirmeler sadece $R_2^2 - R_1^2$ farkından elde edilen ΔR^2 değeri ile yapılmıştır.

DMF içerdiği belirlenen maddelerin hangi gruba avantaj sağladıklarını belirleyebilmek için ise İlişkisiz Örneklem t Testi uygulanmıştır. DMF içerdiği tespit edilen maddelerin düzeylerini belirlemek için Bakan Kalaycıoğlu ve Kelecioğlu (2011) tarafından düzenlenen sınıflama sistemi kullanılmıştır. Bu durumda DMF içeren maddeler, düzeyleri ve hangi gruba avantaj sağladığı ile ilgili bilgiler Tablo 23’de verilmiştir.

Tablo 23: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler

Almanya Türkiye		Finlandiya Türkiye		ABD Türkiye	
Madde	Av. sağ. grup	Madde	Av. sağ. grup	Madde	Av. sağ. grup
1	Almanya	1	Finlandiya	1	Türkiye
2	Almanya	2	Finlandiya	6	ABD
7	Almanya	7	Finlandiya	7	Türkiye
11	Türkiye	11	Finlandiya	9	Türkiye
12	Almanya	12	Finlandiya	13	ABD
15	Almanya	13	Finlandiya	14	ABD
17	Türkiye	14	Finlandiya	15	Türkiye
19	Türkiye	15	Finlandiya	17	ABD
20	Almanya	19	Finlandiya	19	ABD
21	Almanya	20	Finlandiya		

Tablo 23 incelendiğinde bazı maddelerin sürekli olarak Türkiye dışındaki ülkelere avantajlı olacak şekilde işleyen madde olduğu tespit edilmiştir. Türkiye dışındaki ülkelere avantaj sağlayacak şekilde işleyen maddeler 1., 7. ve 15. maddelerdir.

4.6 Altıncı Alt Probleme İlişkin Bulgular

Altıncı alt problem: PISA 2009 Matematik okuryazarlığı alt testinde bulunan maddelerde uygulandığı ülke açısından SIBTEST yöntemine göre DMF gösteren madde var mıdır?

PISA 2009 matematik okuryazarlığı alt testinde bulunan maddeler SIBTEST yöntemi ile analiz edilmiştir. Elde edilen sonuçların tablolaştırılmış hali Tablo 24’de verilmiştir.

Tablo 24: Matematik testine ait SIBTEST yöntemine göre yapılan analiz sonuçları

	Almanya Türkiye		Finlandiya Türkiye		ABD Türkiye	
	β	Düzeyi	β	Düzeyi	β	Düzeyi
1	-0,17	C	-0,14	C	-0,129	C
2	-0,097	C	-0,1	C	-0,015	
3	-0,007		0,007		-0,084	C
4	-0,037		-0,09	C	-0,071	C
5	-0,054		-0,012		0,046	
6	0,044		0,003		0,134	C
7	-0,216	C	-0,176	C	-0,202	C
8	-0,073	C	-0,015		-0,076	C
9	-0,085	C	0,018		-0,174	C
10	-0,039		-0,041		-0,063	C
11	0,234	C	0,159	C	0,016	
12	0,074	C	0,095	C	0,054	
13	0,041		0,071	C	0,11	C
14	0,021		0,096	C	0,144	C
15	-0,151	C	-0,185	C	-0,109	C
16	-0,006		-0,065	C	0,006	
17	0,104	C	0,025		0,184	C
18	-0,036		-0,005		-0,029	
19	0,238	C	0,167	C	0,154	C
20	0,146	C	0,103	C	0,031	
21	-0,117	C	-0,08	C	-0,029	

Verilerin değerlendirilmesi yapılırken β değerine bakılır. β değerinin pozitif olduğu maddeler referans grubu için, negatif olduğu maddeler ise odak grubu için avantaj sağladığını göstermektedir. DMF içeren maddelerin DMF düzeylerinin

belirlenmesinde Roussos ve Stout (1995) tarafından oluşturulan sınıflama sistemi kullanılmıştır. Bu sınıflama sistemine göre DMF içeren maddeler, DMF düzeyleri ve hangi gruba avantaj sağladığı ile ilgili bilgiler Tablo 25’de verilmiştir.

Tablo 25: DMF içeren maddeler, düzeyleri ve avantaj sağladığı grup ile ilgili bilgiler

Almanya Türkiye		Finlandiya Türkiye		ABD Türkiye	
Madde	Av. sağ. grup	Madde	Av. sağ. grup	Madde	Av. sağ. grup
1	Almanya	1	Finlandiya	1	ABD
2	Almanya	2	Finlandiya	3	ABD
7	Almanya	4	Finlandiya	4	ABD
8	Almanya	7	Finlandiya	6	Türkiye
9	Almanya	11	Türkiye	7	ABD
11	Türkiye	12	Türkiye	8	ABD
12	Türkiye	13	Türkiye	9	ABD
15	Almanya	14	Türkiye	10	ABD
17	Türkiye	15	Finlandiya	13	Türkiye
19	Türkiye	16	Finlandiya	14	Türkiye
20	Türkiye	19	Türkiye	15	ABD
21	Almanya	20	Türkiye	17	Türkiye
		21	Finlandiya	19	Türkiye

Tablo 25 incelendiğinde bazı maddelerin sürekli olarak Türkiye dışındaki ülkelere bazı maddelerin ise sürekli olarak Türkiye için avantajlı olacak şekilde işleyen madde olduğu tespit edilmiştir. Türkiye dışındaki ülkelere avantaj sağlayacak şekilde işleyen maddeler 1., 7. ve 15. maddeler; Türkiye için avantajlı olacak şekilde işleyen madde ise 19. maddedir.

BÖLÜM V

5 Sonuçlar ve Öneriler

5.1 Sonuçlar

5.1.1 Cinsiyet Değişkenine ilişkin sonuçlar

PISA 2009 matematik okuryazarlığı alttestindeki maddeler için cinsiyete göre yapılan DMF analizlerinde toplamda 5 maddede DMF bulunduğu tespit edilmiştir.

Tablo 26: Cinsiyet Değişkenine göre Elde Edilen DMF'li Maddeler ve Avantaj sağladıkları gruplar

	MH	LR	SIBTEST
1	B Erkek	B Erkek	C Erkek
5			C Erkek
6			B Erkek
11	C Kız	C Kız	C Kız
17		C Erkek	

MH analizi sonuçlarına göre iki maddede DMF bulunmuş, bu maddelerin biri erkekler için avantajlı diğer madde ise kızlar için avantajlı olduğu görülmüştür. LR analizi sonucunda üç maddede DMF bulunmuş, bu maddelerin iki tanesi erkekler lehine bir tanesi kızlar lehine avantaj sağlamaktadır. SIBTEST analizi sonucunda dört maddede DMF bulunmuş ve bu maddelerin üç tanesi erkekler lehine bir tanesi ise kızlara avantaj sağlamaktadır. Genel anlamda düşünülecek olursa DMF'li olduğu

belirlenen beş maddenin dördü erkekler lehine, sadece bir tane maddenin kızlar lehine avantajlı olduğu söylenebilir. Cinsiyete göre yapılan DMF analizi sonuçlarına göre her üç yöntemde de 1. ve 11. maddelerde DMF olduğu bulunmuştur. Bu iki madde dışında Lojistik Regresyon analizi sonucunda 17. madde, SIBTEST yöntemi ile DMF analizi sonucunda ise 5. ve 6. maddeler erkekler lehine DMF içerdiği bulunmuştur. Bu açıdan düşünüldüğünde Mantel-Haenszel yöntemi sadece üç yöntemde DMF içeren maddeleri belirlemiştir.

Elde edilen sonuçlara göre DMF gösteren madde sayısı bakımından yöntemler arasında fark vardır ancak maddenin avantajlı olduğu grup her bir yöntem için aynıdır (1., 5., 6. ve 17. maddeler erkekler lehine; 11. madde kızlar lehine). Bu açıdan araştırmadan elde edilen sonuç Yurdugül (2003), Bekci (2007) ve Gök ve diğerlerinin (2010) yapmış oldukları çalışmalarında elde ettiği sonuç ile tutarlılık göstermektedir.

Ayan (2011)'in PISA 2009 Fen Bilgisi maddeleri ile yapmış olduğu çalışmada hem MH yöntemi hem de LR yönteminde dört adet DMF gösteren madde tespit edilmiştir. Ancak bu 4'er maddenin tamamı birbirinden farklıdır. Yapılan çalışmada ise DMF'li olduğu tespit edilen maddelerin yöntemlere göre bu kadar farklılık göstermediği görülmüştür. Bu açıdan elde edilen sonuçlar, Ayan (2011)'in yapmış olduğu çalışma ile farklılık göstermektedir.

Yıldırım (2008)'in PISA 2003 Matematik testindeki maddeler ile yapmış olduğu çalışma sonucunda MH yöntemi 13, Olabilirlik Oranı Analizi yöntemi 15 maddeyi farklı işleyen madde olarak tespit ederken Sınırlandırılmış Faktör Çözümlemesi yöntemi yedi maddeyi farklı işleyen madde olarak tespit etmiştir. Bu yedi madde her üç yöntem tarafından da farklı işleyen madde olarak tespit edilmiştir. Yine PISA 2009 Fen Bilgisi maddeleri ile yapılan çalışmada toplam da sekiz maddede DMF tespit edilmiştir. Uluslararası düzeyde yapılmakta olan PISA çalışmalarında cinsiyete göre DMF gösteren maddelerin yıllara göre azalma gösterdiği söylenebilir.

5.1.2 Kültür Değişkenine ilişkin sonuçlar

Almanya ile Türkiye'den elde edilen verilerle yapılan DMF analizleri sonucunda toplamda 12 maddede DMF olduğu tespit edilmiştir.

Tablo 27: Ülke Değişkenine göre Elde Edilen DMF'li Maddeler ve Avantaj sağladıkları gruplar

Madde	Almanya Türkiye		
	MH	LR	SIBTEST
1	Almanya	Almanya	Almanya
2	Almanya	Almanya	Almanya
7	Almanya	Almanya	Almanya
8			Almanya
9			Almanya
11	Türkiye	Türkiye	Türkiye
12	Türkiye	Almanya	Türkiye
15	Almanya	Almanya	Almanya
17	Türkiye	Türkiye	Türkiye
19	Türkiye	Türkiye	Türkiye
20	Türkiye	Almanya	Türkiye
21	Almanya	Almanya	Almanya

MH analizi sonuçlarına göre on maddede DMF bulunmuş, bu maddelerin beşi Almanya için avantajlı diğer maddeler ise Türkiye için avantajlı olduğu görülmüştür. LR analizi sonucunda on maddede DMF bulunmuş, bu maddelerin yedi tanesi Almanya lehine üç tanesi Türkiye lehine avantaj sağlamaktadır. SIBTEST analizi sonucunda on iki maddede DMF bulunmuş ve bu maddelerin yedi tanesi Almanya lehine beş tanesi ise Türkiye lehine avantaj sağlamaktadır. Genel anlamda düşünülecek olursa DMF'li olduğu belirlenen on iki maddenin yedi tanesi Almanya lehine, üç tane maddenin Türkiye lehine avantajlı olduğu ayrıca iki tane maddenin MH ile SIBTEST yöntemlerine göre Türkiye lehine avantaj sağlarken, LR yöntemine göre ise Almanya lehine avantaj sağladığı söylenebilir.

Finlandiya ile Türkiye’den elde edilen verilerle yapılan DMF analizleri sonucunda toplamda 14 maddede DMF olduğu tespit edilmiştir.

Tablo 28: Ülke Değişkenine göre Elde Edilen DMF’li Maddeler ve Avantaj sağladıkları gruplar

Madde	Finlandiya Türkiye		
	MH	LR	SIBTEST
1	Finlandiya	Finlandiya	Finlandiya
2	Finlandiya	Finlandiya	Finlandiya
4	Finlandiya		Finlandiya
7	Finlandiya	Finlandiya	Finlandiya
10	Finlandiya		
11	Türkiye	Finlandiya	Türkiye
12	Türkiye	Finlandiya	Türkiye
13	Türkiye	Finlandiya	Türkiye
14	Türkiye	Finlandiya	Türkiye
15	Finlandiya	Finlandiya	Finlandiya
16			Finlandiya
19	Türkiye	Finlandiya	Türkiye
20	Türkiye	Finlandiya	Türkiye
21			Finlandiya

MH analizi sonuçlarına göre on iki maddede DMF bulunmuş, bu maddelerin altı tanesi Finlandiya için avantajlı altı tanesi ise Türkiye için avantajlı olduğu görülmüştür. LR analizi sonucunda on maddede DMF bulunmuş, bu maddelerin tamamı Finlandiya lehine avantaj sağlamaktadır. SIBTEST analizi sonucunda on üç maddede DMF bulunmuş ve bu maddelerin yedi tanesi Finlandiya lehine altı tanesi ise Türkiye lehine avantaj sağlamaktadır. Genel anlamda düşünülecek olursa DMF’li olduğu belirlenen on dört maddenin sekiz tanesi Finlandiya lehine, altı tane maddenin MH ile SIBTEST yöntemlerine göre Türkiye lehine avantaj sağlarken, LR yöntemine göre ise Finlandiya lehine avantaj sağladığı söylenebilir.

ABD ile Türkiye'den elde edilen verilerle yapılan DMF analizleri sonucunda toplamda 14 maddede DMF olduğu tespit edilmiştir.

Tablo 29: Ülke Değişkenine göre Elde Edilen DMF'li Maddeler ve Avantaj sağladıkları gruplar

ABD Türkiye			
Madde	MH	LR	SIBTEST
1	ABD	ABD	ABD
3			ABD
4			ABD
6	Türkiye	Türkiye	Türkiye
7	ABD	ABD	ABD
8			ABD
9	ABD	ABD	ABD
10	ABD		ABD
12	Türkiye		
13	Türkiye	Türkiye	Türkiye
14	Türkiye	Türkiye	Türkiye
15	ABD	ABD	ABD
17	Türkiye	Türkiye	Türkiye
19	Türkiye	Türkiye	Türkiye

MH analizi sonuçlarına göre on bir maddede DMF bulunmuş, bu maddelerin beş tanesi ABD için avantajlı altı tanesi ise Türkiye için avantajlı olduğu görülmüştür. LR analizi sonucunda dokuz maddede DMF bulunmuş, bu maddelerin dört tanesi ABD lehine, beş tanesi ise Türkiye avantaj sağlamaktadır. SIBTEST analizi sonucunda on üç maddede DMF bulunmuş ve bu maddelerin sekiz tanesi ABD lehine beş tanesi ise Türkiye lehine avantaj sağlamaktadır. Genel anlamda düşünülecek olursa DMF'li olduğu belirlenen on dört maddenin sekiz tanesi ABD lehine, altı tane maddenin ise Türkiye lehine avantaj sağladığı söylenebilir.

Elde edilen verilerin tamamı düşünülecek olursa 1., 7. ve 15. maddelerin Türkiye dışındaki ülkeler lehine avantajlı, 19. maddenin Türkiye lehine avantajlı olduğu ve 12. maddenin MH ile SIBTEST yöntemlerine göre Türkiye lehine avantaj sağlarken, LR yöntemine göre ise Türkiye dışındaki ülkeler lehine avantaj sağladığı söylenebilir. Genel anlamda yöntemler karşılaştırılacak olursa; Mantel-Haenszel ve lojistik regresyon yöntemlerine göre SIBTEST yöntemi DMF içeren maddeleri daha hassas olarak belirleyebilmektedir.

Farklı kültürleri ve farklı dilleri olan ABD, Almanya, Finlandiya ve Türkiye arasında yapılan analizler sonucunda, Matematik testinde bulunan maddelerin yaklaşık olarak yarısında DMF içeren madde bulunduğu görülmüştür. Bu sonuç, dilsel ve kültürel farklılıkların arttıkça DMF gösteren maddelerin arttığını göstermektedir. Elde edilen bu sonuç, Asil (2010) ve Atalay (2010)'un yaptığı çalışma sonuçları ile tutarlılık göstermektedir.

Geçerlik, kuramsal bilgilerin ve gözlenebilir davranışların geliştirilen test veya ölçek puanları ile doğrulanmasıdır. Geçerlik süreci, uygulanan ölçeğin geçerli ve güçlü bir bilimsel temele sahip olduğunu göstermektir. 1954 yılında yayımlanan test standartları raporunda geçerlik kavramının yapı geçerliliği şemsiyesi altında bütünleşmesi gerektiği tartışılmıştır (Cronbach ve Meehl, 1955). Bu tartışmalar ışığında görünüş geçerliği, kapsam geçerliği ve ölçüte dayalı geçerlik yapısal geçerliğin bir göstergesi olarak değerlendirilir (Şencan, 2005). Bu kapsamda düşünüldüğünde geçerliği etkileyen her etmen yapı geçerliğini etkileyecektir. Bu nedenle geçerliği etkileyen en büyük etmenlerden birisi olan madde yanlılığı yapı geçerliği için bir sorundur. Yapılan araştırmada yanlılık analizi için ilk adım olan DMF analizi sonuçlarına göre PISA 2009 uygulaması matematik testinde bulunan maddelerin %60'ı DMF içermektedir ve bu maddeler yapı geçerliği için tehdit teşkil etmektedirler.

5.2 Öneriler

Bu çalışma sonunda yapılacak öneriler iki grupta toplanabilir.

5.2.1 Araştırma Sonuçlarından Çıkan Öneriler

1. PISA uygulamasının diğer yıllarda yapılan çalışmalarında ve PISA uygulamasında kullanılan diğer okuryazarlık alt testlerinde de DMF analizleri yapılarak genel sonuçlar çıkarılabilir, DMF'ye neden olan noktalar belirlenerek sonraki çalışmalarda benzer hatalar yapılmaması noktasında test geliştiricilere tavsiyelerde bulunulabilir.

2. DMF analizleri yanlılık araştırmalarının ilk aşaması olarak değerlendirilir. DMF'li olduğu tespit edilen maddelerin yanlı olup olmadığına alt gruplar testin yapısı ve kapsamı dikkate alınarak yapılan incelemeler sonucunda karar verilir. Farklı alt gruplardaki bireylerin test puanlarının buldukları gruba bağlı olarak değişmesi diğer bir anlamda yanlı olması sistematik hatadır. Bu nedenle yanlılık bir geçerlik sorunudur. Bu nedenle önemli bir psikometrik özellik olan geçerlik durumu için cinsiyet ve kültür değişkenine göre DMF barındırdığı belirlenen maddelerin incelenip yanlı olup olmadıklarının belirlenmesi önerilir.

3. Farklı ülkelerde uygulanan PISA verileri üzerinde elde edilen veriler üzerinden yapılan analizler sonucunda yaklaşık olarak testteki maddelerin %50'sini DMF'li maddelerden oluştuğu görülmektedir. Bu amaçla MEB Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı'nın uygulamalarda görev alan çevirmenlerin seçimi, eğitimi ve PISA çalışmalarında çevirmenler, alan uzmanları ve ölçme değerlendirme uzmanlarının ortak çalışmasının sağlanması konusuna daha fazla önem verilmesi önerilebilir.

5.2.2 Araştırmacılara Yönelik Öneriler

1. Bu çalışmada cinsiyet değişkeni için Türkiye verisi ve dil değişkeni için ise Türkiye, ABD, Finlandiya ve Almanya ile sınırlı tutulmuştur. Araştırmaya başka değişkenler de eklenebilir veya farklı dil seçenekleri de dikkate çalışmalar da yapılabilir.

2. Çalışmada DMF belirleme yöntemlerinden Maentel Haenzsel, Lojistik Regresyon ve SIBTEST yöntemleri kullanılmış ve sonuçları karşılaştırılmıştır. Diğer yöntemler ele alınarak karşılaştırılmalarda yapılması önerilmektedir.

KAYNAKÇA

- Angoff, W.H. (1972). A technique for investigation of cultural differences. *Paper presented at the American Psychological Association meeting, Honolulu* (ERIC Doc. No. ED 069686).
- Angoff, H. W. (1993). Perspectives on Differential Item Functioning Methodolgy. P.W. Holland ve H. Wainer (ed.), *Differential item functioning* (ss. 31-66). Erlbaum: Hillsdale NJ.
- Asil, M. (2010). *Uluslararası Öğrenci Değerlendirme Programı (PISA) 2006 Öğrenci Anketinin Kültürler Arası Eşdeğerliğinin İncelenmesi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, yayınlanmamış doktora tezi.
- Asil M. ve Gelbal S. (2012). PISA Öğrenci Anketinin Kültürler Arası Eşdeğerliği. *Eğitim ve Bilim* 37 (166), 236-249.
- Atalay, K. (2010) *PISA 2006 Öğrenci Anketinde Yer Alan Tutum Maddelerinin Değişen Madde Fonksiyonu Açısından İncelenmesi*, Yayınlanmamış Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Atalay, K., Gök, B, Kelecioğlu H. ve Arslan, N. (2012) Değişen Madde Fonksiyonunun Belirlemesinde Kullanılan Farklı Yöntemlerin Karşılaştırılması: Bir Simülasyon Çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi*, 43 (2), 270-281.
- Ayan, C. (2011). *PISA 2009 Fen Okuryazarlığı Alt Testinin Değişen Madde Fonksiyonu Açısından İncelenmesi*, Yayınlanmamış Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Baghi, H ve Ferrara, S. F. (1990) *Detecting Differential Item Functioning Using IRT and Mantel-Haenszel Techniques: Implementing Procedures and Comparing Results*. Paper presented at the Annual Meeting of the Eastern Educational Research Association (14-17 Şubat, Clearwater).

- Bakan Kalaycıođlu, D. ve Keleciođlu, H. (2011). Öğrenci Seçme Sınavı'nın madde yanlılığı açısından incelenmesi. *Eđitim ve Bilim*, 36 (161), 3-12.
- Bekçi, B. (2007). *Ortaöđretim Kurumları Öğrenci Seçme Sınavının Deđişen Madde Fonksiyonlarının Cinsiyete ve Okul Türüne Göre İncelenmesi*, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Bishop, Y., Fienberg, S. E., ve Holland, P.W. (1975), *Discrete Multivariate Analysis: Theory and Practice*, Cambridge, MA: MIT Press.
- Büyüköztürk, Ş. (2002) Faktör analizi: Temel Kavramlar ve Ölçek Geliştirmede Kullanımı. *Kuram ve Uygulamada Eđitim Yönetimi* 32, 470-480.
- Büyüköztürk, Ş. (2011). *Sosyal Bilimler İçin Veri Analizi El Kitabı*. (14. Basım) Ankara: Pegem Yayınları.
- Camilli, G. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational Measurement (4th ed., Vol. 4, pp. 221-256)*. Westport: American Council on Education & Praeger Publishers.
- Camilli, G. ve Shepard, L. A. (1994). *Methods for Identifying Biased Test Items*. Hollywood: Sage Publications.
- Clauser, B. ve Mazor, K. (1998). Using statistical procedures to identify differential item functioning test items. *Educational Measurement: Issue and Practice*, 17, 31-44.
- Çepni, Z, (2011). *Deđişen Madde Fonksiyonlarının SIBTEST, Mantel Haenszel, Lojistik Regresyon ve Madde Tepki Kuramı Yöntemleriyle İncelenmesi*, Doktora Tezi, Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Çet, S. (2006), *PISA 2003 Matematik Maddeleri Kullanılarak Yanlı Çalışan Maddelerin Tespitinde Çok Boyutlu Eşleştirme Analizi*, Yayınlanmamış Doktora Tezi, Ankara: Orta Dođu Teknik Üniversitesi.

- Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and White students in integrated colleges. *Journal of Educational Measurement*, 5, 115-124.
- Cronbach, L. J. ve Meehl, P. E. (1955) Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Dođan, N. ve Öğretmen, T. (2008). Deđişen Madde Fonksiyonunu Belirlemede Mantel-Haenszel, Ki-kare ve Lojistik Regresyon Tekniklerinin Karşılaştırılması. *Eđitim ve Bilim*, 33, 100-112.
- Dorans, N. J., ve Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. P. W. Holland ve H. Wainer (ed.), *Differential İtem Functioning* (ss. 35-66). Lawrence Erlbaum: Hillsdale, NJ.
- Dorans, N. J. ve Kullick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, 23, 355-368.
- Du Toit, W. F. (2003). *The discriminant validity of the culture assessment instrument: a Comparison of company cultures*. Yayınlanmamış Doktora Tezi, Johannesburg Rand Afrikaans University
- EARGED, (2010) PISA 2009 Uluslararası Öğrenci Deđerlendirme Programı Ulusal Ön Rapor, Ankara, MEB.
- EARGED, (2011) TIMSS 2007 Ulusal Matematik ve Fen Raporu 8. Sınıflar Ankara, MEB.
- Ercikan, K., Gierl, M. J., McCreith, T., Puhan, G., & Koh, K. (2004). Comparability of Bilingual Versions of Assessments: Sources of Incomparability of English and French Versions of Canada's National Achievement Tests. *Applied Measurement in Education*, 17(3), 301-321.
- Gierl, M. J. (2000). Construct Equivalence on Translated Achievement Tests. *Canadian Journal of Education*, 25(4), 280-296.

- Gierl, M. H., Khaliq, S. N. ve Boughton, K. (1999). *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-Haenzsel ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 35, 3-16.
- Holland, P. W. (1985). On the study of differential item performance without IRT. In *Proceedings of the 27th Annual Conference of the Military Testing Association* (Vol. 1, pp. 282-287). San Diego, CA: Navy Personnel Research and Development Center.
- Holland, P.W., ve Thayer, D.T. (1986). *Differential item performance and the Mantel-Haenzsel procedure* (Technical Report No. 86-69). Princeton, NJ: Educational Testing Service.
- Holland P. W. ve Wainer H. (1993). *Differential Item Functioning* (ss. 35-66). Lawrence Erlbaum: Hillsdale, NJ.
- Hubley, A. M., & Zumbo, B. D. (1996). A dialectic on validity: Where we have been and where we are going. *The Journal of General Psychology*, 123, 207-215.
- Jodoin, M. G. ve Gierl, M.J. (2001). Evaluating Type I error and power rates using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14, 329-349.
- Le, L. T. (2009). Investigating Gender Differential Item Functioning across Countries and Test Languages for PISA Science Items *International Journal of Testing*, 9 (2), 122- 133.
- Lee, L.F. (1995), Asymptotic bias in simulated maximum likelihood estimation of discrete choice models. *Econometric Theory*, 11, 437-483.

- Liu, L. O. ve Wilson, M. (2009). Gender Differences in Large-Scale Math Assessments. PISA Trend 2000 and 2003, *Applied Measurement in Education*, 22 (2), 164- 184.
- Mantel, N. ve Haenszel, W. (1959) Statistical aspects of the analysis of data from retrospective studies of disease. *Journal Natural Cancer Institue*, 22 (4), 719–748.
- Mellor, T. L. (1995). *A comparison of four differential item functioning methods for polytomously scored items*. Yayınlanmamış Doktora Tezi. The university of Texas Austin
- Miller, T. R., & Spray, J. A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of Educational Measurement*, 30 (2), 107-122
- OECD, (2009) Organization for Economic Cooperation and Development 2009. *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science*. Paris: OECD.
- Osterlind, S. ve H. Everson (2009). *Differential item functioning*. Londra: Sage Publications.
- Öğretmen, T. (1995). Differential item Functioning analysis of the verbal ability section of the first stage of the university entrance examination in Turkey. Yayınlanmamış Yüksek Lisans Tezi. Middle East Technical University, Ankara.
- Penfield, R. ve Camilli, G. (2007). Test fairness and differential item functioning. C.R. Rao (ed.) *Handbook of statistics: Psychometrics*, 26 (ss. 125-167). Amsterdam: Elsevier.
- Roussos, L. ve Stout, W. (1996) A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement* 20, 355-371.
- Scheuneman, J. (1979) A method for assessing bias in test items. *Journal of Educational Measurement*, 16 (3), 143-152.

- Shealy, R., ve Stout, W. F. (1993b). A model-based standardization approach that separates true bias/DIF from group differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika*, 58, 159-194.
- Swaminathan, H. ve Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayınevi.
- Thissen, D. (2001). *IRTLRDIF v.2.0.b: Software for the Computation of the Statistics Involved in Item Response Theory Likelihood-Ratio tests for Differential Item Functioning*.
- Thissen, D., Steinberg, L. ve Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P.W. Holland ve H. Wainer (Eds.), *Differential item functioning* (pp. 67-113). Hillsdale NJ: Erlbaum.
- Turgut, F. (1997). *Eğitimde Ölçme ve Değerlendirme Metotları*. Ankara: Yargıcı Matbaası.
- Waller N. G. (1998) EZDIF: Detection of Uniform and Nonuniform Differential Item Functioning With the Mantel-Haenszel and Logistic Regression Procedures. *Applied Psychological Measurement*, 22 (4), 391.
- 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. (2006); *Uluslararası Değerlendirme Çalışmalarında Kullanılan Matematik Sorularının Madde Yanlılığı Yöntemleri Kullanılarak İncelenmesi*, Yayınlanmamış Doktora Tezi, Ankara: Orta Doğu Teknik Üniversitesi.
- Yıldırım, D. (2008). Farklı İşleyen Maddelerin Belirlenmesinde Sınırlandırılmış Faktör Çözümlemesinin Olabilirlik-Oranı ve Mantel-Haenszel Yöntemleriyle Karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi*, 34, 297-307.

- Yurdugül, H. (2003). *Ortaöğretim kurumları seçme ve yerleştirme sınavının madde yanlılığı açısından incelenmesi*. Yayımlanmamış doktora tezi, Hacettepe Üniversitesi.
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. P. W. Holland ve H. Wainer (ed.) *Differential item functioning* (ss. 337-347). Erlbaum: Hillsdale, NJ.
- Zumbo, B. D. ve Thomas, D. R. (1997). *A measure of effect size for a model-based approach for studying DIF* (Working paper of the Edgeworth Laboratory for Quantitative Behavioral Science). Prince George, Canada: University of Northern British Columbia.
- 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*. Directorate of Human Resources Research and Evaluation, Department of National Defense. Ottawa: Canada.