



Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Eğitim Bilimleri Anabilim Dalı

Eğitimde Ölçme ve Değerlendirme Bilim Dalı

**PISA 2006 FEN BAŞARI TESTİNİN MADDE YANLILIĞININ KÜLTÜR VE
DİL AÇISINDAN İNCELENMESİ**

N. Bilge Başusta

Doktora Tezi

Ankara, 2013

PISA 2006 FEN BAŞARI TESTİNİN MADDE YANLILIĞININ KÜLTÜR VE DİL
AÇISINDAN İNCELENMESİ

N. Bilge Başusta

Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü

Eğitim Bilimleri Anabilim Dalı

Eğitimde Ölçme ve Değerlendirme Bilim Dalı

Doktora Tezi

Ankara, 2013

KABUL VE ONAY

N. Bilge Başusta tarafından hazırlanan "PISA 2006 Fen Başarı Testinin Madde Yanlılığının Kültür ve Dil Açısından İncelenmesi" başlıklı bu çalışma, 24/01/2013 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Doktora Tezi olarak kabul edilmiştir.



Prof.Dr., Şener Büyüköztürk (Başkan)



Prof. Dr., Selahattin Gelbal (Danışman)



Doç. Dr. Hülya Kelecioğlu



Doç. Dr. Nuri Doğan



Yrd. Doç. Dr. Neşe Güler

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. Yusuf Çelik

Enstitü Müdürü

BİLDİRİM

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

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

24.01.2013



N. Bilge Başusta

TEŞEKKÜRLER

Kendime tüm yaşamım boyunca birçok bakımdan örnek aldığım danışmanım Sn. Prof. Dr. Selahattin Gelbal'a; çalışmamın başından sonuna kadar desteğini ve önerilerini hiç eksik etmeyen hocalarım Doç Dr Hülya Kelecioğlu'na; Doç Dr. Nuri Doğan'a; Doç Dr. Tuncay Öğretmen'e; Yrd. Doç. Dr. Neşe Güler'e; Prof Dr. Şener Büyüköztürk'e;

Eşsiz destekleriyle yanımda olan ve beni yüreklendiren, anlayışlarını benden hiç esirgemeyen Tıp Eğitimi ve Bilişimi Anabilim Dalı çalışanlarına, özellikle kendime olan inancımı artıran, beni benden fazla dinleyen, kıymetli hocam Doç.Dr. Melih Elçin'e;

Tezime görüşleriyle katkı sağlayan ve değerli zamanlarının bir bölümünü benimle paylaşan öğretmenlere ve uzmanlara;

Doktora eğitimim boyunca almış olduğum maddi destekten dolayı TÜBİTAK' a;

çok teşekkür ederim.

Çoğaldığım ve çoğalttığım tüm satırların, doldurduğum tüm zamanların, attığım tüm adımların koşulsuz değişmeyen paydaşları Seçilime, Anneme ve Babama; yine varlıklarıyla hayatımı genişleten Turgay'a ve Poyrazıma ve ... son 4 yıldır her sabah güneşten önce gülen gözleriyle aydınlandığım, sol yanımda, en büyük desteğim, herşeyim Serhatıma ne kadar teşekkür etsem azdır.

Onurla sunarım.

ÖZET

BAŞUSTA N. Bilge, *PISA 2006 Fen Başarı Testinin Madde Yanlılığının Kültür ve Dil Açısından İncelenmesi*, Doktora Tezi, Ankara, 2013.

Karşılaştırmalı araştırmalarında değişen madde fonksiyonunun (DMF) varlığı olası yanlılık sebebi olarak düşünülür. Bu çalışmada, PISA 2006 testi maddelerinin farklı alt gruplarda kültür ve dile göre DMF analizleri yürütülmüştür. DMF analizleri Mantel Haenszel (MH), lojistik regresyon (LR), Alan indeksleri yöntemleriyle gerçekleştirilmiştir.

Çalışma grubunu beşinci kitapçığı alan 1744 öğrenciden oluşan Kanada örneklemini, 1124 öğrenciden oluşan Avustralya örneklemini, 1008 öğrenciden oluşan İngiltere örneklemini ve 377 öğrenciden oluşan Türkiye örneklemini ile birinci kitapçığı alan 1430 İngiltere örneklemini ile 380 Türkiye örneklemini oluşturmaktadır. Bu ülkeler DMF'nin temel nedeni olarak görülen kültürel ve dilsel farklılıklar nedeni ile seçilmiştir. DMF'nin olası nedenlerini belirlemek için alan uzmanlarının görüşlerine başvurulmuştur.

Çalışmada Kanada örnekleminde fransızca formu ve ingilizce formu alan gruplar için; MH tekniğine göre üç maddede B düzeyinde, üç maddede A düzeyinde; LR yöntemine göre üç maddede ihmal edilebilir düzeyde; alan indeksleri yöntemine göre beşer maddede DMF bulunmuştur. Avustralya–İngiltere örnekleminde MH tekniğine göre; dört maddede ihmal edilebilir düzeyde, bir maddede B düzeyinde ve bir maddede C düzeyinde; LR yöntemine göre dört maddede ihmal edilebilir düzeyde; alan indeksleri yöntemine göre ikişer maddede DMF bulunmuştur. Birinci kitapçık için Türkiye–İngiltere örnekleminde MH tekniğine göre beşi A düzeyinde, ikisi B düzeyinde ve üçü C düzeyinde olan on madde DMF içermektedir, bu maddelerin sekizi İngiltere ikisi ise Türkiye lehinedir. Lojistik regresyon yöntemine göre; beş maddede ihmal edilebilir düzeyde ve bir maddede B düzeyinde; alan indeksleri yöntemine göre altışar maddede DMF bulunmuştur.

Dilsel ve kültürel farklılıklar arttıkça DMF'li madde sayısının da arttığı gözlenmiştir. Kullanılan tekniğe göre DMF'li madde sayısı değişmektedir. LR ve MH arasındaki

korelasyon ve alan indeksleri yöntemleri arasındaki korelasyon aynı kültür farklı dil gruplaması için 0,01 düzeyinde anlamlıdır. Farklı kültür aynı dil için 0,05 düzeyinde ve farklı kültür farklı dil beşinci kitapçık için gruplaması için MH ve LR arasındaki korelasyon 0,01 düzeyinde; farklı kültür farklı dil alan indekleri tekniğiyle elde edilen korelasyon 0,05 düzeyinde anlamlı bulunmuştur.

Genel olarak; yanlılık araştırmaları DMF'nin olası temel nedenini kültür ve dile bağlı farklılıklar ve program farklılıkları olarak belirlemiştir.

Anahtar Sözcükler

Değişen madde fonksiyonu, Mantel Haenszel, Lojistik Regresyon, Alan İndeksleri Yöntemleri, PISA 2006

ABSTRACT

BAŞUSTA N. Bilge, *An Investigation of Item Bias in PISA 2006 Science Test in Terms of The Language and Culture*, Ph. D. Dissertation, Ankara, 2013.

In comparability investigations, the presence of differential item functioning (DIF) is considered to be an indication of possible bias. In this study, differential item functioning (DIF) analyses of Science items of PISA 2006 tests were carried out between different samplings in respect to language and culture.. Mantel Haenszel (MH), logistic regression(LR) and signed - unsigned area indexes methods were used for DIF detection analyses.

The research group of this study consists of the Australia sample comprising 1124 students, the Canada sample comprising 1744 students; the England sample comprising 1008 students, the Turkey sample comprising 377 students; took the fifth booklets and the England sample comprising 1430 students, the Turkey sample comprising 380 students; took the first booklets. These countries were selected due to the differences in cultural relevance and linguistic are the possible main reasons for differential item functioning (DIF). In order to investigate the sources of DIF field specialist opinions were consulted.

In the study, in Canadian sampling, DIF was found in three items at B level and three items at negligible level according to the MH technique and in three item at negligible level according to LR technique, in five items according to each fo signed - unsigned area indexes methods. In Australia- England sampling DIF was found in one item at B level and C level, four items at negligible level according to the MH technique and in four items at negligible level according to LR technique, in two items according to each fo signed - unsigned area indexes methods. In England- Turkey sampling for the first booklet; ten items included DIF according to MH results; five of them were at A level, two of them were at B level and three of them were at C level according to the MH technique those of the items eight of them favored English form, where two of them favored Turkish form. DIF was found in five items at negligible level and one item at B level according to LR technique, in six items according to each for signed-unsigned area indexes methods. In England-Turkey sampling for the fifth booklet; in two items at A

level and in four items at B level and in five items at C level according to the MH technique and in four items at negligible level and two items at B level according to LR technique, in six items according to each of signed-unsigned area indexes methods had DIF.

It is observed that as the linguistic and cultural differences increased between countries, the number of DIF items increased. The number of DIF items varied significantly according to the procedure used. The correlation coefficients for the same culture-different language between LR and MH were significant, Non-signed area indexes and Signed area indexes were significant at $\alpha = 0,01$. The correlation coefficients for the different culture-same language between LR and MH were significant at $\alpha = 0,05$, For the different culture and language LR and MH were significant at $\alpha = 0,01$ for the first booklet and Non-signed area indexes and Signed area indexes were significant at $\alpha = 0,01$ for the fifth booklet.

Generally; results of bias researchs indicated that the main possible reasons for DIF is due to differences in cultural relevance, linguistic differences and differences in curriculum.

Keywords

Differential Item Functioning, Mantel-Haenszel, Logistic Regression, Signed-Unsigned Area Indexes, PISA 2006.

İÇİNDEKİLER

KABUL VE ONAY.....	i
BİLDİRİM.....	ii
TEŞEKKÜR.....	iii
ÖZET.....	iv
ABSTRACT.....	iv
İÇİNDEKİLER.....	vii
KISALTMALAR DİZİNİ.....	xi
TABLolar DİZİNİ.....	xii
ŞEKİLLER DİZİNİ.....	xvi
BÖLÜM I: GİRİŞ.....	1
1.1. PROBLEM DURUMU.....	1
1.1.1. Klasik Test Kuramı	7
1.1.2. Madde Tepki Kuramı.....	8
1.1.3. Madde Karakteristik Eğrisi ve Fonksiyonu.....	10
1.1.4. Değişen Madde Fonksiyonu Belirleme Teknikleri.....	11
1.1.4.1. KTK Kapsamındaki Başlıca Değişen Madde Fonksiyonu Belirleme Teknikleri.....	14
1.1.4.2. MTK Kapsamındaki Başlıca Değişen Madde Fonksiyonu Belirleme Teknikleri	15
1.2. ULUSLARARASI ÖĞRENCİ DEĞERLENDİRME PROGRAMI – PISA (PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT)	18
1.3. PROBLEM CÜMLESİ.....	19
1.4. ALT PROBLEMLER.....	19
1.5.ARAŞTIRMANIN ÖNEMİ VE AMACI.....	20
1.6. SINIRLILIKLAR.....	22
1.7.İLGİLİ ÇALIŞMALAR.....	22

BÖLÜM II: YÖNTEM.....	34
2.1. ARAŞTIRMA MODELİ.....	34
2.2. ÇALIŞMA GRUBU	34
2.3. VERİLERİN ELDE EDİLMESİ.....	38
2.4. VERİ ÇÖZÜMLEME TEKNİKLERİ.....	38
2.4.1. PISA–2006 Verilerine Ait Betimsel İstatistiklerin Hesaplanması.....	40
2.4.2. Madde Tepki Kuramının Sayıtlılarının Kontrolü.....	41
2.4.3. Model Uyumu.....	46
2.4.3.1. Birinci Kitapçık İçin Model-Veri Uyumu	46
2.4.3.2. Beşinci Kitapçık İçin Model Veri Uyumu.....	47
2.4.4. Değişen Madde Fonksiyonu Analizleri.....	49
2.5. ARAŞTIRMADA KULLANILAN TEKNİKLER.....	49
2.5.1. Mantel – Haenszel Tekniği.....	49
2.5.2. Logistik Regresyon Yöntemi	51
2.5.3. Madde Tepki Kuramı Alan İndeksleri Tekniği.....	52
 BÖLÜM III: BULGULAR VE YORUM.....	 55
3.1.ALT PROBLEM 1 İÇİN ELDE EDİLEN BULGULAR.....	55
3.1.1. Aynı Kültür- Farklı Dil (Kanada örneği) için yapılan Değişen Madde Fonksiyonu analizleri.....	55
3.1.2. Farklı Kültür Aynı Dil (Avustralya-İngiltere)İçin Yapılan Değişen Madde Fonksiyonu Analizleri.....	58
3.1.3. Farklı Kültür- Farklı Dil (Türkiye-İngiltere) için Yapılan Değişen Madde Fonksiyonu Analizleri.....	61
3.2. ALT PROBLEM 2 İÇİN ELDE EDİLEN BULGULAR.....	65
3.3. ALT PROBLEM 3 İÇİN ELDE EDİLEN BULGULAR.....	71
3.3.1.Birinci Forma Ait Bulgular.....	71
3.3.2. İkinci Forma Ait Bulgular.....	81

BÖLÜM IV: SONUÇ VE ÖNERİLER.....	87
4.1 SONUÇLAR.....	87
4.1.1 Alt Problem 1’e İlişkin Sonuçlar.....	87
4.1.2 Alt Problem 2’ye İlişkin Sonuçlar.....	90
4.1.3 Alt Problem 3’e İlişkin Sonuçlar.....	90
4.2. ÖNERİLER.....	91
 KAYNAKÇA.....	 94
EKLER.....	104
EK 1: Uzman Kanısı Formu 1.....	108
EK 2: Uzman Kanısı Formu 2.....	115
EK 3: EZDIFF MH ve LR Analiz Çıktıları.....	125
EK 4: CALCBİAS Programı Analiz Çıktıları.....	144
ÖZGEÇMİŞ.....	148

KISALTMALAR DİZİNİ

KTK: Klasik Test Kuramı

MTK: Madde Tepki Kuramı

DMF: Değişen Madde Fonksiyonu

LR: Lojistik Regresyon

MH: Mantel-Haenszel

İAİ: İşaretli Alan İndeksleri Tekniği (tablolarda kullanılmıştır)

İSAİ: İşaretsiz Alan İndeksleri Tekniği (tablolarda kullanılmıştır)

TBDMF: Tek Biçimli Değişen Madde Fonksiyonu

ÇBDMF: Çok Biçimli Değişen Madde Fonksiyonu

PISA: Uluslararası Öğrenci Değerlendirme Programı

TABLOLAR DİZİNİ

Tablo 1: PISA'ya katılan ve Beşinci Kitapçığı Cevaplayan Öğrenci sayılarına Ait İstatistikler	35
Tablo 2: Soruların Kitapçıklara Dağılımı	36
Tablo 3: Araştırmadaki Çalışma Grubuna İlişkin Betimsel İstatistikler	40
Tablo 4: PISA-2006 Birinci Kitapçıkta yeralan 19 soruluk teste ait faktör analizi sonuçları	42
Tablo 5: PISA-2006 Beşinci Kitapçıkta Yeralan 21 Soruluk Teste Ait Faktör Analizi Sonuçları.....	44
Tablo 6: Birinci Kitapçık Model Veri Uyumu İçin Hesaplanan –2 Log (Olabilirlik) Değerleri	46
Tablo 7: Beşinci Kitapçık Model Veri Uyumu İçin Hesaplanan –2 Log (Olabilirlik) Değerleri	47
Tablo 8: Lojistik Regresyon Tekniğinde Değişen Madde Fonksiyonu Düzeyi İçin R^2 Ölçütleri	52
Tablo 9: Madde Tepki Kuramına Bağlı Değişen Madde Fonksiyonu Belirleme Yöntemlerinden Alan İndeksleri Tekniğinin Yorumlanmasında Kullanılan İç Ölçütlere İlişkin İstatistikler	54
Tablo 10: Aynı Kültür- Farklı Dil (Kanada Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Mantel – Haenszel Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları.....	56
Tablo 11: Aynı Kültür- Farklı Dil (Kanada Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Logistik Regresyon Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları	56
Tablo 12: Aynı Kültür- Farklı Dil (Kanada Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Alan İndeksleri Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları	57
Tablo 13: Farklı Kültür Aynı Dil (Avustralya-İngiltere Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Mantel – Haenszel Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları.....	59
Tablo 14: Farklı Kültür Aynı Dil (Avustralya-İngiltere Örneği) Bakımından	59

PISA 2006 Beşinci Kitapçık Alt Testinde Lojistik Regresyon tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları	
Tablo 15: Farklı Kültür Aynı Dil (Avustralya-İngiltere Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Alan İndeksleri Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları.....	60
Tablo 16: Farklı Kültür Farklı Dil (Türkiye-İngiltere Örneği) Bakımından PISA 2006 Birinci Kitapçık ve Beşinci Kitapçık Alt Testinde Mantel – Haenszel Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları	62
Tablo 17: Farklı Kültür Farklı Dil (Türkiye-İngiltere Örneği) Bakımından PISA 2006 Beşinci Kitapçık ve Birinci Kitapçık Alt Testinde Lojistik Regresyon tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları	63
Tablo 18: Farklı Kültür Farklı Dil (Türkiye-İngiltere Örneği) Bakımından PISA 2006 Birinci Ve Beşinci Kitapçık Alt Testinde Alan İndeksleri Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları	64
Tablo 19: Değişen Madde Fonksiyonu İçeren Maddelerin Tekniklere Göre Dağılımı.....	66
Tablo 20: Mantel – Haenszel, Logistik Regresyon ve Alan İndeksleri Tekniği İle Elde Edilen Değişen Madde Fonksiyonu Değerleri Arasındaki Sıra Farkları Korelasyonları	68
Tablo 21: Uzmanların PISA Fen Testi Maddelerinin Kültürler ve Diller Arasında Değişen Madde Fonksiyonu Göstermesinin Olası Nedenlerine İlişkin Görüş Dağılımları (Form 1).....	72
Tablo 22: Uzmanların PISA Fen Testi Maddelerinin Kültürler ve Diller Arasında DMF Göstermesinin Olası Nedenlerine İlişkin Görüş Dağılımları (Form 2)	81
Tablo 23: Farklı Formlara Yanıt Veren Uzmanlardan Elde Edilen Phi Katsayısı Değerleri	84
Tablo 24: Kültür ve Dil Farklılıklarına Göre DMF Gösteren Maddelerin Tekniklerdeki Sayısal Dağılımı	87

ŞEKİLLER DİZİNİ

Şekil 1: Madde Karakteristik Eğrisi	11
Şekil 2a: Maddenin Tek Biçimli Fonksiyonlaşması (TBDMF).....	16
Şekil 2b: Maddenin Çok Biçimli Fonksiyonlaşması (ÇBDMF).....	16
Şekil 3: PISA Birinci Kitapçıkta Yeralan 19 Soruluk Teste Ait Özdeğer-Bileşen Grafığı	43
Şekil 4: PISA Beşinci Kitapçıkta Yeralan 21 Soruluk Teste Ait Özdeğer-Bileşen Grafığı	45
Şekil 5: Alt Gruplar İçin PISA Beşinci Kitapçıkta Yer Alan 21 Soruluk Teste Ait Özdeğer-Bileşen Grafığı	45

BÖLÜM I

GİRİŞ

Bu bölümde problem durumu, Klasik Test Kuramı (KTK) ve Madde Tepki Kuramı (MTK) ile ilgili temel bilgiler, Değişen Madde Fonksiyonu Belirleme (DMF) teknikleri, problem cümlesi, araştırmanın amacı ve önemi, sınırlamalar ve araştırma ile ilgili olduğu düşünülen çalışmalara yer verilmiştir.

1.1 PROBLEM DURUMU

Sosyal bilimlerde yapılan çalışmalarda; bireylere ve olaylara ilişkin sayısal bilgileri içeren verilerin toplanması, düzenlenmesi, özetlenmesi, analizi ve bu analizler aracılığıyla elde edilen sonuçların yorumlanması ve karar verilmesi alana çok çeşitli katkılar sağlamaktadır.

Ülkemizde her yıl farklı amaçlarla birçok sınav düzenlenmektedir. Bu sınavların bir kısmı ulusal bir kısmı da uluslararası çaptadır. Özellikle uluslararası sınavların her geçen gün önem ve değer düzeyinin arttığını belirtmek yanlış olmayacaktır. Bu sınavlara katılan bireyler çeşitli açılardan farklılaşmaktadır. Sınav dilinin değişmesi, farklı demografik özelliklere ait olma, farklı sosyo ekonomik düzeye sahip olma, cinsiyet, kültürel özellikler gibi farklılıklar sınavlara katılan bireylerin performanslarını etkilemektedir. Bu nedenle ölçme aracından elde edilen puanların yorumlanmasında, bireylerin performanslarını etkileyebilecek değişkenler ile alt grupların özellikleri dikkate alınmalıdır. Eşdeğerlik ve karşılaştırılabilirlik; bu sınavlarda eşdeğerlik ve karşılaştırılabilirlik üzerinde çalışılması gereken önemli unsurlardandır. Uluslararası sınavların değerlendirilmesinde çeşitli yöntemlerin kullanılması ve yapılan karşılaştırma çalışmalarının geçerliğinin farklı yöntemlerle araştırılması gerekmektedir (Ercikan, 1998). Çünkü sınavların karmaşık yapısı değerlendirmelerin doğruluğunu etkileyebilmektedir.

Bir ölçme aracı geliştirildiğinde “ölçme aracı uygulanan bireylerde aynı özelliği ölçer” kabulü ile hazırlanır. Dolayısıyla ölçme aracından alınan puanların farklı alt gruplara bağımlı olmadığı kabul edilir. Ölçme sonuçlarının geçerliğine ilişkin kanıtlar

toplarken, ölçme aracını cevaplayan grup üyeleri için değişmez olması beklenir (Meredith ve Millsap, 1992; Millsap, 2005). Farklı kültür toplulukları, alt-kültür öbekleri, kültür çevreleri tüm dünyada yanyana yaşamakta, birçok bakımdan ve türlü derecelerde farklılıklar göstermektedir. Bu nedenle bir ölçme aracı geliştirilirken aracın farklı gruplar için geçerli olabileceği bir çerçeve sunulmalıdır. Bir başka deyişle test puanlarının yanlılık içermemesi önemsenmelidir. Eğer, ölçme aracının ölçtüğü özellikler farklı gruplar için geçerli değilse ve yanlılık gösteriyorsa, bu ölçme aracından elde edilen verilerle yapılan karşılaştırmaların, yorumların, alınan kararların geçerliği sorgulanabilir. Çünkü yapılan ölçmeye hata karışmıştır. Hatalı ölçmenin sonuçlarına bağlı olarak yapılan değerlendirmeler hatalı olur.

Davranış bilimlerinde yapılan ölçme ve değerlendirme çalışmalarında ölçmelere karışan sistematik hatalara yanlılık (bias) denilmektedir ve yanlılık çalışmaları geçerlik çalışmaları kapsamında ele alınır. Yanlılık; testlerin önemli psikometrik özelliklerinden olan geçerliği düşüren bir durumdur. Yanlılık çalışmalarında sıklıkla üzerinde durulan hata türü sistematik hatalardır. Davranış bilimlerinde alanda kullanılan ölçme araçlarının ölçmeye çalıştığı değişkenlerin birer psikolojik değişken olması nedeniyle sistematik hataların kaynaklarının belirlenmesi önem kazanmaktadır. Bu nedenle özellikle son yıllarda “değişen madde fonksiyonu (DMF) ve madde yanlılığı” konularının araştırılmasına yoğunlaşılmaktadır. Bunun yanı sıra gerek kültürlerarası karşılaştırmalarda, gerek aynı kültür içerisindeki gruplararası karşılaştırmalarda DMF ile yakından ilgili olan ölçme eşdeğerliği çalışmaları artmaktadır. Çok kültürlü çalışmalarda farklı kültürler için ölçme aracının eşdeğerlik düzeyi DMF’ nin varlığı ve yokluğu ile yakından ilgilidir (Van de Vijver ve Leung; 1997). Farklı alt gruplar için ölçme eşdeğerliğinin sağlanamaması, ölçme ve değerlendirmede yöntemsel bir problem olarak karşımıza çıkmaktadır. Çünkü davranış bilimlerindeki birçok çalışma gruplararası karşılaştırmaları içermektedir. Gruplararası karşılaştırmalarda üzerinde çalışılan değişken dışında farklılığa neden olacak değişkenler kontrol altına alınarak (sistematik hata kaynakları elimine edilerek), karşılaştırma yapılan değişken açısından grupların eşleştirilmesi amaçlanmaktadır. Bu durum sağlanamazsa gruplar arası yapılan karşılaştırmaların doğruluğu tartışmalı hale gelir. Bir test etme durumunda, aynı yetenek düzeyinde olup, farklı gruplarda bulunan cevaplayıcıların, maddelerin cevaplanmasında avantajlı ya da dezavantajlı durumda

olmaması gerekir (Childs,1990). Çünkü farklı alt gruplardan gelen aynı yetenek seviyesindeki bireylerin test puanlarında farklılıklar; bireylerin belli bir yetenek ölçüsünde sıralamalarında adil olmayan, yanlış, kararlar alınmasına neden olur.

Zaman içinde aynı alt gruptan alınmış farklı ölçümlerin ya da aynı zamanda farklı alt gruplardan alınmış ölçümlerin birbirine denkliğinden emin olunması, test geliştirme sürecinin önemli yönlerinden birisidir. Bir test heterojen bir gruba uygulanmışsa, grup içinde yer alan alt gruplarda ölçme işlemine ait özelliklerin birbirine eşit olduğunun gösterilmesi gerekir. Bu bakımdan madde yanlılığı çalışmalarının önemi giderek artmaktadır. Çünkü eğitim ve psikolojide kullanılan ölçme araçlarının geçerlik kanıtlarından birisi de test ve madde yanlılığı belirleme işlemleridir. Test maddelerinin, aynı yetenek düzeyinde olan fakat farklı gruplardan gelen bireylerin cevapları açısından farklı özellikler göstermemesi; grupların ölçülen özellik bakımından doğru tanımlanması önem kazanmaktadır. Test maddelerinin; farklı kültür, dil, cinsiyet veya sosyo-ekonomik düzey gruplarında bulunan aynı yetenek düzeyindeki bireylerden herhangi birinin lehine ya da aleyhine işlemiyor olması gerekir. Böylesine istenmeyen bir durumun varlığı; test sonuçlarının geçerliğinin tartışılmasına neden olabilir.

Madde yanlılığı; aynı yetenek düzeyinde olan fakat farklı gruplardan gelen iki kişinin bir maddeyi doğru cevaplama olasılıklarının aynı olmaması durumu olarak tanımlanmaktadır (Osterlind, 1983; Adam ve Rowe 1988; Tittle, 1988; Cole ve Moss, 1989; Hambleton ve Rogers,1996; Raju ve Ellis, 2002; Zumbo,1999). Bir başka ifade; madde parametrelerinin kestirim değeri ile gerçek değeri arasındaki farklılıktır. Bu tanımdan anlaşılacağı gibi, madde yanlılığı analizleri her bir test maddesinin aynı evrenden elde edilen farklı alt gruplarda benzer şekilde işleyip işlemediği sorusu üzerine odaklanmaktadır. (Mellenberg, 1983)

Tarihsel olarak incelendiğinde madde yanlılığı konusunun çok yeni bir konu olmadığı anlaşılmaktadır (Camilli ve Shepard, 1994). Yanlılık araştırmaları 1910'lu yıllarda, Binet'in düşük sosyo - ekonomik düzeylerden gelen çocukları test etmesiyle başlamıştır (Harold, 1989; Camilli ve Shepard, 1994) Binet anadili İngilizce olmayan ve sosyo ekonomik düzeyi düşük çocuklarda bazı zeka testi maddelerinin zihinsel

kapasiteden ziyade çocukların evde veya okuldaki kültürel eğitim-öğretimlerinin etkileriyle ilişkili olduğunu saptamıştır. Bazı kategorilerdeki maddeleri testten çıkarmıştır. Sovyet psikolog Lev Vygotsky’de kültürler arası çalışmalarda gençlerin problem çözme becerilerinin doğasında kültürel önyargıların varlığını vurgulamış ve kendi gibi psikolog olan arkadaşı Alexander Luria (1978) ile problem çözme stratejilerinde kültürel önyargıların nedenlerini (olası yanlılık nedenleri) belirlemeye çalıştıkları bir araştırma yapmıştır. Vygotsky ve Luria, yapılan değerlendirmelerin çoğu zaman referans alınan kültürün özelliklerine göre şekillendiğini ve bu nedenle elde edilen sonuçların kültürlere göre yanlılık gösterdiğini belirtmişlerdir.

Madde yanlılığı istatistikleri, maddelerin bir gruptaki cevaplayıcıları aynı yetenek düzeyindeki diğer bir alt gruptaki cevaplayıcılardan farklı bir şekilde ölçüp ölçmediğini göstermeyi amaçlamaktadır. Eğer bir farklılık varsa, bu farklılığın gerçekten testin kapsamından mı yoksa başka bir kaynaktan mı olduğu araştırılmalıdır. Madde yanlılığının araştırılması süreci hem istatistiksel yönden, hem de bu maddelerin farklılığının kaynağı açısından ne olduğunun araştırılmasını içerir. Farklılığın sadece istatistiksel olarak gösterildiği duruma “değişen madde fonksiyonu (DMF)” adı verilir. Yanlılık genel olarak DMF sonuçlarına dayalı olarak değerlendirilmektedirler. Zumbo (1999, s:6) DMF’yi yanlılık analizlerindeki yeni standart olarak belirtmiştir. Ancak DMF doğrudan test yanlılığının kanıtı olarak değerlendirilmemelidir. Geliştirilecek ölçme aracı farklı gruplarda kullanılacaksa DMF analizleri ölçme aracına madde seçim sürecinin en önemli bölümlerinden biri olmaktadır (Crocker ve Algina, 1986; Mellenberg, 1983).

Ölçmelerin yanlılığı DMF’den daha karmaşıktır. Ölçme ve değerlendirmede yanlılık basit istatistiksel teknikler ve klasik test teorisi metodları ile kolayca belirlenemezler. Yanlılığın belirlenmesinde grup - madde etkileşimleri için yapılan varyans analizleri ve klasik test teorisi ile elde edilen istatistiksel anlamlılık (p) değerleri etkili değildir (Idaszak, Bottom ve Drasgow, 1987). Bu biçimde yapılan analizlerde ortaya çıkan DMF doğrudan yanlılık olarak yorumlanmamalıdır. Bu durum maddenin alt gruplardaki gerçek etkisinden de (gerçek farklılığından=true difference) kaynaklanabilir. Buradan da anlaşılacağı gibi DMF’nin ortaya çıkmasının iki nedeni vardır. Bunlar; alt gruplar arasındaki gerçek farklılık (true difference) ve madde

yanlılığıdır (item bias) (Camilli ve Shepard,1994). Her zaman bir testte bir grubun puanının başka bir grubun puanından farklı olması testin ya da incelenen maddelerin yanlı olduğu anlamına gelmez. Örneğin ergenlik döneminde cinsiyet grupları arasındaki gelişimsel farklılıklar, cinsiyet bağlamına ilişkin büyüme eğrileri arasındaki farklılıklar gerçek farklılıklardır. Bir yüzme yarışında erkeklerin kızlara göre daha iyi derece yapması, bu yarışın erkeklerin lehine olduğu anlamına gelmez. Yanlılık çalışmaları gruplar arası gerçek farklılık dışındaki farklılıklara neden olan etkenler üzerine yoğunlaşır. Yanlılık sistemattir; çünkü grubun belirli üyeleri için test sonuçlarında belirli bir farklılık yaratır.

Yanlılığın olası nedenleri temel olarak dil, kültür ve program çevresinde yoğunlaşmaktadır. Çalışmalardaki olası yanlılık nedenleri; sosyo-ekonomik farklılıklardan, müfredat kapsamındaki farklılıklardan, yetersiz çevirilerden, madde içeriğinin belirli kültürlere aşinalığı/uygunluğundan, negatif, edilgen ve mecazi anlam içeren cümlelerin kullanımından, madde yazılış tarzı ile ilgili kültüre özgü imalardan, kelimelerin değişken içeriksel anlamlara sahip olmasından, gramer hatalarının tespitini gerektiren maddelerden...vs. biçiminde sıralanmaktadır (Van de Vijver ve Tanzer, 1997; Uiterwijk ve Valen, 2005; Ercikan, 1998; Grisay ve diğerleri, 2007). Hambleton ve Jones (1994) bir maddenin kapsamı ya da dilinin alt gruplarda farklılaşmasından, madde yapısı ya da biçiminin farklı alt gruplarda farklılaşmasından ya da madde yapısı ve biçiminin farklı gruplardaki zorluk derecelerinin değişmesinden dolayı farklı fonksiyonlaşabileceğini belirtmiştir. Uluslararası çalışmalarda ölçme aracı kaynak dilden başka dillere çevrilir. Çeviri çalışmaları sonucu ortaya çıkan ölçme aracı kimi zaman bazı ifadelerin karşılığını kaynak kültürde bulamaması ya da farklılaşması nedeniyle geliştirilen ilk ölçme aracından farklı olabilmektedir. Bu nedenle kültürler arası kullanılacak olan ölçme araçlarında en azından çeviriden kaynaklanan farklılığı en aza indirmek gerekir. Rodrigues ve Bau (1956) iyi bir çevirinin; anlam, içerik, vurgusu ile orijinal versiyonun stilini tamamen yansıtması gerektiğini vurgulamıştır (aktaran: Ercikan; 1998). Uzmanlar çevrilen bir ölçme aracının çeviri kalitesini ortaya koyabilmek adına anlamsal karşılaştırmalar; odak ve referans grup için elde edilen yetenek değerleri, ölçme aracının farklı formlarına verilen yanıtlar ve performans ölçümleri ile ilgilenmektedir (Brislin; 1973). DMF analizi daha çok maddelerin farklı dil gruplarında ölçme eşdeğerliğini sağlamak amacıyla yürütülmektedir. Kültürler

arası yapılan birçok çalışmada; uyarlama çalışmaları ile oluşan kültüre özgü anlamsal farklılıklar, farklı dillere çevrilen cümlelerin farklı zorluktaki kelime gruplarından oluşması, çeviriden kaynaklanan cümle karışıklıklarının...vs madde yanlılığına neden olduğu, uyarlanan ölçme araçlarının eşdeğer olmadığı yönünde sonuçlara ulaşılmıştır (Beaton, 1998; Çet, 2006; Yıldırım, 2006; Ercikan ve Koh, 2005; Ercikan ve arkadaşları, 2010). Her ne kadar değişmezlik çalışmalarının yapılmasına özen gösterilse de ya da kültürler arası çalışmalarda bu tür çalışmaların yapıldığı ölçme araçları kullanılsa da; yanlılıklar ve olası kültür etkileri ölçek maddelerinin anlamlarını değiştirmektedir. Tüm bu olası etkenler farklı dil ve kültür gruplarından gelen birey performanslarının karşılaştırılmasında problem yaratabilmektedir.

Madde ve test yanlılığı çalışmalarında dilin önemine odaklanmak yararlı olabilir. Çalışmalar bir ölçme aracının iç yapısının çevirinin bir fonksiyonu olarak değişebileceğini göstermiştir (Marin ve Vanoss Marin, 1991). Nitekim bir kelime bile başlı başına yanlılık nedeni olabilir. Bazen de kültürler arası farklılıklar ölçme aracının ya da ölçme aracında yer alan bazı maddelerin kaynak kültürde farklı işlemesine neden olabilmektedir. Ercikan (2002) yanlılığın olası nedenlerinin sadece karşılaştırılan gruptan değil, değerlendirmenin tipi ve amacına bağlı olarak da farklılaşabileceğini vurgulamıştır. Çevirisi bulunan ölçme araçlarına yönelik yapılan DMF analizleri sonucu, DMF gösteren maddelerin çokluğu da dikkat çekmektedir. Örneğin, Gierl, Rogers ve Klinger (1999) Fransızca ve İngilizce çevirisi bulunan 6. sınıflar sosyal bilgiler başarı testinde yer alan 50 maddenin 26'sında (52%); Ercikan (1999) Kanada, İngiltere ve Fransız örnekleminde yapmış olduğu çalışmada 140 TIMSS fen maddesinin 58'inde (41%) ve 158 matematik maddesinin de 29'unda (18%) yüksek derecede DMF bulunduğunu ifade etmiştir. Benzer biçimde Salano, Flores ve Lii (2006) yapmış oldukları genellenebilirlik çalışmasında öğrenci, madde ve dil çaprazlanmış deseninde her üçünün etkileşiminden kaynaklanan artık varyansın büyüklüğüne dikkat çekerek; bireyin performansı üzerinde dil etkisinin yapı geçerliğini tehdit eden bir durum olduğu sonucuna ulaşmışlardır. Çalışmalarda ortaya konan bu gibi örnekleri irdeleyen madde yanlılığı çalışmaları, ölçme aracının niteliklerinin ortaya konulması ve grubun ölçülen özelliğe ilişkin doğru değerlendirilmesi bakımından kritik bir öneme sahiptir.

DMF analizleri iki farklı kurama dayanır. Bu kuramlar Klasik Test Kuramı (KTK) ve Madde Tepki Kuramıdır (MTK). Aşağıdaki bölümde bu kuramlara ilişkin temel bilgiler ve bu kuramlara ait belli başlı DMF belirleme teknikleri ile ilgili bilgiler yer almaktadır.

1.1.1. Klasik Test Kuramı

Ölçme aracı geliştirme ve değerlendirmenin psikometrik temellerini Klasik Test Kuramı oluşturmaktadır (Embretson ve Reise; 2000). Psikolojik ölçme tarihinin başlangıcından itibaren test geliştirme, analiz ve psikolojik ölçeklerin puanlanmasında yaygın olarak kullanılan kuramdır. Gözlenen puanları, gerçek puan ve hata puanlarının toplamı olarak modelleyen bu kuramın bulgularının yardımı ile madde analizleri, güvenirlik ve geçerlik çalışmaları yürütülmektedir.

Klasik test kuramının temel denklemi eşitlik 1'deki gibi ifade edilmektedir.

$$X=T+E \quad \text{(Eşitlik 1)}$$

X: Gözlenen puan;

T: Gerçek puan;

E: Hata Puanı

KTK' nın temel denklemi olan bu eşitlikte; gerçek puanlar ile gözlenen puanlar arasındaki fonksiyonel bağıntıdan yararlanarak doğrusal model kestirilmektedir. Eşitlik 1' e göre bir bireyin ölçülen özelliğine ilişkin gözlenen puanı, gerçek puan ve hata puanlarının bileşiminden oluşmaktadır (Crocker ve Algina, 1986). Bu, KTK'nın en temel varsayımıdır. Eşitlikte hata puanı değeri azaldıkça bireyin gözlenen puanı gerçek puanına yaklaşacaktır. Ancak hatasız ölçme olmadığından; gözlenen puan gerçek puana pratikte hiçbir zaman eşit olamayacaktır. Bireyin gerçek puanı hipotetik madde evrenindeki puandır. Bir bireye aynı ölçme aracının (ya da bu aracın paralel formunun) sonsuz sayıda uygulandığı varsayılırsa, birey her seferinde farklı gözlenen puanlar alacaktır. KTK'na göre ölçme aracı ile ölçülen birey özellikleri sabittir ve gözlenen puanlardaki değişim, çeşitli faktörlerin sonucu olan tesadüfi hatalardan kaynaklanır (Crocker ve Algina; 1986).

KTK'nın en temel avantajı varsayımlarının karşılanması nispeten kolay olması ve farklı test etme durumlarına uygulanmasının kolay olmasıdır (Hambleton ve Jones, 1993). KTK bir takım sınırlılıklara sahiptir. Bu kuramın en temel sınırlılığı; maddeye, teste veya bir bireye ait parametrelerin testi alan gruba bağımlı olmasıdır. Bu sebeple maddeye ilişkin karakteristikler çok fazla bilgi sağlamaz. KTK'ya göre bir örneklemdaki yetenek düzeyi ve dağılımı, madde güçlük indeksi ve madde ayıricılık indeksi gibi madde ve test istatistiklerini önemli ölçüde etkilemektedir. Bir diğer sınırlılığı ise; bireylerin yeteneklerinin kendilerine uygulanan maddelere bağımlı olmasıdır. Bir birey, aynı özelliği ölçen farklı güçlükteki testlerden farklı puanlar alabilir (Hambleton ve Swaminathan; 1985). Daha açık bir ifadeyle; bireyin yeteneği testten aldığı puan ile ilişkilendirildiğinden; aynı özellik ölçüldüğü halde, farklı zorluk derecelerine sahip test sonuçlarına bakılarak bireyin aynı özellik bakımından farklı yetenek düzeyine sahip olduğu sonucuna ulaşılır. Başka bir ifadeyle KTK'da bireylerin aldıkları puanlar o testin güçlük düzeyine göre değişmektedir (Lord ve Novic, 1968).

KTK'na dayalı olarak ölçme aracı geliştirilirken grubun yetenek düzeyine bağlı olarak elde edilen madde istatistiklerinin değişmesi gibi temel sınırlılıkları alternatif kuram arayışlarına yol açmış ve bu doğrultuda madde tepki kuramı (MTK) geliştirilmiştir. MTK ile ilgili temel bilgiler aşağıda yer almaktadır.

1.1.2. Madde Tepki Kuramı

Örtük Özellikler Kuramı olarak da bilinen Madde Tepki Kuramı (MTK), Klasik Test Kuramına alternatif olarak ortaya çıkmıştır. 1930'lu yıllarda ilk olarak ortaya atılmış, konu ile ilgili asıl çalışmalar ise 1950'li yıllarda yapılmaya başlanmıştır. MTK ile birlikte ölçmenin psikometrik temelleri değişime uğramıştır. Bu kuramın en güçlü yanı; yetenek kestirimlerini uygulanan maddelerden bağımsız olarak vermesidir. Bireyler farklı sorularla test edilseler bile MTK farklı kişilerin yeteneklerini kestirebilmemiz için standart bir çerçeve sunmaktadır (Hambleton; Swaminathan ve Rogers, 1991,s:11).

Klasik Test Kuramındaki bazı sınırlılıkları ortadan kaldırmak adına üzerinde çalışılan Madde Tepki Kuramı, her ne kadar daha ileri bir kuram olarak görülse de

uygulanmasında ve hesaplanmasındaki zorluklar sebebiyle, Klasik Test Kuramının yerini tam anlamıyla alamamıştır. Günümüzde hala KTK kolay ve uygulanabilir olması sebebiyle daha çok tercih edilmektedir. MTK, test geliştirme, soru bankası oluşturma, bireye uyarlanmış test geliştirme, madde yanlılığının belirlenmesi, seçenekleri ağırlıklandırma ve test eşitleme konularında karşılaşılan sorunlara alternatif çözümler sunmaktadır (Hambelton ve Swaminathan, 1985).

Madde tepki kuramı (MTK), ölçülen özellik bakımından farklı yetenek düzeylerindeki cevaplayıcıların maddeyi nasıl yanıtladıklarına ilişkin matematiksel bir model sunan bir kuramdır (Crocker ve Algina, 1986).

Madde Tepki Kuramına göre:

1. Bir madde, o maddeyi cevaplandırabilecek herhangi bir cevaplayıcı grubundan bağımsızdır.
2. Bir bireyin yetenek düzeyi, o bireye uygulanan herhangi bir madde grubundan bağımsızdır.

Yukarda belirtilen iki özellik, KTK'nın sınırlılıkları olarak belirtilen özelliklerin tam tersidir. Bu sayede sınırlılıkların ortadan kaldırıldığı varsayılmaktadır. Elde edilen parametrelerin gruba veya bireye uygulanan maddelere bağımlı olmaması madde tepki kuramını daha avantajlı hale getirmektedir.

MTK için iki temel varsayımdan söz edilmektedir (Hambleton ve Swaminathan, 1985):

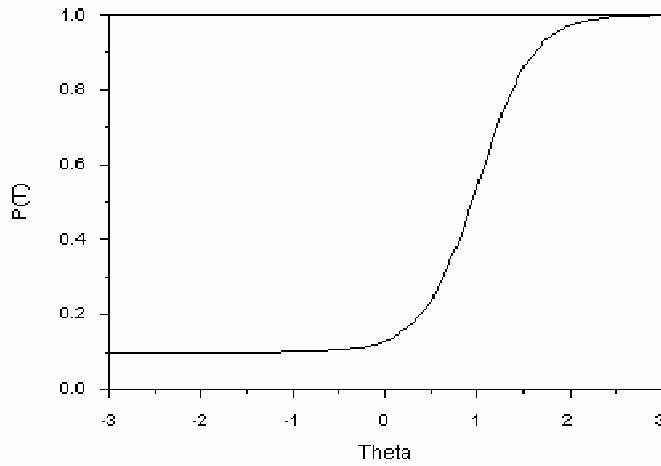
a) Tek Boyutluluk: Bu varsayım testi oluşturan maddelerin, yalnızca tek bir yeteneği ölçeceğini kabul etmektedir. Başka bir ifadeyle; bireyin testteki performansını belirleyen faktörün testin ölçmeye çalıştığı faktör olmasıdır. Ancak bu varsayımın karşılanması bireyin performansını etkileyen birden fazla faktör olduğundan oldukça zordur. Bireyin bilişsel ve kişisel özellikleri ile motivasyon düzeyi, sınav kaygısı, cevaplar arasında çelişkiye düşüp tartışma, vs. davranışları test performansını etkiler ve çoğunlukla kontrol edilemez. İçerisinde tek bir baskın yeteneğin yer aldığı düşünüldüğünde bir testin tek boyutluluğundan söz edilebilir. Testin tek boyutlu olup

olmadığının saptanmasında çoğunlukla faktör analizi tekniğinden yararlanılmaktadır. Kuramın, her yeteneği açıklayan tek bir faktörün var olması şeklindeki bu varsayımın teoride sağlanması zordur. Örneğin bir İngilizce yeterliliği sınavının tek bir faktörü ölçtüğünü söylemek yanlış olur; çünkü bu yeterlilik sınavı okuduğunu anlama, kelime bilgisi, dinleme, yazma gibi faktörler içerir. Benzer biçimde çoktan seçmeli bir test ile yoklanan herhangi bir özelliğin yanı sıra, öğrencinin test çözme becerisi de ölçmede başka bir faktör olarak karşımıza çıkabilir.

b) Yerel Bağımsızlık: Bu varsayım aynı yetenek düzeylerinde testi oluşturan maddelerin birbirinden bağımsız olduğunu tanımlamaktadır. Başka bir ifadeyle, testteki maddelerden birine verilen cevap bir başka maddeye verilecek cevabı etkilememelidir. Bu varsayımın test edilmesinde kullanılan yollardan biri farklı yetenek düzeyleri için ikiye ayrılmış olarak bireylerin cevaplarının dikkate alınması suretiyle χ^2 (ki-kare) testi yapmaktır. Bir testin tek boyutlu olduğunun belirlenmesi aynı zamanda aynı yetenek düzeyindeki cevaplayıcılar için maddeler arasındaki kovaryansın sıfır olduğu anlamına gelmektedir. Bu durum, tek boyutluluk varsayımının karşılanması durumunda yerel bağımsızlık varsayımının da karşılandığını göstermektedir (Hambleton, Swaminathan ve Rogers, 1991, 11).

1.1.3. Madde Karakteristik Eğrisi ve Fonksiyonu

Madde tepki kuramında bireyin bir maddeye ilişkin performansı ile ölçülen özelliği arasındaki ilişkinin matematiksel olarak gösterilebileceği varsayılmaktadır. Bu matematiksel tanım madde karakteristik fonksiyonu ve eğrisidir (Crocker-Algina, 1986). Madde karakteristik fonksiyonu, madde puanının θ (theta) yetenek vektörü üzerindeki regresyonudur ve “ $P_i(\theta)$ ” şeklinde ifade edilir. Ölçülen örtük özellik tek boyutlu olduğunda, fonksiyonu belirleyen parametreler sabit olur ve fonksiyon madde karakteristik eğrisi adını alır (Lord, Novick, 1968). $P_i(\theta)$ fonksiyonu gözlenen yanıtlardan hareketle gözlenmeyen özelliklerin kestirilmesinde kullanılır; doğrusal olmayan bir eğridir ve tam örtük uzay tanımlanmışsa farklı gruplarda değişmezliğini korur. Madde karakteristik eğrisi, belirli bir yetenek düzeyindeki kişinin maddeyi doğru yanıtlandırma olasılığı hakkında bilgi verir (Hambleton ve Swaminathan, 1985).



Şekil 1: Madde Karakteristik Eğrisi

Şekil 1’de yatay ekseninde yer alan “theta” yeteneği; dikey ekseninde yer alan “ $P(\theta)$ ” ise maddeye doğru yanıt verme olasılığını temsil etmektedir. Madde karakteristik eğrisinin önemi, madde güçlüğü ve ayıricılığı istatistiklerinin tersine, örtük özelliklerin doğru cevap verme olasılığına nasıl bağlı olduğunu görmemize izin vermesidir. (Crocker-Algina, 1986)

Bir maddenin DMF göstermemesi için, a, b ve c madde parametreleri yardımı ile çizilen madde karakteristik eğrisinin iki farklı evrende aynı olması gerektiği belirtilmektedir (Raju, Laffitte ve Byrne, 2002). Gruplar arası karşılaştırmalar yapıldığında aynı yetenek düzeyindeki grupların aynı maddeler üzerindeki madde karakteristik eğrileri karşılaştırılabilir. Yetenek düzeyleri aynı olan gruplar için elde edilen madde karakteristik eğrileri farklı ise söz konusu maddede gruplar arası DMF ‘nın varlığından söz edilebilir.

1.1.4. Değişen Madde Fonksiyonu Belirleme Teknikleri

KTk kapsamında olanlar

- Madde ayıricılık gücü tekniği
- Madde güçlüğü tekniği
- Varyans analizi tekniği
- Ki-kare tekniği

- Mantel-Haenszel (MH) tekniği
- Madde güçlük dönüşümü (MGD) tekniği
- Lojistik regresyon (LR) tekniği

MTK kapsamında olanlar

- Alan indeksleri tekniği
- Madde parametreleri tekniği
- Olabilirlik oran analizi

DMF belirlemede kullanılan KTK yöntemleri MTK kapsamında olan yöntemlere göre daha az varsayıma sahiptir ve görelî olarak uygulamaları daha kolaydır ancak uygulamaların sonuçları örnekleme özgü ölçme eşdeğerliğini sağlamak için yeterli olmamaktadır. MTK madde yanıtlarını ölçerken birkaç parametrik modellemeyi kullanarak klasik test kuramına göre avantajlı ve güçlü bir teknik olabilmektedir (Embretson, 1996). Ölçme eşdeğerliği kapsamında ele alınılabilecek test ve madde yanlılığını belirleme metodlarının mümkün olduğunca kullanılması gerekmektedir (Kristjansson ve diğerleri, 1997). Test ve madde yanlılığı belirlemeye dönük çalışmalar geçerlik çalışmaları kapsamında incelenmektedir. Madde yanlılığına yönelik analizler, testi alan grubun alt grupları arasında karşılaştırmalar yapılması tekniğine dayanır (aktaran: Öğretmen ve Doğan, 2004, Ironson ve Craig, 1982; Shepard, Camilli ve Williams, 1985). Eğer bir test çok farklı örneklemelerin karşılaştırılması için kullanılacaksa, değişen madde fonksiyonu analizleri madde seçimi sürecinin önemli bir parçası haline gelir. Çünkü, bu ölçme işleminde test uzmanı için en önemli kriter farklı alt gruplardan gelen ve testi alan bireyler için en adil ve en doğru sonuçları elde etmek olacaktır.

Yöntemlere ait tarihsel sürece göz atıldığında; DMF yöntem çalışmalarının 1960'ların başlarında başladığı görülmektedir. DMF belirleme yöntemlerinin kullanımına ilk etapta özellikle farklı kalıtsal özelliklere sahip olan öğrenciler arasındaki test performansı arasındaki farklılığın sebeplerini araştırmak ve kültürel farklılıkları araştırmak için yeni yöntemler geliştirmek adına başlanmıştır. Cardall ve Coffman 1963 yılında SAT testini alan çoğunluğu beyaz ve siyah deneklerden elde edilen verilerde test ve madde etkileşimlerini ortaya koymak için varyans analizi

uygulamışlardır. Varyans analizi çalışmaları ile kültürel farklılıkların ortaya çıkarılmasına aynı biçimde Cleary ve Hilton (1968) ile Angoff ve Sharon (1974) devam etmiş; ancak bu yöntem yaygınlaşmamıştır.

Angoff 1972 yılında kültürel farklılıkları çalışmak için delta plot ya da dönüştürülmüş madde güçlüğü olarak bilinen yöntemi önermiştir. Mantığının ve pratikte kullanımının basit olduğu bu yöntem kültürel farklılıkların çalışılmasında kısa sürede popüler olmuştur.

Lord'un madde tepki kuramı ile ilgili kitabının yayınlanmasının ardından Shepard (1981) DMF Rash indeksi ile DMF'ye ait dönüştürülmüş madde güçlüğü indeksi arasında .99 dan yüksek korelasyon bulmuştur. Sonraki yıllarda klasik test kuramına bağlı yöntemlerin yanı sıra madde tepki kuramına dayalı yöntemler üzerine çalışmalar yapılmaya başlanmıştır.

Sheuneman (1979) DIF değerlendirilmesinde ki kare yöntemine benzer bir yöntem önermiştir. Bu yöntem Baker (1981) tarafından sonucun örneklem büyüklüğü ve örneklem dağılımından fazlaca etkilenmesinden dolayı eleştirilmiştir. Holland ve Thayer (1988) Scheuneman' ın tekniğini geliştirmiş ve Mantel Haenzel prosedürleri kullanılmaya başlanmıştır. Kültürler arası çalışmalar arttıkça ve kuramlar genişledikçe yeni yöntemler kullanılmış ve aynı veri seti üzerinde yöntemlerin etkililiği çalışmaları üzerinde durulmuştur. Yapılan DMF çalışmaları incelendiğinde hemen hemen her tekniğin kullanıldığı ve birbiriyle karşılaştırıldığı çalışmalara rastlamak mümkündür.

Bu çalışmada KTK teknikleri içerisinde yer alan Mantel-Haenszel (MH) tekniği, Lojistik regresyon (LR) tekniği ve MTK kapsamında yer alan madde karakteristik eğrileri arasındaki alanlar hesaplanarak elde edilen alan indeksleri tekniğine bağlı değişen madde fonksiyonu belirleme teknikleri kullanılmıştır. Bu bölümde yaygın biçimde kullanılan tekniklerden bir kısmına ilişkin açıklamalar bulunmaktadır. Çalışmada DMF belirlemek için kullanılan üç teknikle (MH, LR, alan indeksleri yöntemi) ilgili daha geniş açıklamalar yöntem kısmında yer almaktadır.

1.1.4.1. KTK Kapsamındaki Başlıca Değişen Madde Fonksiyonu Belirleme Teknikleri

Ki-Kare Tekniği: Bu tekniğe göre bir maddenin yansızlığı, aynı test puanı kategorisinde fakat farklı alt gruplarda yer alan cevaplayıcıların, maddeyi doğru cevaplama oranlarının aynı olması olarak tanımlanmaktadır. Ki-kare tekniğinde, gözlenen puanlar ölçeği bir kaç kategoriye ayrılmaktadır. Her bir kategori için de alt gruplar bir maddeyi doğru cevaplama oranları açısından karşılaştırılır. Eğer oranlar gruplar arasında farklılık gösteriyorsa bu durum madde yanlılığı için bir kanıt olarak kullanılabilir.

Mantel-Haenszel Tekniği: KTK kapsamında DMF belirlemede madde tepki fonksiyonunun özel formlarına ihtiyaç duymaması, küçük örneklem için de kullanılabilmesi ve karmaşık hesaplamalar gerektirmemesi yönüyle sıkça kullanılan; belli avantajlar sağlayan Mantel-Haenszel (MH) yaklaşımı bir ki-kare tekniğidir. Bir serbestlik derecesindeki ki-kare dağılımıdır. Diğer ki-kare tekniklerinden güçlü olan yanı maddedeki yanlılık miktarının ölçüsünü hesaplamayı sağlamasıdır. Yanlılık ölçümü Mantel Haenszel oranı olarak verilir. Madde yanlılığını belirlemek için daha karmaşık olan örtük özellikler tekniklerine güçlü bir alternatiftir.

Bu teknikte benzer test puanı kategorisindeki farklı alt gruplarda yer alan cevaplayıcılar iki grup içerisinde yer alırlar. Bu gruplar; odak ve referans gruplardır. Odak grup; maddenin avantajlı ya da dezavantajlı olup olmadığı araştırılan (ilgilenilen); referans grup ise karşılaştırmaya esas olan gruptur (Ayala, 2009).

Teknikte odak ve referans gruplarının her ikisi de kendi içinde yanıtlayıcıların yeteneklerini eşitlemek için iki ya da beş alt gruba bölünür. Denkleştirilmiş iki gruptaki performans olasılık oran değerleri dikkate alınarak karşılaştırılır. Olasılık oranı bir gruptaki yanıtlayıcıların bir maddeye doğru cevap verme olasılığının diğer gruptan daha yüksek olup olmadığını belirler. Olasılık oranının daha kolay yorumlanabilmesi için bu oran “delta ölçeği” adı verilen bir ölçek değerine dönüştürülür. Hesaplanan değer, 0’a eşitse maddede DMF yoktur. Hesaplanan değer 0’dan büyükse madde odak grup lehine, 0’dan küçükse referans grup lehine çalışmış olduğuna karar verilir.

Lojistik Regresyon (LR) Tekniđi: Lojistik regresyon analizleri odak ve referans gruptaki yanıtlayıcıların belli bir ölçüm kriterleri çerçevesinde bir maddeye doğru cevap verme olasılıđını hesaplamaya dayanır. Madde bağımlı deđişken olarak düşünülür. Bağımlı deđişken test maddesinin 0-1 şeklinde kodladığı deđerden oluşur. Odak ve referans grup bağımsız deđişkendir. Bağımsız deđişken ise ikili veya sürekli veri niteliğindedir. Puanlar (madde puanları) üzerinde grup etkisinin anlamlı olup olmadığına bakılır. Eđer madde grup etkileşimi anlamlı ise bu maddenin yanlı olabileceđine karar verilir.

1.1.4.2. MTK Kapsamındaki Başlıca Deđişen Madde Fonksiyonu Belirleme Teknikleri

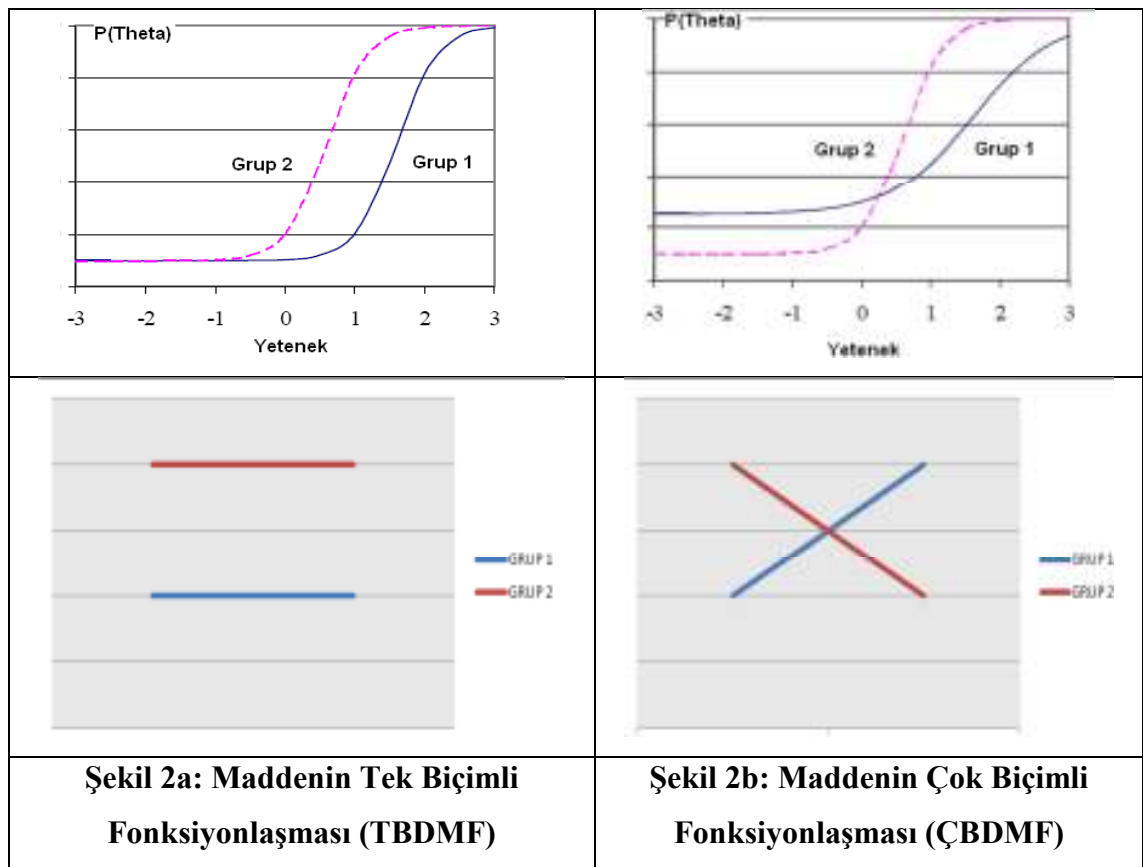
Madde parametrelerinin bir gruptan diđerine deđişkenlik göstermesi; parametrelerin sabit olmaması veya örnekleme bağımlı olması nedeniyle klasik test kuramı ile yapılan madde yanlılıđı analizlerinin, hatalı yorumlara yol açacağı öne sürölmektedir (Lord ve Novic 1968; Hambleton, and Swaminathan 1991; Baker 2001). Madde tepki kuramı deđişen madde fonksiyonun (DMF) incelenmesi ve daha az yanlı ölçeklerin geliştirilebilmesi için tercih edilmesi daha muhtemel bir yöntem sunmaktadır. Çünkü belli koşullar altında bu modeller evrene bağımlı olmayan sabit parametreler üretirler. Bu nedenle, kuramsal olarak da parametreler düzensizlik veya madde yaygınlığına ilişkin oran farklılıklarından da az etkilenirler. KTK kuramından farklı olarak, MTK, madde güçlüđünü yetenek düzeyine bağlar; maddeler ve kişiler ortak bir ölçeđe yerleştirilirler (Öğretmen, 2006). Aşağıda MTK'ya bađlı olan ve çalışmalarda sıklıkla karşılaşılan tekniklerden bazılarının açıklamalarına yer verilmiştir.

Alan İndeksleri Tekniđi (Madde Karakteristik Eğrileri Arasındaki Alan Hesaplamalarına Bađlı DMF Belirleme): MTK kuramı yöntemlerinde aynı yetenek düzeyine sahip odak ve referans grubun bir maddeye doğru cevap verme olasılıkları arasındaki farklar incelenir. Başka bir ifade ile MTK ile yapılan madde yanlılıđı belirleme yöntemlerinde alt gruplardan (odak ve referans grup) elde edilen madde parametreleri veya madde karakteristik eğrileri karşılaştırılır.

Madde karakteristik eğrilerinin aynı düzlem üzerindeki görünömlerinin incelenerek yanlılık hakkında kanıt aranır. Bu yöntemle madde yanlılıđı belirleme yöntemine Raju'nun alan ölçümü denilmektedir (Raju,1988,1990; Kim ve Cohen,1991). Aynı

düzlem üzerine konulan madde karakteristik eğrileri eğer çakışır veya madde karakteristik eğrileri arasındaki alanın ölçümü sıfır olursa, yanlılığın olmadığına karar verilir. Madde karakteristik eğrileri arasındaki alanın büyüklüğü sıfırdan uzaklaştıkça, maddenin yanlılığı da artmaktadır (aktaran: Deveci, 2008; Lord, 1980; Raju, 1988).

Madde karakteristik eğrilerinin durumuna bağlı olarak maddenin tek biçimli ve çok biçimli fonksiyonlaşması söz konusu olabilmektedir. Aşağıdaki şekilde maddenin tek ve çok biçimli farklı fonksiyonlaşmasının grafikte gösterimlerinde yetenek düzeyi ile grup kategorileri arasındaki ilişki modellenmiştir.



Şekil 2a ve 2b 'de y eksen, maddeye doğru cevap verilme olasılığını $P(\theta)$ gösterirken, x eksen yetenek düzeyini (θ) göstermektedir. Tüm yetenek düzeyleri bakımından bir grubun bir maddeyi doğru cevaplama olasılığının diğer gruptan yüksek olma durumu maddenin tek biçimli fonksiyonlaşmasıdır (Şekil 2a). Başka bir ifade ile; gruplar arası maddeyi doğru yanıtlama olasılığının tüm yetenek düzeylerinde sabit kalması biçiminde tanımlanabilir. Tek biçimli değişen madde fonksiyonunda (TBDMF) sadece madde güçlük parametresi farklılaşmaktadır. Maddenin çok biçimli

fonksiyonlaşmasında ise yetenek düzeyi ile grup kategorileri arasında bir kesişim söz konusudur. Gruplar arası başarı olasılığı tüm yetenek düzeylerinde sabit kalmamaktadır (Şekil 2b). Maddenin hem ayırt edicilik hem de güçlük parametreleri gruplar arasında farklılık göstermektedir.

Madde karakteristik eğrileri arasındaki alan hesaplamalarına bağlı yanlılık belirlemede farklı yöntemler kullanılmaktadır. İşaretli ve işaretsiz alan indeksleri, ağırlıklandırılmış ve ağırlıklandırılmamış işaretli ve işaretsiz alan indeksleri, yöntemleri bunlardan bazılarıdır (Raju ve Arenson 2002, Crocker ve Algina 1986). Bu yöntemleri yorumlarken kullanılan ölçütler, işaretli ve işaretsiz alan ölçümlerinin ortalama, medyan değerleri, bu değerlerin bir standart kayma veya çeyrek kayma üstü ile alan ölçümlerinin z değerlerinin, manidarlığıdır. İşaretli alan indeksleri iki alt grubun maddeye ait madde karakteristik eğrileri arasındaki farklılığın; işaretsiz alan indeksi ise iki maddeye ait madde karakteristik eğrileri arasındaki uzaklığın bir ölçüsüdür (Raju N., 1988).

Madde Parametreleri Tekniği: Parametre değerleri farklarından yararlanarak madde yanlılığı araştırması yapabilmek için, farklı iki grubun madde güçlük parametre değerleri karşılaştırılır. Testin uygulandığı gruptan odak ve referans grup olmak üzere iki alt grup belirlenir. Bu grupların, madde güçlük parametreleri aynı ölçek üzerinde eşitlenerek karşılaştırılır. Karşılaştırma işleminde odak grubun b parametre değeri referans grubun b parametre değerinden çıkarılır. Bulunan farkın manidarlığı test edilerek veya iç ölçüt kullanılarak maddelerin yanlılığa yol açıp açmadığına karar verilir.

Olabilirlik Oran Analizi: MTK'na bağlı, yanlılık belirlemede kullanılan diğer bir yöntem ise, sınırlı ve serbest model olarak tanımlanan iki modelin, olabilirlik oran farklarının anlamlılığının test edilmesidir (Thissen, Seinberg ve Wainer,1993). Yöntemin eşitliği aşağıda verilmiştir:

$$\text{Olabilirlik Oran (LR)} = L^*(\text{Sınırlı Model})/L^*(\text{Serbest Model})$$

Serbest model, bir veya daha fazla parametrenin değerlerinin kestirimi için serbest bırakılan modeldir. Örneğin, a ve b parametrelerinin gruptan gruba farklılık

göstermesine izin verilir. Hesaplanan G^2 istatistiği χ^2 (ki-kare) dağılımı gösterir. Bu değer ilgili serbestlik derecesi ile manidarlık düzeyindeki tablo değerinden büyükse iki gruptan kestirilen madde parametrelerinde farklılık olduğuna ve DMF içerdiğine karar verilir. Serbestlik derecesi olarak sınırlı ve serbest modelin parametre sayısının farkı alınır. G^2 değeri anlamlı değil ise serbest modeldeki parametrelerin hiçbirinin 0'dan farklı olmadığına karar verilir (Teresi, Kleinman, Welikson, 2000). Serbest modelde, modele ilave edilen, parametrelere ihtiyaç olup olmadığını ve ilave parametreleri modelin uyum derecesini artırmada katkı sağlayıp sağlamadığı tespit edilir. Burada H_0 hipotezi (sınırlı model) K sayıdaki parametre içerdiğini, H_1 hipotezi (serbest model), L sayıda parametre içerdiği şeklinde kurulur. Olabilirlik oran analizi ile K sayıdaki madde parametresi uyum iyiliğini sağlamakta mıdır? Başka bir ifadeyle, “L sayıdaki parametre modelin uyum iyiliğine katkı getirmekte midir?” sorusuna cevap aranır.

Olabilirlik oran analizi ile DMF incelenirken karşılaştırma yapılan grupların parametrelerini aynı metrik üzerinde eşitlemek üzere iç kriter olarak testin kendi maddelerinden oluşan ortak (ancor - connection) maddeler kullanılır (Thissen, Seinberg ve Wainer, 1993). Sınırlı modelde gruplar arasında madde parametrelerinin değişmediği varsayılarak, serbest modelde ortak maddeler dışında kalan maddeler için DMF incelenerek, iki modelin olabilirlik oran farklarının anlamlılığı testi edilir. Ortak maddeler DMF içermeyen maddeler olarak kabul edilir ve DMF analizine katılmaz (Kim ve Cohen, 1995).

1.2. ULUSLARARASI ÖĞRENCİ DEĞERLENDİRME PROGRAMI – PISA (PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT)

Eğitimde belirli periyotlarla ve içeriklerle uluslararası karşılaştırmalar yapılmaktadır. Bu karşılaştırmaların sonuçlarına bakılarak, ülkeler kendi eğitim düzeyindeki gelişmeleri izleyebilmekte ve varsa aksaklıkları görerek programlarını güncellemektedirler. Ekonomik İşbirliği ve Kalkınma Örgütü OECD'nin eğitim projelerinden biri olan PISA 15 yaş grubu öğrencilerin kazandıkları bilgi ve becerilerin değerlendirilmesine yönelik yapılan bir tarama araştırmasıdır. PISA projesi, zorunlu eğitimin sonunda örgün eğitime devam eden 15 yaş grubu öğrencilerin öğretim programlarında ele alınan konuları (matematik, fen bilimleri ve

okuma becerileri) ne dereceye kadar öğrendikleri değil, günümüz bilgi toplumunda karşılaşabilecekleri durumlar karşısında sahip oldukları bilgi ve becerileri kullanabilme yeteneğini ölçmeyi amaçlamaktadır. PISA projesinde Matematik okuryazarlığı, Fen Bilimleri okuryazarlığı, Okuma Becerileri konu alanları ve öğrencilerin motivasyonları, kendileri hakkındaki görüşleri, öğrenme biçimleri, okul ortamları ve aileleri ile ilgili veriler toplanmaktadır. Üçer yıllık dönemler halinde uygulanan PISA projesinde, her bir dönemde bir konu alanına ağırlık verilmektedir. PISA projesinin ilki 2000 yılında uygulanmış, bu uygulamada okuma becerilerine ağırlık verilmiştir. 2003 yılında yapılan ikinci uygulamada Matematik okuryazarlığı alanına, 2006 yılında yapılan uygulamada ise Fen Bilimleri alanına yoğunlaşmıştır.

PISA 2006 uygulamasına aralarında Türkiye’inde yer aldığı 30’u OECD ülkesi olmak üzere toplam 57 ülke katılmıştır. PISA projesinde çoktan seçmeli, karmaşık çoktan seçmeli, açık uçlu, kapalı uçlu gibi değişik soru türleri kullanılmaktadır. TIMMS, PISA gibi çalışmalar uluslararası alanda başka kültürleri araştıran ve uluslararası araştırmalar yapan sosyal bilimciler için fırsatlar sağlamaktadır.

Bu çalışma kapsamında, ülkemizde 2006 yılında katıldığı PISA projesine ait fen testinin psikometrik özellikleri kültürler arası karşılaştırmalı olarak incelenmiştir. Çalışmada Türkiye, Avusturalya, İngiltere, Kanada ve İspanya başarı testine ilişkin veriler kullanılmıştır.

1.3. PROBLEM CÜMLESİ

PISA 2006 yılında yapılan Uluslararası Öğrenci Değerlendirme Programında (PISA) yer alan fen başarı testi maddeleri dile ve kültüre göre farklı fonksiyonlaşmakta mıdır? Farklı fonksiyonlaşan maddelerin olası yanlışlık sebepleri neler olabilir?

1.4. ALT PROBLEMLER

1. PISA 2006 yılında yapılan Uluslararası Öğrenci Değerlendirme Programında (PISA) yer alan fen başarı testi maddeleri; Mantel-Haenszel (MH), Logistik Regresyon(LR) ve Alan İndeksleri Teknikleri (AİT) ile yapılan analizlerde kültür ve dile göre farklı fonksiyonlaşmakta mıdır?

1.a.Aynı kültür - Farklı dil (Kanada örneği) için fen başarı testi maddeleri Mantel-Haenszel (MH), Logistik Regresyon(LR) ve Alan İndeksleri Teknikleri (AİT) ile yapılan analizlerde kültür ve dile göre farklı fonksiyonlaşmakta mıdır?

1.b.Farklı Kültür - Aynı Dil (Avustralya- İngiltere Örneği) için fen başarı testi maddeleri Mantel-Haenszel (MH), Logistik Regresyon(LR) ve alan indeksleri teknikleri ile yapılan analizlerde kültür ve dile göre farklı fonksiyonlaşmakta mıdır?

1.c.Farklı Kültür - Farklı Dil (İngiltere – Türkiye Örneği) birinci kitapçık ve (İngiltere – Türkiye Örneği) beşinci kitapçık için fen başarı testi maddeleri Mantel-Haenszel (MH), Logistik Regresyon(LR) ve Alan İndeksleri Teknikleri (AİT) ile yapılan analizlerde kültür ve dile göre farklı fonksiyonlaşmakta mıdır?

2. Kültür ve dile göre Mantel-Haenszel (MH), Logistik Regresyon(LR) ve Alan İndeksleri Teknikleri (AİT) ile elde edilen DMF düzeyleri farklılaşmakta mıdır?

3. Uzman kanılarına göre maddelerin kültürler ve diller arasında(Farklı Kültür- Farklı Dil gruplaması için) DMF göstermesinin olası nedenleri nelerdir?

1.5. ARAŞTIRMANIN ÖNEMİ VE AMACI

Küreselleşen dünyamızda, eğitim alanında yapılan ulusal değerlendirme çalışmalarının yanı sıra, uluslararası düzeyde konumumuzu belirlemek amacıyla eğitim göstergelerine ihtiyaç duyulmaktadır. Uluslar ekonomik, sağlık ya da bilimsel alanlarda gittikçe daha fazla işbirliği içine girmektedirler. PISA, TIMSS, PIRLS eğitim alanında bu girişimi destekleyen tarama çalışmalarıdır. PISA, TIMSS, PIRLS gibi kültürler-arası çalışmalar, ülkelerdeki eğitim sistemini değerlendirmek ve geliştirmek için geniş bakış açıları sağlamaktadır. Belirli referans noktalarına göre yürütülen çalışmalar ile ülkelerin eğitim alanında hangi düzeyde olduğunun, giderilmesi gereken eksikliklerin ve alınması gereken tedbirlerin belirlenmesi, bu sayede de eğitim düzeyinin yükseltilmesi amaçlanmıştır. Büyük çaplı sınav verileri kullanılarak yapılan yanlılık çalışmalarının, ülkeler arası karşılaştırmaların anlamlılığı bakımından önemli bir ihtiyaç olduğu düşünülmektedir.

Uygulamaya katılan tüm ülkeler için büyük önem taşıyan PISA uygulamasının sonuçları, ülkelerarası karşılaştırmalar için çeşitli araştırmalarda farklı şekillerde yer almıştır. Bu araştırmaların tamamına yakını PISA'dan elde edilen aile ya da öğrenci değişkenlerine ait verilerin öğrencilerin başarılarıyla ilişkilerinin incelenmesini konu almaktadır. Ancak PISA'nın bilişsel özellikleri ölçen maddelerine odaklanan çalışmalar sınırlı sayıdadır. Değişen madde fonksiyonu ve yanlılık çalışmalarında PISA gibi uluslararası bir sınavda dil ve kültürden kaynaklanabilecek olası yanlılık nedenlerine odaklanmak önemlidir. Bilindiği üzere yanlılık çalışmaları ölçme aracının geçerliği bakımından kritik çalışmalardır ve eğitim politikalarında önemli değişikliklere yol açan uluslararası karşılaştırılmalı çalışmaların geçerlilikleri uluslararası düzeyde büyük önem taşımaktadır. Geçerliliğin düşük olduğu durumlarda yapılan karşılaştırmaların vereceği bilginin hiç bir bilimsel değeri ve anlamı olmayacaktır. Bu bakımdan PISA bilişsel testinin geçerliğine ilişkin bulguların farklı ülkeler bakımından değerlendirilmesinin önemli olduğu düşünülmektedir.

Bu çalışmanın amacı, PISA 2006 öğrenci anketinde yer alan maddelerin aynı kültür farklı dil, farklı kültür aynı dil ve farklı kültür farklı dil çerçevesinde değişen madde fonksiyonunun farklı teknikler ile analiz edilmesi ve karşılaştırılmasıdır. Böylece çalışmanın; DMF belirlemek için klasik test kuramı ve madde tepki kuramı dahilinde kullanılabilecek tekniklerden elde edilen sonuçların benzerlik ve farklılıklarının ortaya çıkarılması ve teknikler arasındaki uyumun belirlenmesi amaçlanmıştır. Bu çalışmanın DMF belirleme tekniklerinin karşılaştırılmasına yönelik çalışmalarda ortaya çıkan çelişkili sonuçların aydınlanmasına katkı sunabileceği düşünülmektedir.

Diğer bir amaç ise değişen madde fonksiyonu tespit edilen maddelerin olası DMF sebeplerinin araştırılmasıdır. Yapılan DMF analizleri sonucunda çok az çalışma DMF'nin olası sebeplerine değinmektedir. DMF'nin sebepleri belirlenebilirse, farklı test formlarının değişik alt gruplar açısından karşılaştırılabilirliğini sağlayabilmek için test geliştirme ve uyarlama süreçlerinde daha dikkatli olunması gereken aşamalar belirlenebilir.

1.6. SINIRLILIKLAR

Bu çalışma, OECD'nin eğitim projesi olan ve 2006 yılında yapılan PISA verilerinin bilişsel alanda yapılan ve fen okur yazarlığına yönelik birinci ve beşinci kitapçıkta yer alan çoktan seçmeli maddelere ait veriler ve çalışmanın yürütüldüğü ülkeler ile sınırlıdır.

Alan indeksleri yönteminde kullanılan Calcbias programının 2 PLM parametrelerini kullanarak işaretli ve işaretsiz alan indeksleri değerlerini türetmesi teknik bir sınırlılıktır.

Çalışmanın üçüncü alt probleminin (DMF'nin olası nedenleri) sadece açıklanan maddeler üzerinden yürütülmesi de diğer bir sınırlılıktır.

1.7.İLGİLİ ÇALIŞMALAR

Engelhard (1990) bilişsel karmaşıklık ve içerik kategorisi ile ilgili test maddeleri seti üzerinde cinsiyet ve performans arasındaki ilişkiyi araştırmıştır. Bu amaç için TIMSS'e katılan Tayland ve Amerika Birleşik Devletleri örneklemini çalışma grubu olarak kullanmıştır. DMF'yi belirlemek için Mantel-Haenszel tekniğini kullanmıştır. Bilişsel düzey (cebir, aritmetik, geometri) ve içerikle (hesaplama, anlama, analiz) ilgili olduğu belirlenen maddelerde cinsiyet bakımından farklılıkları test etmek için tekrarlı ölçümlerde anova testi (repeated measures of ANOVA) yapılmıştır. Sonuçlar bilişsel karmaşıklık ve içerik kategorisi düzeylerinde cinsiyet farklılıkları olduğunu göstermiştir. Cinsiyet farklılıkları; bilişsel karmaşıklık düzeyi yükseldikçe erkeklerin kızlara göre daha başarılı olduğu yönündedir.

1993 yılında Harris and Carlton SAT'ın matematik başarı testi maddelerinde cinsiyete göre farklılaşmaları test etmiştir. DMF'yi belirlemek için Mantel-Haenszel tekniğini ve tek yönlü varyans analizini kullanmışlardır. Maddeler madde biçimleri ve madde içerik kategorilerinde analiz edilmiştir. Çalışmanın sonuçları aynı puanı alan kız ve erkek öğrencilerin; bu sonuçları farklı soru çiftlerini yanıtlayarak aldıkları görülmüştür. Erkek öğrencilerin nispeten geometride kız öğrencilere göre daha iyi

performans sergiledikleri, kız öğrencilerinde cebirle ilgili maddelerde erkeklerden daha iyi performans gösterdiği sonucuna ulaşılmıştır.

Öğretmen (1995) 1993 yılı ÖSS birinci basamak Türkçe ve sosyal bilimler alt testlerindeki maddelerin konu alanlarını ve taksonomik seviyelerini cinsiyet gruplarına göre karşılaştırdığı madde yanlılığı analizini dört farklı alan indeksi kullanarak yapmıştır. Türkçe alt testinde kullanılan yöntemle ilgili olarak farklı sonuçlar elde edilmiştir. Sosyal bilimler alt testinde ise tarih ve coğrafya maddeleri erkek öğrenciler lehine çalışırken, psikoloji ve sosyoloji maddelerinin kız öğrenciler lehine çalıştığı tespit edilmiştir.

Yenal (1995), 1993 Öğrenci Seçme Sınavı (ÖSS) sayısal testinin cinsiyete göre yanlılığını incelemiştir. Bu çalışmanın verileri 1993 ÖSS'ye giren Ankara'daki beş devlet ve özel okulun öğrencilerinin cevaplarından elde edilmiştir. Araştırmada Madde Tepki Kuramı modellerine göre elde edilen dört alan indeksi madde yanlılığı açısından değerlendirilmiştir. Bulgular, işaretli ve işaretsiz alan indekslerinin madde karakteristik eğrilerinde düzgün olmayan bir yanlılık olduğunu göstermektedir. Konu alanlarının karşılaştırılması sonucunda geometri maddelerinin matematik testinin diğer konu alanlarına göre daha fazla yanlılık gösterdiği ve bu yanlılığın erkeklerin aleyhine olduğu gözlenmiştir. Fen bilgisi testinde ise yanlılık en fazla biyoloji maddelerinde ve erkeklerin aleyhine olmuştur. Ayrıca, analiz sonuçları fen bilgisi testinin matematik testinden daha fazla yanlılık içerdiğini göstermiştir. Diğer taraftan, alan indeksleri arasında anlamlı bir ilişki olduğu gözlenmiştir.

Berberoğlu (1996) öğrenci Seçme Sınavları ile ilgili çalışmalarda, sınavda yer alan Türkçe, Sosyal Bilimler, Fen Bilimleri ve Matematik alt testlerinde olası madde yanlılıkları incelemiştir. 1995 yılı Üniversite giriş sınavının Türkçe alt testindeki maddelerin %58'inde, Sosyal Bilimler alt testinde yer alan maddelerin %68'in de Fen Bilimleri ve Matematik testindeki maddelerin %53 'ün de erkekler lehine yanlılık gözlemlendiğini belirtmiştir. Fen bilimleri testindeki biyoloji maddelerinin kızların, fizik maddelerinin erkeklerin lehine yanlılık gösterdiği ayrıca, bu sınavda yer alan testlerde hesaplamaya yönelik maddelerde erkek öğrencilerin, sözel ve uzamsal ilişkilere yönelik maddelerde ise, kız öğrencilerin daha avantajlı olduklarını bulunmuştur.

Beaton (1998) yaptığı çalışmada “ülkelerin farklı müfredatlara sahip olması ve farklı öğretim yöntemleri” çıkış noktasından hareketle TIMSS gibi uluslararası çalışmalarla yapılan karşılaştırmaların sınırlı olabileceği üzerinde durmuştur. Bu teoriyi güçlendirmek için müfredat eşleştirme analizi ile farklı matematik ve fen maddelerinden oluşan farklı alt testler kullanmıştır. Genel uluslararası sıralama görüntüsü değişmese de, alt alanlara ilişkin ülke performanslarının birbiriyle ilişkisinin yüksek olduğunu saptamıştır. Öte yandan bir alt alanda yüksek puan alan öğrenciler, diğer alt testlerde de yüksek puanlar almışlardır.

Girl ve Khaliq (2001) 1995 TIMSS matematik ve fen maddeleri üzerinde yürüttüğü nicel ve nitel analizler ile olası yanlışlık nedenlerini araştırmıştır. SIBTEST kullanılarak yürütülen istatistiksel analizlerin yanında 11 test geliştiricinin DMF belirlenen maddelere ilişkin görüşleri alınarak yanlışlık nedenleri ortaya çıkarılmaya çalışılmıştır. Çalışmada temel yanlışlık nedeni olarak çeviriden kaynaklanan farklılıklar vurgulanmıştır. Çevirilerde yer alan kelime ve ifadede ki anlamsal farklılıklar, maddelerdeki cümle yapılarının farklılaşmasının SIBTEST ile DMF gösterdiği belirlenen maddelerde yanlışlık nedeni olarak ön plana çıkmıştır.

Klieme ve Baumert (2001) TIMSS verilerini kullanarak lise öğrencilerinin yeterli profillerinin belirlenmesi için yürüttükleri çalışmada ülkelerin güçlü ve zayıf yönlerini belirlemek için DMF yöntemlerinden biri olan MTK yaklaşımlarını kullanmışlardır. Referans grup olarak Almanya seçilmiştir. MTK yaklaşımı ile Almanya’ya göre diğer ülkeler incelendiğinde belirgin farklılıklar bulunduğu tespit edilmiştir. Örneğin; ABD’nin açıklayıcı ve yöntemsel bilgide daha başarılı olduğu; Almanya’nın gelişmiş bilgi ve anlayışında zayıf kaldığı ama görsel ve grafik sunumlarının kullanımında ABD’ye göre daha başarılı olduğu görülmüştür.

Ercikan, Gierl, McCreith, Puhan ve Koh (2002); Koh ve Ercikan (2002) okul başarı göstergeleri programına (School Achievement Indicators Program=SAIP) yönelik yaptıkları çalışmada testin matematik, fen ve dil bilimleri ile ilgili değerlendirmelerinde İngilizce ve Fransızca formu alan öğrenciler bakımından DMF içerdiğini belirlemiştir. DMF nedenleri araştırıldığında farklılığın müfredattan, anlam

bakımından kullanılan kelimelerin farklı zorluklarda olmasından kaynaklandığı ortaya çıkarılmıştır.

Zenisky, Hambleton ve Robin (2003) fen maddeleri üzerinde cinsiyete ilişkin değişen madde fonksiyonlarını belirlemeyi amaçlayan bir çalışma planlamışlardır. Bu çalışmada DMF'ye neden olabilecek faktörleri araştırmışlardır (içerik, bilişsel talepler, madde türü, madde metni ve görsel-mekansal veya referans faktörler). Çalışma İlkokul, orta ve lise düzeyi yaklaşık 360.000 öğrenci ile yürütülmüştür. Çoktan seçmeli ve açık uçlu sorulardan oluşan 32-42 maddelik testler kullanılmıştır. İçerik kategorisi, görsel- mekansal ve referans bileşenleri ve madde türü ile ilgili olası madde desenleri her bir düzey için araştırılmıştır. İçerik kategorisi, görsel-mekansal bileşeni ve madde türü boyutlarında farklılıklar bulunmuştur.

Doğan ve Öğretmen (2004) OKÖSYS matematik alt testine ait maddelerin yanlılık analizi çalışmasını yapmışlardır. Ankara'daki ilköğretim okullarının 1684 kız ve 1660 erkek olmak üzere toplam 3344 ilköğretim sekizinci sınıf öğrencisinin OKÖSY Sınavındaki maddelere verdikleri cevaplar kullanılarak yanlılık çalışılmıştır. Yapılan analizler sonucunda; matematik alt testindeki maddelerin cinsiyete göre yanlılık gösterdiğini ortaya koymaktadır. Sonuçlara göre, testteki çoğu madde kızların lehine yanlılık göstermektedir.

Fen bilimleri ve sosyal bilimler gibi farklı akademik özgeçmişleri olan Kore öğrencilerinin 1998 üniversite ve yüksekokul giriş sınavının, İngilizce alt testi maddelerinin yanlılık taşıyıp taşımadığını araştıran Pae (2004) Mantel-Haenszel ve likelihood ratio yöntemini kullanmıştır. Testte okuma ve dinleme kapsamlarında elli beş madde vardır. Testte 28 maddede DMF belirlenmiştir. Yanlılık kaynaklarının belirlenmesi için içerik araştırması yapılması önerilmiştir.

Yurdugül ve Aşkar (2004) Ortaöğretim Kurumları Sınavı üzerinde MH yöntemi ile yaptıkları çalışmada, Türkçe, matematik, fen bilimleri ve sosyal bilimler alt testlerindeki maddelerin yanlılık analizlerini yapmışlardır. Türkçe ve sosyal bilimler alt testlerindeki maddelerin farklı fonksiyonlaşmadığını veya fonksiyonlaşmanın ihmal edilebilir düzeyde olduğunu, matematik alt testinde ise simgesel işlemler, sembollere dayalı gösterimler ve basit denklem maddelerinin kız öğrenciler lehine çalışırken,

sözel problem maddelerinin erkek öğrenciler lehine çalıştığını ifade etmişlerdir. Ayrıca, fen bilimleri alt testindeki fizik maddelerinin erkek öğrenciler, biyoloji maddelerinin ise kız öğrenciler lehine çalıştığını belirtmişlerdir.

Ercikan ve Koh (2005), TIMSS 1995 İngilizce ve Fransızca testlerinin eşdeğerliğini Kanada, İngiltere, Fransa ve ABD örneklemi üzerinde değerlendirmişlerdir. Araştırmacılar çalışmalarında çoklu grupta doğrulayıcı faktör analizi ve madde tepki kuramı parametre karşılaştırma yöntemlerini kullanmışlardır. Analizler sonucunda matematik ve fen bilimleri testlerinin İngilizce ve Fransızca versiyonlarının ölçtüğü yapılar arasında önemli farklılıklar olduğu tespit edilmiştir. Karşılaştırmalarda matematik maddelerinin %14 ile %39, fen bilimleri maddelerinin ise %37 ile %79 arasında DMF gösterdiği belirlenmiştir.

Çet (2006) PISA 2003 Matematik sorularının Türkçe ve İngilizce formları arasındaki madde yanlılığını araştırmıştır. 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. Tek boyutlu DIF analizleri ile çok boyutlu eşleştirme analizlerinin sonuçları Mantel-Haenszel (MH) ve Logistik Regresyon (LR) metotları kullanılarak karşılaştırılmıştır. Bu karşılaştırma sonucunda çok boyutlu eşleştirme yöntemleri ile yapılan analizlerde madde yanlılığı gösteren maddelerde her iki kitapçıkta da bir farklılık görülmüştür. Yanlı çalıştığı tespit edilen maddelerde 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 Amerika'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.

İngilizce'yi yabancı dil olarak öğrenenlere uygulanan dört pasaj, yirmi okuma sorusundan oluşan bir dil testinin cinsiyet yanlılığı içerip içermediğine bakan Ross ve Okabe, (2006) araştırmalarını beş Japon enstitüsünden, 468 kız öğrenciyi odak, 357

erkek öğrenciyi referans grup olarak belirleyerek yürütmüşlerdir. Madde yanlılığına, objektif ve subjektif yöntemler olarak bakılmıştır. Objektif yöntemler olarak, Mantel-Haenszel (MH) delta-plot, Lojistik regresyon ve Simultane yöntemleri kullanılmıştır. Subjektif yöntem ise uzman görüşleridir. Uzman görüşü olarak, deneyimli İngilizce dil öğretmenlerinden testte yer alan maddelerin yanlı olup olmadığına dair görüşleri alınmıştır. Bulgular, testte konu ve cinsiyet yanlılığı olduğunu, subjektif yöntemle madde yanlılığı belirleme yönteminin, daha fazla maddeyi yanlı gösterdiği yönündedir.

Doğan ve Öğretmen 2006 yılında yaptıkları çalışmada madde yanlılıklarını belirlemede kullanılan tekniklerden elde edilen madde yanlılık değerleri arasındaki ilişkileri belirlemeyi amaçlanmıştır. Çalışmalarında “madde güçlük dönüşümü, düzeltilmiş madde güçlük dönüşümü, Mantel – Haenszel, işaretli ve işaretli alan indeksleri teknikleriyle elde edilen madde yanlılık değerleri arasında nasıl bir ilişki vardır?” sorusuna yanıt aramışlardır. Teknikleri karşılaştırmak için 2003 yılı Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavının Fen Bilgisi alt testine cevap veren 550000 kişi arasından rastgele seçilmiş, evreni temsil edebildiği belirlenen 3344 kişilik örneklem kullanılmıştır. Yanlılık analizleri cinsiyet bağlamında yürütülmüştür. Beş teknikte de yanlılık gösteren 4 madde bulunmuştur. Bu maddelerin dördü de erkekler lehine işlemiştir. Beş farklı teknikte elde edilen madde yanlılık değerlerinin büyüklük sıraları arasında manidar ilişkilere rastlanmıştır. Söz konusu ilişkiler, madde güçlük dönüşümü ve düzeltilmiş madde güçlük dönüşümü değerleri; madde güçlük dönüşümü ve düzeltilmiş madde güçlük dönüşümü değerleri ile Mantel – Haenszel değerleri; işaretli ve işaretli alan indeksleri arasında 0,01 düzeyinde anlamlıdır. İşaretli alan indeksleri ile Mantel – Haenszel değerleri arasındaki ilişki ise 0,05 düzeyinde anlamlı bulunmuştur.

Pae ve Park (2006) 2003 yılı Kore üniversite giriş testindeki İngilizce alt testinin (KCSAT) test ve madde yanlılığını araştırmıştır. Madde yanlılığını belirlemede olasılık oran ölçümü (Likelihood-Ratio) ve doğrulayıcı faktör analizi (Confirmatory factor analysis) yöntemleri kullanmışlardır. Araştırma 15000 öğrenci ile yürütülmüştür. Araştırmaya katılan öğrencilerin 3500 kız-sosyal bilimler mezunu, 3500 erkek ve örneklem yarısı kız yarısı erkek, kızların ve erkeklerin yarısı ise fen bilimleri ve sosyal bilimlerden seçkisiz olarak alınmıştır. Toplam 33 okuduğunu anlama

sorusundan 22'sinde madde ayrışık fonksiyonu (DMF) bulunmuştur. 13 madde erkekler, 9 madde kızlar için kolay maddeler olarak çalışmıştır. Erkeklerle, maddeler iş danışmanlığı, gazetelik, basit zihinsel maddelerle ilgili maddeler, kızlara ise seyahat, macera, sanat, TV muhabirliği gibi konularla ilgili maddeler kolay gelmiştir.

Öğretmen (2006) uluslararası okuma becerilerinde gelişim projesi (PIRLS) 2001 testi kapsamında uygulanan okuma parçalarına ait testler üzerinde MTK parametre karşılaştırma ve olabilirlik oranı testi model karşılaştırma yöntemleriyle DMF analizleri yapmıştır. Her iki yönteme göre de PIRLS testlerine ait maddelerin çoğunun kültürel karşılaştırmada farklı fonksiyonlaştığı ortaya çıkmıştır.

Yıldırım (2006) TIMSS-1999 ve PISA 2003 uluslararası sınavlarının matematik başarı testlerinin kültürler arası denkliliğini ve matematik başarı kavramının altında kültürlere has özgül yapılar olup olmadığını araştırmıştır. Bu amaçla, TIMSS-1999 ve PISA 2003 Türkçe ve İngilizce versiyonlarındaki matematik başarı testi maddeleri, farklı dildeki test maddelerinin psikometrik özellikleri ve 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. Sonuçlar, hem TIMSS hem de PISA çalışmalarında kullanılan bazı maddelerde kültürel denklik açısından problemler olduğunu ortaya koymuştur. Çalışmada PISA çalışmasında, madde yanlılığı tespitinde kullanılan farklı metotların sonuçları arasında yüksek derecede uyum tespit edilmiştir. Ancak; TIMSS çalışmasında, çok sayıda yanlı çalışan madde tespit edildiğinden; maddelerin tahmin ve ayırt edicilik indislerindeki farklılıklar bu uyumu kayda değer ölçüde bozmuştur.

Kurnaz (2006), Peabody resim kelime testinde yer alan maddelerin sosyo ekonomik düzey ve cinsiyete göre DMF içerip içermediğini MH ve LR teknikleri ile belirlemeye çalışmıştır. Çalışmasının örneklemini Peabody resim kelime testini alan 592 çocuk oluşturmuştur. Çalışmadan elde edilen bulgulara göre; cinsiyete göre MH tekniği ile 16, LR tekniği ile 14; sosyo ekon omik düzey değişkenine göre ise MH tekniği ile 33, LR tekniği ile 20 maddede DMF olduğu sonucuna ulaşılmıştır. Araştırmacı

çalışmasından elde ettiği bulgulardan yola çıkarak MH tekniğinin çeldirici kaynaklı yanlışlığa da duyarlı olduğu sonucuna ulaşmıştır.

Kan (2007) yaptığı araştırmada, Hacettepe Üniversitesi yabancı diller yüksek okulunun uyguladığı 85 maddeden oluşan yabancı dil muafiyet sınavının cinsiyet ve bölümlere göre yanlışlık içerip içermediğini araştırmıştır. Araştırma, 2071 öğrenci üzerinde yürütülmüştür. Değişen madde fonksiyonunun araştırılması için Mantel-Haenszel istatistiğinden yararlanılmıştır. Çalışma sonucunda, Hacettepe üniversitesi yabancı diller yüksek okulunun uyguladığı yabancı dil muafiyet sınavında kullanılan bazı maddelerin cinsiyete ve bölümlere göre anlamlı büyüklükte DMF içerdiği saptanmıştır.

Deveci (2008) 2005 Mart dönemi ÜDS-Sosyal Bilimler-İngilizce testi maddelerinin, Eğitim Bilimleri ve Öğretmen Yetiştirme (EBÖY) alanı ile Beşeri ve İdari Bilimler (BİB) alanlarından gelen katılımcıların eğitim alanları ve cinsiyetine göre, yanlışlık gösterip göstermediği çalışmıştır. Araştırma, testi yanıtlayan, 1933, EBÖY alanı ve 1798, BİB alanı katılımcıları ile yürütülmüştür. Verilerin çözümlenmesinde, Madde Tepki Kuramı'na dayalı, Değişen Madde Fonksiyonu(DMF) belirleme yöntemleri olan, b parametre farkları, işaretli ve işaretsiz alan indeksleri yöntemleri kullanılmıştır. Ayrıca bu yöntemlere göre DMF içermeyen maddeler, ortak madde olarak kabul edilip, olabilirlik oran analizi ile DMF belirlenmiştir. Yanlışlık incelemeleri, alan, cinsiyet ve alanların kendi içindeki cinsiyet gruplarında yürütülmüştür. Araştırmada elde edilen bulgulara göre; testteki on madde kapsamından, “uygun kelime, fiil veya deyim bulma, cloze test (kelime, cümle tamamlama), paragrafta uygun olmayan cümleyi bulma ve metin inceleme ve anlatım” madde kapsamı DMF'na yol açıyor görünmektedir. Bu madde kapsamlarında yer alan, siyaset, ekonomi, bankacılık ve bilgisayar konu içerikli maddeler, BİB alanına, tarih, askeri operasyon ve arkeoloji konu içerikli maddeler ise, EBÖY alana avantaj sağlamıştır. Aynı şekilde, doğum oranları, ekonomi, para ve sanat olan maddeler, kız yanıtlayıcı grubuna, askerlik, döviz kuru, tarih ve ticaret konu içerikli maddeler ise erkek yanıtlayıcı grubuna avantaj sağlamıştır. Genel olarak, ÜDS-SB İngilizce testi yanıtlayan, BİB alanı yanıtlayıcıları, EBÖY alanı yanıtlayıcılarına göre ve tüm grupta, kızlar, erkeklere göre daha avantaj sağlamıştır.

Kalaycıoğlu (2008) 2005 yılı Öğrenci Seçme Sınavında (ÖSS) yer alan maddelerin cinsiyet ve okul türü değişkenlerine göre farklı fonksiyonlaşıp fonksiyonlaşmadığı incelemiştir. Farklı fonksiyonlaştığı belirlenen maddelerin yanlış olup olmadığına karar vermek için uzman görüşüne başvurulmuştur. Araştırmada, Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) yöntemleri kullanılmıştır. Analizler ÖSS'ye giren 599,330 lise son sınıf öğrencisinin her biri 45 maddeden oluşan Türkçe, sosyal bilimler, matematik ve fen bilimleri alt testlerine verdikleri cevaplar kullanılmıştır. Araştırmanın bulgularına göre indeks değerleri göz önüne alındığında, MH ve LR yöntemleri birbiriyle uyumlu sonuçlar vermiştir; fakat literatürdeki kriter puanlar göz önüne alındığında, bu değerlerin LR yöntemi için yüksek olduğu saptanmıştır. Bu sebeple, çalışmada, yöntemler birbirini tamamlayacak şekilde farklı bir kriter belirlenmiş ve bu kritere uyan maddelerin farklı fonksiyonlaştığı kabul edilmiştir. Araştırmada elde edilen bulgular doğrultusunda, cinsiyete göre alt testler göz önüne alındığında, Türkçe alt testini oluşturan maddelerin farklı fonksiyonlaşmadığı sonucuna varılmıştır. Sosyal bilimler alt testinde 7, matematik ve fen bilimleri alt testlerinde de 3 maddenin farklı fonksiyonlaştığı ve bu maddelerden de 1 fen bilimleri maddesinin yanlışlık gösterdiği belirlenmiştir. Okul türlerine göre karşılaştırma yapıldığında, matematik alt testini oluşturan maddelerin farklı fonksiyonlaşmadığı sonucuna varılmıştır. Türkçe alt testinde; genel lise Anadolu lisesi karşılaştırmasında 14, genel lise özel lise karşılaştırmasında 9 ve Anadolu lisesi özel lise karşılaştırmasında 9 maddenin; sosyal bilimler alt testinde, genel lise Anadolu lisesi karşılaştırmasında 16, genel lise özel lise karşılaştırmasında 1 ve Anadolu lisesi özel lise karşılaştırmasında 6 maddenin; fen bilimleri alt testinde genel lise Anadolu lisesi karşılaştırmasında ise 1 maddenin farklı fonksiyonlaştığı ve bu maddelerden de sadece 1 fen bilimleri maddesinin yanlışlık gösterdiği belirlenmiştir.

Acar (2008) 2006 yılında uygulanmış olan Ortaöğretim Kurumları Sınavının (OKS) alt test maddelerinde cinsiyete ve bölge alt gruplarına göre genelleştirilmiş aşamalı doğrusal modelleme (GADM), lojistik regresyon (LR) ve madde tepki kuramı-olabilirlik oranı (MTK-OO) teknikleri ile maddenin farklı fonksiyonlaşıp fonksiyonlaşmadığını belirlemeye çalışmıştır. Araştırmanın örneklemini, 2006-OKS sınavına giren 10727 öğrenci oluşturmuştur. Araştırmanın bulgularına göre GADM, LR ve MTK-OO teknikleri ile bakıldığında; tüm alt testlerde en fazla cinsiyete göre

DMF gözlenmiştir. Tüm alt testlerde cinsiyete ve bölge alt gruplarına göre LR tekniği ile belirlenen DMF'lerin göz ardı edilebilir düzeyde olduğu; MTK-OO tekniğinde belirlenen DMF'li maddelerin bazılarında da orta ve önemli düzeylerde DMF'ler olduğu bulunmuştur. Tekniklerin DMF'li maddeyi belirleme bakımından aralarındaki ilişkiye bakıldığında her üç alt testte cinsiyete göre MTK-OO ile LR arasında anlamlı bir ilişki olduğu görülmüştür. Bölge alt gruplarına göre GADM ile LR, MTK-OO ile LR ve GADM ile MTK-OO arasında anlamlı ilişkilerin farklı alt testlerde olduğu görülmüştür.

Doğan ve Öğretmen (2008) değişen madde fonksiyonunu (DMF) belirleme tekniklerinden kıkare, Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırarak uygulamada ortaya çıkan benzerlik veya farklılıklarını belirlemeyi amaçlayarak yürüttükleri çalışmada 2003 yılında Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı'na (OKÖSYS) katılan yaklaşık 600.000 öğrenci arasından yansız olarak seçilen 3345 öğrencinin fen bilgisi alt testine verdiği yanıtları kullanmışlardır. Değişen madde fonksiyonu analizi sadece cinsiyet grupları üzerinde yürütülmüştür. Araştırmanın sonuçlarına göre çalışılan DMF belirleme tekniklerinden MH, Ki-kare ve lojistik regresyon tekniklerinin Ki-kare değerlerinin büyüklüğü bakımından benzer; DMF verdiği belirlenen madde sayısı bakımından farklı sonuçlar ürettiği bulunmuştur.

Qi ve Marley Okul Öncesi Dil Ölçeği-4 (PLS-4)'te madde yanlılığının olup olmadığı incelenmiştir (2009). Katılımcılar ana programa kayıtlı olan 440 öğrenciden (3-5 yaşlarında: %86 İspanyol %14 Avrupalı) oluşmaktadır. PLS-4 maddeleri tek parametrelili madde tepki kuramı (MTK) modeli ve lojistik regresyon kullanarak analiz edilmiştir. Çalışmanın sonuçlarına göre İngilizce konuşan İspanyol çocuklarına karşı büyük ölçekli sistematik madde yanlılığını bulunmamıştır. Ölçekteki 4 madde de DIF bulunduğu ilişkin sonuç çıkarılmıştır. İki madde İspanyol çocuklarının aleyhine olduğu; bir maddenin Avrupalı çocuklar için lehine işlediği ve sonuncu maddenin düzgün olmayan DIF gösterdiği görülmüştür. Madde analizleri açısından bu çalışmanın sonucu İngilizce konuşan İspanyol çocukları için kültürel yanlılığın güçlü bir kanıtı bulunmamıştır.

Adedoyin (2010) JSS sertifika sınavında yeralan 38 matematik maddesinin cinsiyete göre DMF içerip içermediğini araştırmıştır. Çalışmayı sınavı alan 36000 öğrenciden tesadüfi örnekleme yolu ile seçilen 2000'i erkek, 2000'i de kız olmak üzere 4000 öğrenci üzerinde yürütmüştür. Analizlerde veri 3 parametrelili model uyum göstermiştir. Her iki gruptan elde edilen madde karakteristik eğrileri karşılaştırılmış ve 5 maddenin cinsiyete göre DMF içerdiği sonucuna ulaşılmıştır.

Abedalaziz (2011) 3 parametrelili lojistik modelle elde edilen alan indeksleri ve 2 parametrelili modelle elde edilen alan indekslerinin matematik yeteneğini ortaya koymaya çalışan bir ölçme aracında cinsiyet alt gruplarında DMF belirlemede benzer sonuçlar verip vermediğini araştırmıştır. Bu amaçla araştırmacı 10. Sınıflar için kullanılan bağıntı ve fonksiyonlar, trigonometrik fonksiyonlar, polinomlar konularına ait sorulardan oluşan bir başarı testi kullanmıştır. Test 664'ü erkek; 654'ü kız olmak üzere toplam 1318 10. Sınıf öğrencisine uygulanmıştır. Sonuçlar 2 ve 3 parametrelili modelle elde edilen alan indekslerinin göreceli olarak benzer olduğu ve Matematikteki konuların cinsiyete göre DMF gösterdiği sonucuna ulaşılmıştır.

Atalay ve arkadaşları (2012) gözlenen puan ve örtük puana dayalı DMF belirleme tekniklerini simülasyon çalışması ile karşılaştırmışlardır. Çalışmada gözlenen puan yöntemleri arasında yer alan MH ve LR; örtük puan yöntemlerinden MTK-OO ve SIBTEST karşılaştırılmıştır. Araştırmada örtük puan yöntemleri ile DMF belirlemenin gözlenen puan ile DMF belirleme yöntemlerine oranla daha duyarlı olduğu sonucuna ulaşılmıştır. Araştırmada MH, SIBTEST ve MTK-OO yöntemlerinin birbiriyle uyumlu olduğu sonucuna ulaşılmıştır. Tek biçimli DMF'yi belirlemede LR, SIBTEST, MTK-OO yöntemleri birbiri ile uyumlu çıkmıştır.

Karakaya ve Kutlu (2012) 2009 yılı Seviye Belirleme Sınavı (SBS) içerisindeki Türkçe alt testlerinin öğrencilerin cinsiyetine ve okul türlerine göre madde yanlılığı gösterip göstermediğini araştırmıştır. Değişen madde fonksiyonlarını belirlemek için Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) yöntemleri kullanılmıştır. Araştırma 22656 öğrenci üzerinde yürütülmüştür. Araştırmada cinsiyet değişkenine göre; 6. sınıfta hiçbir maddede DMF'ye rastlanmamıştır, 7. sınıfta MH tekniğine göre 2 madde, LR tekniğine göre 1 madde, 8. sınıfta MH tekniğine göre 5 madde, LR

teknikğine göre ise 3 maddede DMF bulunduđu sonucuna ulařılmıştır. Okul deđiřkenine göre; 6. sınıfta MH teknikğine göre 7 maddenin, LR teknikğine göre 5 madde, 7. sınıfta MH teknikğine göre 2, LR teknikğine göre 1 maddenin ve 8. sınıfta MH teknikğine göre 2 maddenin DMF'ye sahip olduđuna ulařılmıştır. Yanlılık analizi amacı ile uzman görüşlerine başvurulmuş ve uzman görüşleri dođrultusunda DMF içeren maddeler arasından yalnızca bir maddenin madde yanlılıđı oluřturduđu sonucuna ulařılmıştır.

II. BÖLÜM

YÖNTEM

Bu bölümde, araştırma modeli, çalışma grubu, verilerin elde edilmesi, verilerin analizi ve bu analizlerde kullanılan bilgisayar programları açıklanmıştır. Ayrıca, bu bölümde PISA 2006 verilerine ait betimsel istatistikleri ile kuramların varsayımlarının testlerine ilişkin bilgiler bulunmaktadır.

2.1. ARAŞTIRMA MODELİ

Bu araştırma; PISA bilişsel alanda yer alan maddelerin kültüre ve dile göre yanlılığının farklı tekniklerle incelenmesine yönelik tasarlanmış betimsel bir çalışmadır.

2.2. ÇALIŞMA GRUBU

Bu çalışmada kullanılan PISA 2006 verileri EARGED ve OECD'nin resmi sitesinden elde edilmiştir. Genel olarak; projeye katılan ülkelerde örgün öğretimde kayıtlı olan 15 yaş grubu öğrencilerin bulunduğu tüm okullar PISA projesinin evrenini oluşturmaktadır.

PISA 2006 ulusal raporunda uygulamanın örnekleme deseni oluşturulurken aşağıdaki basamakların izlendiği belirtilmektedir (PISA ulusal rapor, 2007).

Öncelikle tabakalandırma işlemi yapılarak okullar ve öğrenciler OECD tarafından tesadüfi yöntemle seçilmektedir. Örneğin ülkemizde, 2006 yılında yapılan uygulamada, coğrafi bölgelere göre tabakalandırma yapılmış, bu bölgelerdeki okullaşma yüzdesine göre kayıtlı 15 yaş grubu öğrencisi bulunan ve projeye katılacak okul sayısı belirlenmiş, seçkisiz olarak okullar belirlendikten sonra bu okullarda kayıtlı tüm 15 yaş grubu öğrenciler arasından her bir okul için 35 öğrenci yine seçkisiz yöntemle belirlenmiştir. PISA 2006 değerlendirilmesine 57 katılımcı ülkeden yaklaşık 32.000.000 15 yaş gurubu öğrenciyi temsilen yaklaşık 400.000 öğrenci katılmıştır. Türkiye’de yapılan uygulamaya 7 coğrafi bölgemizden bölgelere

ve okul türlerine göre tabakalandırılarak OECD tarafından rastgele seçilen 160 okuldan toplam 4942 öğrenci katılmıştır.

Bu araştırmanın çalışma grubunu, 2006 yılında yapılan PISA araştırma taramasına katılan birinci ve beşinci kitapçıkları alan Türkiye, Avustralya, İngiltere, Kanada öğrenci örnekleme oluşturmaktadır. Kitapçıklarda puanı olmayan veya analiz sonucu aykırı değer sayılan veriler çıkarılmıştır. Örnekleme alınan veriler Tablo 1’de gösterildiği gibi oluşturulmuştur.

Tablo 1: PISA’ya Katılan ve Beşinci Kitapçığı Cevaplayan Öğrenci Sayılarına Ait İstatistikler

Ülke	Sınava Giren Öğrenci Sayısı	5. kitapçığı alan öğrenci sayısı	1. kitapçığı alan öğrenci sayısı
Avustralya	14.170	1124	
Kanada	22.646	1744	
Kanada (İngilizce form)	17.555	1352	
Kanada (Fransızca form)	5.091	392	
İngiltere	13.152	1008	1430
Türkiye	4.942	377	380

Tablo 1’deki istatistikler incelendiğinde, araştırmada karşılaştırılan gruplar için kullanılan örneklemelerin büyüklüğü 377 ile 1744 arasında değişmektedir. Ülkelere ait dağılımda öğrenci sayısı bakımından ilk sırayı Kanada alırken, araştırmada kullanılacak veriler içinde en az öğrenci sayısı Türkiye ‘ye aittir.

PISA 2006 yılı test maddeleri ile öğrencilerin fen bilimleri okuryazarlığını, fen bilimleri değerlendirmesinde yer alan görevlerin ne kadarına sahip oldukları değerlendirilmiştir. Bu nedenle öğrencilere dağıtılan 13 kitapçık türünden birinci ve beşinci kitapçık tüm çoktan seçmeli soruların fen okur-yazarlığı ile ilgili olması ve en çok çoktan seçmeli soru sayısını barındırması bakımından çalışılacak kitapçıklar olarak belirlenmiştir. Soruların kitapçıklara dağılımına ilişkin sonuçlar Tablo 2’de verilmiştir.

Tablo 2: Soruların Kitapçıklara Dağılımı

Kitapçık	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	K11	K12	K13
Soru													
S213Q02*	X				X		X						X
S256Q01				X	X						X	X	
S268Q01		X	X		X				X				
S268Q06		X	X		X				X				
S304Q02					X	X		X		X			
S326Q03	X		X	X		X							
S408Q01	X		X	X		X							
S408Q05	X		X	X		X							
S413Q05				X	X						X	X	
S415Q02	X		X	X		X							
S425Q02	X				X		X						X
S425Q05	X				X		X						X
S426Q03*	X								X	X		X	
S426Q05*	X								X	X		X	
S428Q01					X	X		X		X			
S428Q03					X	X		X		X			
S437Q01	X		X	X		X							
S437Q03	X		X	X		X							
S437Q04	X		X	X		X							
S438Q02					X	X		X		X			
S447Q02*				X	X						X	X	
S447Q03*				X	X						X	X	
S447Q04*				X	X						X	X	
S465Q02				X	X						X	X	
S465Q04				X	X						X	X	
S466Q05					X	X		X		X			
S476Q01	X	X						X			X		
S476Q02	X	X						X			X		
S476Q03	X	X						X			X		
S477Q02*		X	X		X				X				
S477Q03*		X	X		X				X				
S478Q01				X	X						X	X	
S485Q03*	X								X	X		X	
S498Q03		X	X		X				X				
S508Q03*	X								X	X		X	
S521Q02	X	X						X			X		
S521Q06	X	X						X			X		

“*” PISA 2006 açıklanan soruları belirtmektedir.

Tablo 2’de koyu karakterlerle belirlenen sütunlar bu çalışmada verileri kullanılan kitapçıkları temsil etmektedir. Buna göre birinci kitapçıkta yer alan 19 soru ve beşinci kitapçıkta yer alan 21 soru ile çalışma yürütülmüştür. PISA uygulamasında kullanılan 13 kitapçıkta yer alan toplam 37 çoktan seçmeli sorudan 10’u açıklanmıştır. Farklı yöntemlerle yapılan DMF analizleri 37 soru üzerinden, yanlılığın olası nedenlerinin araştırılması ise açıklanan 10 soru üzerinden yürütülmüştür. Tablo 2 incelendiğinde 1.

ve 5. kitapçıkların en fazla çoktan seçmeli maddeye sahip kitapçıklar olduğu görülmektedir. Alt problem üçü yanıtlamak için yürütülecek analizler PISA sınavı için sadece açıklanan maddeler üzerinden yürütülebileceğinden kitapçıklarda açıklanan madde sayısı da önem kazanmaktadır. Yapılacak yanlılık analizi farklı kültür farklı dil gruplaması için yapılacağından bu grup için açıklanan maddelerin hepsine ulaşılabilindiğinden hem birinci ve hem de beşinci kitapçık verileri kullanılacaktır.

1-0 biçiminde puanlanan verilerle yapılacak MTK'ya dayalı yapılacak DMF analizlerinde 1 parametrelili model için veri büyüklüğünün en az 200 ($N > 200$); 2 parametrelili model için, en az 500 ($N > 500$) ve 3 parametrelili model için veri büyüklüğünün en az 1000 ($N > 1000$) olması gerekmektedir (Crocker ve Algina, 1986). Zieky (1993)'e göre ise, DMF istatistiklerinin kullanımı için örneklem büyüklüğü, karşılaştırılan küçük gruplar için (odak ya da referans gruplar için) en az 200, küçük ve büyük gruplar birlikte toplam en az 600 olmalıdır. Bu bilgi referans alınarak örneklem büyüklüğünün yeterli olduğu kabul edilmiş ve analizlere devam edilmiştir.

Bu araştırmada yanlılık çalışması odak ve referans grup temel alınarak yürütülmüştür. Çalışmada farklı veriler kullanılarak yapılan DMF analizinde referans ve odak gruplar aşağıdaki biçimde belirlenmiştir:

1. DMF analizi çalışması: Aynı kültür- Farklı dil
Referans grup 1: :Kanada (İngilizce)
Odak grup 1: Kanada (Fransızca)
2. DMF analizi çalışması: Farklı Kültür Aynı Dil
Referans grup 1: Avustralya
Odak grup 1: İngiltere
3. DMF analizi ve yanlılık çalışması: Farklı Kültür Farklı Dil
Referans grup 1: İngiltere
Odak grup 1: Türkiye
Referans grup 2: İngiltere (1. Kitapçık)
Odak grup 2: Türkiye(1. Kitapçık)

2.3. VERİLERİN ELDE EDİLMESİ

Bu araştırmada 2006 yılı PISA fen başarı testine ait veriler kullanılmıştır. PISA testinde farklı soru türleri kullanılmaktadır. Ancak çalışmada çoktan seçmeli test maddelerine verilen öğrenci cevapları kullanılmıştır. Bu test maddeleri öğrencilerin fen alanındaki yeterliğini, bilgisini doğrudan ya da kapsamlı ilişkilendirerek değerlendirmeyi hedefleyen sorulardır. Tüm PISA 2006 sorularında olduğu gibi çoktan seçmeli fen bilimleri maddelerini yanıtlarken, öğrencilerden bilimsel soruları tanımlamaları, bilimsel olguları açıklamaları ve bilimsel delilleri kullanmaları istenmektedir. Bu üç temel yeterliğin seçilme nedeni, bilimsel uygulamalarda etkili olmaları ve temel zihinsel yeteneklerle ilişkili olmalarıdır. Çalışmada kullanılan ülkelerin verileri ülke ve kitapçık biçiminde gruplanmıştır. Sonrasında fen bilimleri çoktan seçmeli test maddeleri dışında olan tüm veriler silinerek analizler yapılmıştır.

Ayrıca Türkçe ve İngilizce form için açıklanan maddelerdeki olası DMF nedenlerini araştırma amacıyla alan uzmanlarının görüşlerini almak için iki form kullanılmıştır. Birinci form (EK 1) ile açıklanan maddelerin Türkçe ve İngilizce formunun yanlışlık içerip içermediği uzman görüşlerine başvurularak değerlendirilmiştir. Bu formu yanıtlayan uzmanların fen bilimleri alanından mezun olmaları (fen bilimleri alanında öğretmenlik yapmaları) ve iyi derecede İngilizce bilmeleri dikkate alınmıştır. Bu form ile kültür ve dilin yanlışlığa olası etkileri üzerinde durulması planlanmıştır. Birinci form doğrultusunda 15 uzmanın görüşlerine başvurulmuştur. İkinci form (EK 2) ile açıklanan maddelerin Türkçe formunun yanlışlık içerip içermediği konusunda uzman görüşlerine başvurulmuştur. Bu formu yanıtlayan uzmanların fen bilimleri alanından mezun olmaları ve fen bilimleri alanında öğretmenlik yapıyor olmaları yeterli görülmüştür. Bu form için görüş bildiren 30 uzman bulunmaktadır.

2.4. VERİ ÇÖZÜMLEME TEKNİKLERİ

Verilerin çözümlemek için; faktör analizi ve klasik test kuramına dayalı olarak yapılan test istatistikleri; madde tepki kuramına ilişkin analizler ve DMF analizlerinin yapılması için uygun bilgisayar programları kullanılmıştır.

Alt problem 1 ve 2'nin yanıtı olan DMF analizlerine geçmeden önce sırasıyla verilere ait betimsel istatistikler hesaplanmış, madde tepki kuramının sayıltıları kontrol edilmiş ve model veri uyumu test edilmiştir.

Alt problem 1'de kullanılan DMF belirleme yöntemlerinden MH ve LR yöntemlerini uygulayabilmek için her bir alt grupta betimsel istatistikler hesaplanmış ve kullanılan testlerin güvenilir olup olmadıkları belirlenmiştir. Alt problem 1'de kullanılan diğer bir teknik madde tepki kuramı yöntemlerinden biri olan alan indeksleri yöntemini kullanıp kullanamayacağımızı belirlemek, kullandığımız taktirde mevcut sınırlılıklarımızı görebilmek amacı ile madde tepki kuramının sayıltıları kontrol edilmiştir. Bunun için SPSS programı kullanılmıştır. Sonrasında ise BILOG programı yardımı ile model veri uyumu test edilerek uygun modele karar verilmiştir. Uygun modelle elde edilen parametreler yardımıyla alan indeksleri tekniği için kullanılacak DMF ölçütleri hesaplanmıştır.

Tüm bu kontrol ve işlemlerden sonra DMF analizlerine geçilmiştir. MH ve LR analizi için EZDIFF programı; alan indeksleri tekniği için BILOG OLD ve Frans Oort tarafından geliştirilen CALCBIAS programı kullanılmıştır.

Alt problem 2'ye yanıt bulabilmek için alt problemde hesaplanan DMF değerleri arasındaki sıra farkları korelasyonlarından yararlanılmıştır. Aynı zamanda "DMF sonuçları tekniklere göre farklılık göstermekte midir?" sorusunun cevabı Cohran q testi ile de araştırılmıştır.

Alt problem 3 için ise uzman görüşleri ve uzman görüşlerine ait yüzde ve frekanslar elde edilerek olası DMF sebepleri belirlenmeye çalışılmıştır. Farklı formlara yanıt veren uzmanların yanlılık kararlarının birbirine benzer olup olmadığını belirlemek amacıyla her bir soru için kısmi mertebe katsayısı olarakta bilinen Phi katsayısı hesaplanmıştır.

Alt problemlere yanıt aramadan önce yapılacak DMF analizlerinin yapılma yapılmama kararında etkili olacak ve alınacak kararın doğruluğunun yorumlanabileceği ön analizler aşağıdaki başlıklar altında belirtilmiştir (2.4.1/2/3/4).

2.4.1. PISA–2006 Verilerine Ait Betimsel İstatistiklerin Hesaplanması

Verilerin analizlere uygunluğunu değerlendirebilmek ve veri yapısı hakkında bilgi sahibi olabilmek için bir takım ön analizlere başvurulmuştur. Bu amaçla öncelikle kullanılan PISA–2006 verilerine ait betimsel istatistikler hesaplanmıştır. Kitapçıklara ve alt gruplara göre ayrı ayrı elde edilen sonuçlar tablo 3’te verilmiştir.

Tablo 3: Araştırmadaki Çalışma Gruplarına İlişkin Betimsel İstatistikler

	Ülke adı	N	Ortalama	Basıklık	Çarpıklık	S.S	KR-20
5.Kitapçık	AVUSTRALYA	1067	13,98	-,457	-,495	4,13	0,78
	KANADA	1647	13,78	-,664	-,421	4,25	0,79
	TÜRKİYE	333	10,38	-,568	,326	4,08	0,74
	İNGİLTERE	968	12,84	-,927	-,188	4,45	0,81
	TOPLAM	3829	13,35	-,781	-,328	4,36	0,80
1.Kitapçık	TÜRKİYE	365	9,99	-,728	,288	3,84	0,74
	İNGİLTERE	790	12,83	-,563	-,421	3,82	0,79
	TOPLAM	1158	11,93	-,871	-,195	4,05	0,80

Tablo 3’te yer alan değerler incelendiğinde; beşinci kitapçığı alıp çalışma grubunu oluşturan ülkeler arasında en düşük ortalamaya Türkiye’nin, en yüksek ortalamaya ise Avustralya’nın sahip olduğu görülmektedir. Dağılımların normalliği hakkında bilgi edinmek için çarpıklık ve basıklık değerlerine bakılmıştır. Çarpıklık; dağılımın ortalama etrafında simetriden ne kadar saptığını belirleyen bir ölçüdür. Basıklık ise dağılımın dikliği veya düzlüğü hakkında bilgi verir. Çarpıklık değeri sıfıra yakın ise normal dağılıma yakın, pozitif ise puanların ortalamanın altında yığıldığını, negatif ise puanların ortalamanın üzerinde yığıldığını gösterir. Basıklık değeri sıfıra yakın ise normal dağılıma yakın, pozitif ise normalden daha dik, negatif ise normalden daha düz bir şekil oluşturur. Beşinci kitapçığı alan Türkiye örneklemini dışındaki verilere ilişkin başarı puanları dağılımı, standart normal dağılıma göre çok hafif sola çarpık ve basıktır. Çarpıklık değerinin sıfıra yakın bulunmuş olması verilerin oluşturduğu dağılımın simetriğe çok yakın bir şekil aldığı göstergesi olarak kabul edilebilir. Tablo 3’teki dağılımların biçimi, çarpıklık ve basıklık değerleri sıfır değerine yakın olduğundan; dağılımların gruplar için simetrik ve normale yakın olduğu kabul edilebilir. Testin hesaplanan güvenilirlik katsayısının alt gruplar için 0,74 ile 0,81 arasında değiştiği, kullanılan verilerin güvenilir olduğu söylenebilir.

2.4.2. Madde Tepki Kuramının Sayıtlılarının Kontrolü

Alan indeksleri tekniği ile değişen madde fonksiyonu (DMF) belirleme işlemleri için Madde Tepki Kuramının (MTK) sayıtlıları kontrol edilmiştir.

MTK'nın en önemli sayıtlılarından biri kullanılan testlerin tek boyutlu bir niteliği ölçmesidir (Hambleton ve Swaminathan, 1985). Bu sayıtlı madde setine uygulanan temel bileşenler analizi ile özdeğerlerin incelenmesi yoluyla test edilebilir. Bir madde setinin temel bileşenler ile analizi sonrası birden fazla faktörün özdeğeri birden büyük olabilir. Bu MTK modellemelerindeki “Tekboyutluluk” varsayımının o veri seti için varlığını tamamen ortadan kaldırmaz. Bu varsayımı belirlemek için ikinci bir yol da yamaç-eğim grafiğinin incelenmesidir (Cattell, 1966; 1978; Loehlin, 1987). Bu nedenle tek boyutluluk sayıtlısını kontrol etmek için, PISA'ya ait veriler faktör analiziyle incelenmiştir. Yapılan analiz sonucunda tüm gruptan elde edilen özdeğerler dikkate alınmıştır. Tek boyutluluk aynı zamanda testteki maddelerin yerel bağımsızlığı için de kanıt olarak kullanılmıştır.

Kitapçıklarda üzerinde çalışılan soruların tek boyutluluk ve yerel bağımsızlık analizler temel bileşenler faktör analizi ile yapılmıştır. Bir testteki tüm maddeler aynı özelliği ölçüyorsa, bireyin testten aldığı puan, onun bu boyuttaki (uzaydaki) yerini gösterir. Bir ölçme aracının tek boyutlu olup olmadığının kontrolü konusunda faktör analizi teknikleri kullanılabilir (Hambleton ve Swaminathan, 198, s:21). Temel bileşenler faktör analizi yöntemi, korelasyon matrisi tarafından açıklanan ortak varyansı dikkate alır. Değişkenlerin paylaştıkları ortak varyansın ortak faktörlerden kaynaklandığını ve gözlenebilen ve ölçülebilen çok sayıda özelliğin arkasında yatan ölçülemeyen gerçek nedenlerin (gizil boyutların) olduğu varsayılmaktadır (Cattell, 1966). IRT modellemeleri ile ilgili çalışmalarda birinci ve ikinci özdeğer arasındaki farklılığın fazlalığı ve yamaç eğim grafiğindeki keskin düşüş tek boyutluluk için kanıt olarak gösterilmektedir (Lord, 1980; Orlando, Sherbourne, ve Thissen, 2001; Doğan ve Öğretmen. 2004, 2006, 2008).

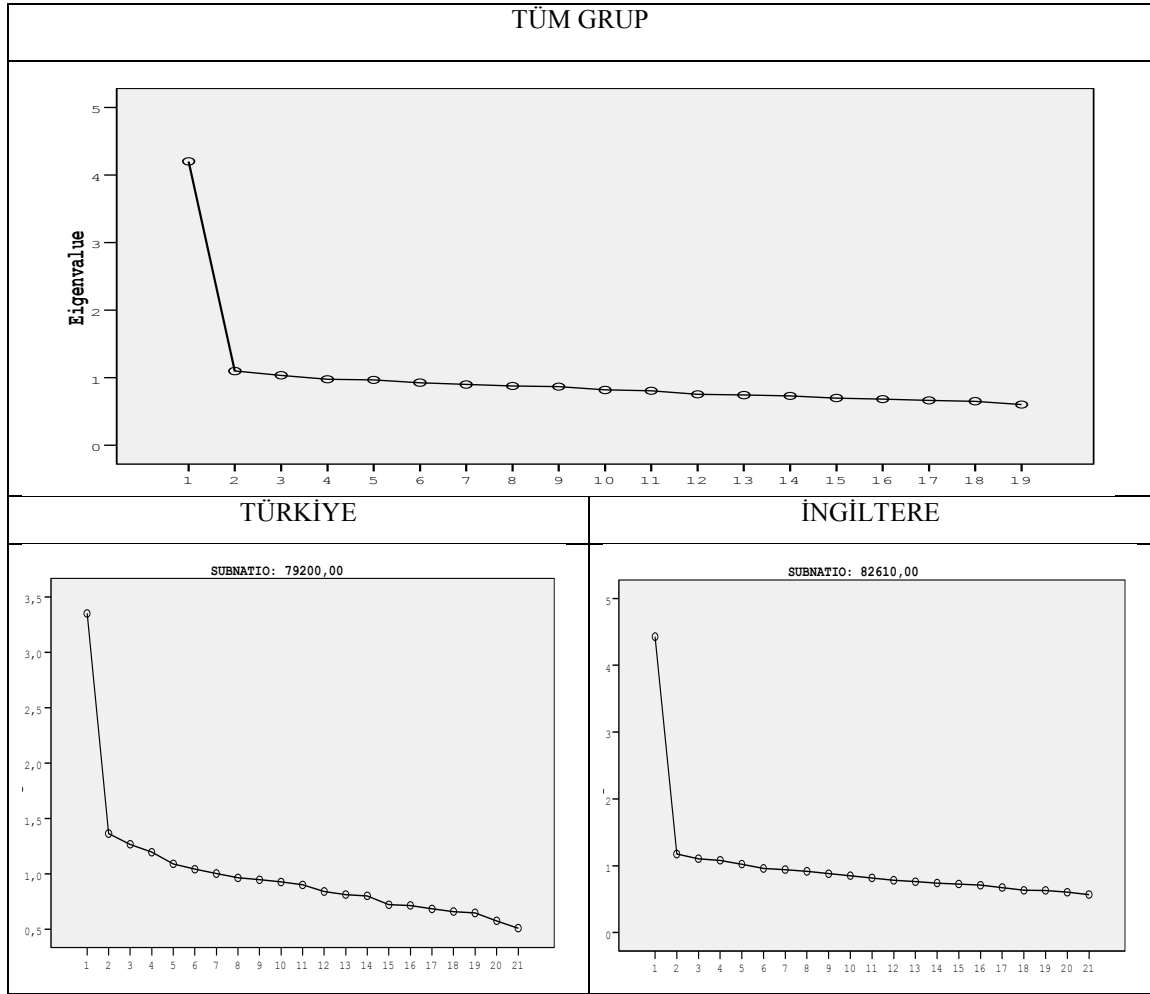
Birinci kitapçıkta yeralan 19 soruluk testin tek boyutluluk ve yerel bağımsızlık değerlendirmeleri için özdeğerlere ilişkin faktörler ile açıkladığı varyans oranları Tablo 4'te gösterilmiştir.

Tablo 4: PISA-2006 Birinci Kitapçıkta Yeralan 19 Soruluk Teste Ait Faktör Analizi Sonuçları

		Özdeğer	Açıklanan Varyans
TÜM GRUP	1. faktör	4,199	22,102
	2.faktör	1,098	5,778
	3. faktör	1,036	5,452
İNGİLTERE	1. faktör	4,153	21,855
	2.faktör	1,177	6,196
	3. faktör	1,046	5,507
TÜRKİYE	1. faktör	3,535	18,605
	2.faktör	1,279	6,732
	3. faktör	1,186	6,242
	4. faktör	1,136	5,979

Tablo 4 incelendiğinde birinci kitapçık fen bilgisi alt testinin birinci faktörü açıklama varyansı %22.102 ve ikinci faktörü açıklama varyansı %5,778’dir. Ayrıca yamaç eğim grafiği (scree plot) ile birinci ve ikinci faktörün özdeğerleri incelendiğinde; birinci özdeğerden ikinci özdeğere keskin bir düşüş söz konusudur. Bu keskin düşüş tek boyutluluğa kanıt oluşturmaktadır. Aslında tek boyutluk varsayımın karşılanması oldukça güçtür. Çünkü uygulama esnasında ölçülmek istenen özellikle ilgili ya da ilgisiz çok sayıda etken ölçülmek istenen değişkenin belirgin bir biçimde ortaya konmasını engeller. Yapılan faktör analizi sonuçlarına göre, ilk özdeğer ile ikinci özdeğer arasındaki farkın büyük; ikinci ile diğer özdeğerlerin birbirine yakın değerler almasından dolayı testin tek boyutlu olarak yorumlanması mümkündür. Tek boyutluluk aynı zamanda testteki maddelerin yerel bağımsızlığı için de kanıt sayılabilir (Lord ve Novic 1968; Hambleton and Swaminathan 1985; Baker 2001).

Tablo 4’teki birinci kitapçığı alan alt gruplar için yapılan faktör analizi incelendiğinde, birinci özdeğer ile ikinci özdeğer arasındaki farkın büyük; ikinci ile diğer özdeğerlerin birbirine yakın değerler aldığı görülmektedir. Yapılan faktör analizi sonucu incelendiğinde birinci ve ikinci faktör arasındaki keskin düşüşler tek boyutluluk için kanıt olarak kabul edilmiştir. Şekil 3’te birinci kitapçığı alan tüm grup; Türkiye ve İngiltere için elde edilen özdeğer-bileşen grafikleri bulunmaktadır.



Şekil 3: PISA Birinci Kitapçıkta Yeralan 19 Soruluk Teste Ait Özdeğer-Bileşen Grafiği

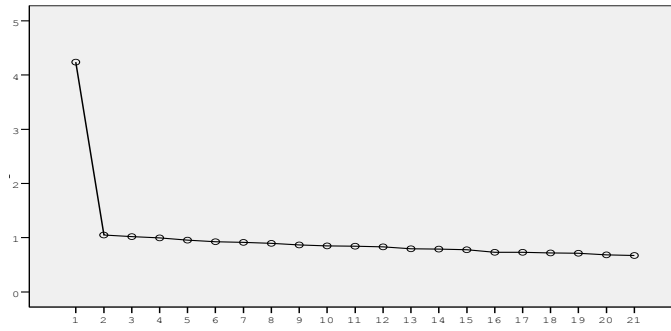
Çalışmada kullanılan birinci kitapçığı yanıtlayan Türkiye ve İngiltere veri grubunun 19 soruluk teste ait özdeğer-bileşen grafiği (Tüm Grup) incelendiğinde dikey eksenindeki birinci faktörün özdeğeri ile ikinci faktörün özdeğeri arasındaki keskin düşüş dikkati çekmektedir. Şekil 3'te çizilen özdeğer bileşen grafiklerinde de her iki ülke elde edilen birinci ve ikinci özdeğer arasındaki keskin düşüş açıkça görülmektedir.

Çalışmada kullanılan beşinci kitapçığı yanıtlayan Avustralya, Kanada, Türkiye ve İngiltere veri grubun oluşturduğu 21 soruluk teste ait belirlenen ilk dört faktöre ait özdeğerler ve açıkladıkları toplam varyans tablo 5'de açık bir biçimde görülmektedir.

Tablo 5: PISA-2006 Beşinci Kitapçıkta Yeralan 21 Soruluk Teste Ait Faktör Analizi Sonuçları

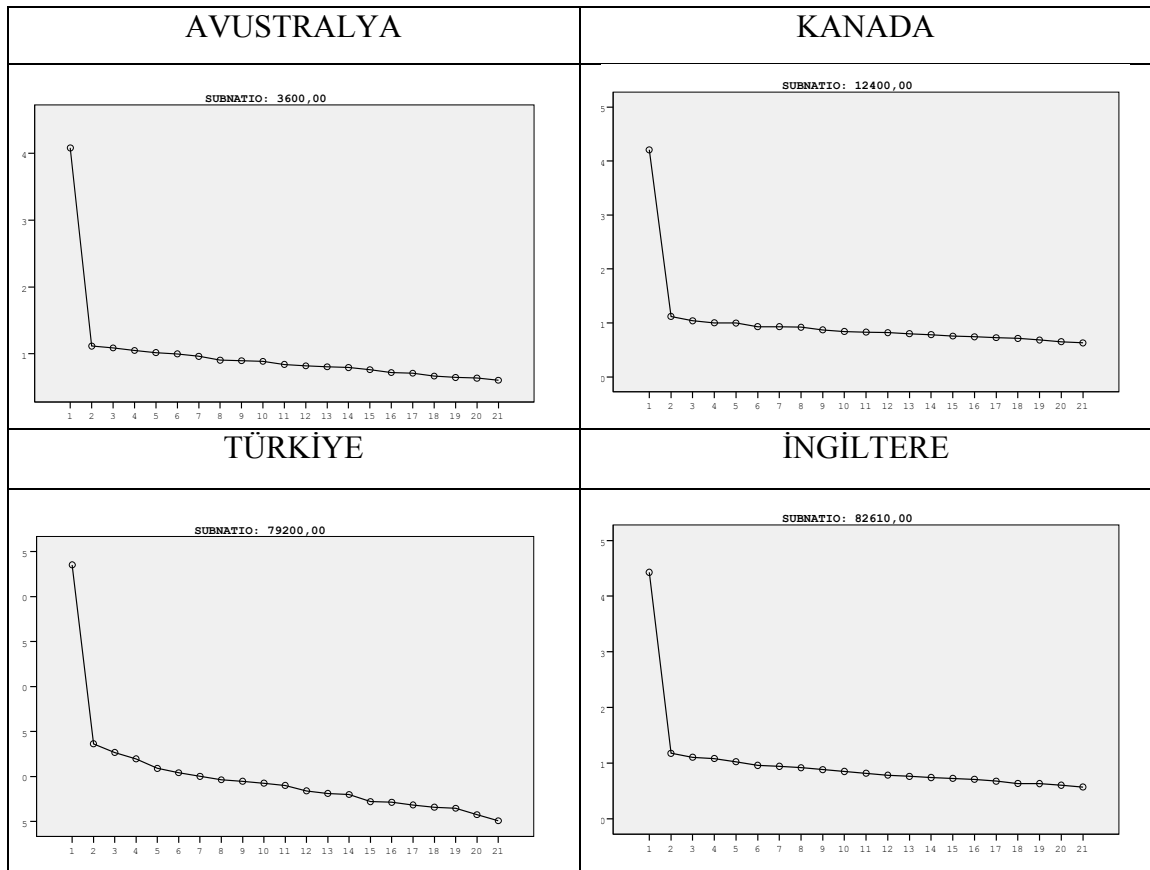
		Özdeğer	Açıklanan Varyans
TÜM GRUP	1. Faktör	4,37	20,811
	2.Faktör	1,023	4,872
	3. Faktör	1	4,761
AVUSTRALYA	1. Faktör	4,152	19,773
	2.Faktör	1,11	5,285
	3. Faktör	1,101	5,243
	4. Faktör	1,043	4,98
KANADA	1. Faktör	4,301	20,481
	2.Faktör	1,065	5,072
	3. Faktör	1,013	4,825
	4. Faktör	1,01	4,81
TÜRKİYE	1. Faktör	3,524	16,782
	2.Faktör	1,4	6,665
	3. Faktör	1,265	6,023
	4. Faktör	1,153	5,491
İNGİLTERE	1. Faktör	4,519	21,517
	2.Faktör	1,151	5,481
	3. Faktör	1,094	5,21
	4. Faktör	1,044	4,969

Tek boyutluluk varsayımının karşılanıp karşılanmadığını belirlemek üzere temel bileşenler faktör analizi ile döndürülmemiş faktör çözümlemesi uygulanmıştır. Analiz sonucuna göre fen bilgisi alt testinin birinci faktörü açıklama varyansı %20.81 birinci ve ikinci faktörü açıklama varyansı % 4,8'dir. Ayrıca özdeğer-bileşen grafiği incelendiğinde ilk özdeğerden ikinci özdeğere keskin bir düşüş söz konusudur. Bu keskin düşüş tek boyutluluğa da kanıt oluşturmaktadır. Tablo 5'deki alt gruplar için yapılan faktör analizi sonuçları incelendiğinde tüm alt gruplar için ilk boyuttaki özdeğerlerin ikinci boyuttan büyük olduğunu, ikinci ve üçüncü boyutlardaki özdeğerlerin büyük ölçüde benzer büyüklükte olduğu bulunmuştur. Beşinci kitapçıkta yeralan 21 soruluk testin tek boyutluluk değerlendirmeleri için özdeğerlere ilişkin özdeğer- bileşen grafiği sonuçları (scree plot) şekil 4'de (yatay eksen bileşenleri, dikey eksen ise özdeğerleri göstermektedir.)



Şekil 4: PISA Beşinci Kitaplık ta Yeralan 21 Soruluk Teste Ait Özdeğer-Bileşen Grafiği

Şekil 5’da bu düşüşlerin grafiksel olarak görüldüğü alt gruplara ilişkin özdeğer-bileşen grafikleri bulunmaktadır



Şekil 5: Alt Gruplar İçin PISA Beşinci Kitaplık ta Yer Alan 21 Soruluk Teste Ait Özdeğer-Bileşen Grafiği

Şekil 5’da yer alan her bir alt grup için ayrı ayrı çizilen özdeğer-bileşen grafikleri ve özdeğerler incelendiğinde grafikleri de ilk boyuttaki özdeğerlerin her iki test içinde ikinci boyuttan büyük olduğunu, ikinci ve üçüncü boyutlardaki özdeğerlerin büyük ölçüde benzer büyüklükte olduğu gözlenmiştir. Birinci boyuttan ikinci boyuta özdeğerlerdeki keskin düşüş, testlerin tek boyutlu olduğuna işaret eden bir kanıt olarak kullanılmıştır.

2.4.3. Model Uyumu

Bu aşamaya kadar olan kısımda PISA fen testinden elde edilen verilerin MTK’ya ilişkin temel varsayımları sağlayıp sağlamadığına bakılmıştır. Bu aşamada ise her model için yetenek kestirimlerine yönelik elde edilen $-2 \log$ (olabilirlik) değerlerinin farkını almak suretiyle bir uyum iyiliği testi (“ χ^2 ” Kay Kare) yapılarak hangi modelin uyumlu olduğuna karar verilmiştir.

2.4.3.1. Birinci Kitapçık İçin Model-Veri Uyumu

Birinci kitapçığa ait verilerle, her model için BILOGMG 3.0 yazılımı ile ayrı ayrı elde edilen $-2 \log$ (olabilirlik) değerleri tablo 8’de bulunmaktadır.

Tablo 6: Birinci Kitapçık Model Veri Uyumu İçin Hesaplanan $-2 \log$ (Olabilirlik) Değerleri

Model	$-2 \log$ (olabilirlik)
Bir Parametrel Model (1PLM)	24428.1389
İki Parametrel Model (2PLM)	24275.5208
Üç Parametrel Model (3PLM)	24211.9699

Tablo 6’da görülen bir parametrel model ve iki parametrel modele ilişkin $-2 \log$ (olabilirlik) değerlerinin farkı alındığında;

$$\chi^2 = -2\log \text{Olabilirlik}_{1\text{PLM}} - (-2\log \text{Olabilirlik}_{2\text{PLM}})$$

$$\chi^2 = 24428.1389 - 24275.5208 = 152,6181$$

sonucuna ulaşılmıştır. Serbestlik derecesi olarak daha karmaşık modele eklenen parametre sayısı alınmıştır. Buna göre iki parametrel modele testteki madde sayısı

kadar, yani 19 adet “a parametresi” eklendiği göz önüne alınırsa, serbestlik derecesi 19’dur. Anlamlılık düzeyi 0,05 alındığında χ^2 ’nin tablo değeri; $\chi^2_{19,0,05} = 31.41$ ’dir. Bu sonuca bakılarak; iki parametrelili modelde yapılan yetenek kestiriminin olabirliğin, bir parametrelili modelde yapılan yetenek kestiriminden anlamlı olarak daha iyi olduğı söylenebilir. Benzer olarak yine Tablo 8’de görülen iki parametrelili model ve üç parametrelili modele ilişkin $-2 \log$ (olabilirlik) değeriinin farkı alındığında;

$$\chi^2 = -2\log \text{Olabilirlik}_{2\text{PLM}} - (-2\log \text{Olabilirlik}_{3\text{PLM}})$$

$$\chi^2 = 24275.5208 - 24211.9699 = 63,5509$$

sonucuna ulaşılmıştır. Serbestlik derecesi olarak daha karmaşık modele eklenen parametre sayısı alınmıştır. Buna göre üç parametrelili modele testteki madde sayısı kadar, yani 19 adet “c parametresi” eklendiği göz önüne alınırsa, serbestlik derecesi 19’dur. Anlamlılık düzeyi 0,05 alındığında χ^2 ’nin tablo değeri; $\chi^2_{19,0,05} = 31.41$ ’dir. Bu sonuca göre; χ^2 değeri 19 serbestlik derecesinde anlamlı olduğı için, üç parametrelili modelle elde edilen yetenek kestiriminin, iki parametrelili modelden daha iyi olduğı söylenebilir. Bu durumda PISA Fen Testi’ne ait veriler için en uygun modelin üç parametrelili olduğunu söylemek mümkündür.

2.4.3.2. Beşinci Kitapçık İçin Model Veri Uyumu

Bu aşamaya kadar olan kısımda PISA Fen Testi’nden elde edilen verilerin MTK ‘ya ilişkin temel varsayımları sağlayıp sağlamadığına bakılmıştır. Bu aşamada ise her model için yetenek kestirimlerine yönelik elde edilen $-2 \log$ (olabilirlik) değeriinin farkını almak suretiyle bir uyum iyiliği testi (“ χ^2 ” Ki- Kare) yapılarak hangi modelin uyumlu olduğuna karar verilmiştir. Her model için BILOGMG 3.0 yazılımı ile elde edilen $-2 \log$ (olabilirlik) değeri Tablo 9’da sunulmuştur.

Tablo 7: Beşinci Kitapçık Model Veri Uyumu İçin Hesaplanan -2 Log (Olabilirlik) Değeri

Model	$-2 \log$ (olabilirlik)
Bir Parametrelili (1PLM)	96348.5878
İki Parametrelili (2PLM)	95598.4384
Üç Parametrelili (3PLM)	95445.8956

Tablo 7’de görülen bir parametrelili model ve iki parametrelili modele ilişkin $-2 \log$ (olabilirlik) değerlerinin farkı alındığında;

$$\chi^2 = -2\log \text{Olabilirlik}_{1\text{PLM}} - (-2\log \text{Olabilirlik}_{2\text{PLM}})$$

$$\chi^2 = 96348.5878 - 95598.4384 = 750,1494$$

sonucuna ulaşılmıştır. Serbestlik derecesi olarak daha karmaşık modele eklenen parametre sayısı kullanılmıştır. Buna göre iki parametrelili modele testteki madde sayısı kadar “a parametresi” eklendiği göz önüne alınırsa, serbestlik derecesi 21’dir. Anlamlılık düzeyi 0,05 alındığında χ^2 ’nin tablo değeri; $\chi^2_{21;0,05} = 32.67$ ’dir. Bu sonuca bakılarak; iki parametrelili modelde yapılan yetenek kestiriminin olabilirliğinin, bir parametrelili modelde yapılan yetenek kestiriminden anlamlı olarak daha iyi olduğu söylenebilir. Benzer olarak yine tablo 9’da görülen iki parametrelili model ve üç parametrelili modele ilişkin $-2 \log$ (olabilirlik) değerlerinin farkı alındığında;

$$\chi^2 = -2\log \text{Olabilirlik}_{2\text{PLM}} - (-2\log \text{Olabilirlik}_{3\text{PLM}})$$

$$\chi^2 = 95598.44 - 95445.90 = 152,54$$

sonucuna ulaşılmıştır. Serbestlik derecesi olarak daha karmaşık modele eklenen parametre sayısı (21 adet “c parametresi”) kullanıldığında, modelin serbestlik derecesi 21 olacaktır. Anlamlılık düzeyi 0,05 alındığında χ^2 ’nin tablo değeri; $\chi^2_{21;0,05} = 32.67$ ’dir. Bu sonuca göre; χ^2 değeri 21 serbestlik derecesinde anlamlı olduğu için, üç parametrelili modelle elde edilen yetenek kestiriminin, iki parametrelili modelden daha iyi olduğu söylenebilir. Bu durumda PISA fen testine ait veriler için en uygun modelin üç parametrelili olduğunu söylemek mümkündür.

Yapılan uyum testi sonucunda her iki kitapçığa ilişkin veriler için, parametrelerinin kestirimlerinde üç parametrelili lojistik model kullanılmıştır. Üç parametrelili lojistik modelin madde karakteristik fonksiyonu aşağıdaki eşitlikle gösterilir. (Hambleton ve Swaminathan 1985, 49)

$$P_i(\theta_j) = c_i + (1 - c_i) / [1 + \text{Exp}(Da_i(\theta - b_i))]^{-1}$$

Eşitlik2

Yukarıdaki eşitlikte ; “D” ölçekleme sabitini temsil etmektedir ve sayısal değeri 1.7’dir; “ a_i ” i maddesinin ayırtediciliğini, “ b_i ” i maddesinin güçlüğünü ve “ c_i ” i maddesinin şansla doğru cevaplanmasını vermektedir.

2.4.4. Değişen Madde Fonksiyonu Analizleri

Bu analizler alt problem 1 ve alt problem 2’ye yanıt bulabilmek için yürütülmüştür. DMF belirlemek için Klasik Test Kuramı (KTK) DMF belirleme yöntemlerinden Mantel Haenzel (MH) ve Logistik Regresyon (LR) ve Madde Tepki Kuramı (MTK) DMF belirleme yöntemlerinden alan indeksleri teknikleri kullanılmıştır.

Aynı biçimde uygun model kullanılarak uzmanlarca seçilen maddelere ilişkin odak ve referans grupları için madde karakteristik eğrileri elde edilerek ve eğriler altında kalan alanların hesaplanması yoluyla maddelerin DMF içerip içermediği test edilmiştir.

Son olarak KTK ve MTK yöntemleri ile elde edilen DMF bulgularının birbiriyle uyumlu olup olmadığı test edilmiştir.

2.5. ARAŞTIRMADA KULLANILAN TEKNİKLER

2.5.1. Mantel – Haenszel Tekniği

Bu yöntemde grupların puanları eşleştirilir ve belirli bir testteki tüm maddeler için iki cevaplayıcı grubunun performansı karşılaştırılır. MH tekniğinde kurulan yokluk hipotezi benzer yetenek düzeyindeki bireylerin yer aldıkları grup ve testte yer alan bir madde için gösterdiği performans arasında herhangi bir ilişki yoktur biçimindedir. Bir ki-kare tekniği olan MH yaklaşımında bir maddeye verilen cevaplar iki grup için aşağıdaki gibi gösterilebilir. MH istatistiği eşitlik (3) ve (4) kullanılarak hesaplanabilir.

Madde Cevaplarının İki Grup İçin Dağılımı:

Gruplar	Doğru (1)	Yanlış (0)	Toplam
Grup 1	A	B	A+B
Grup 2	C	D	C+D
Toplam	A+C	B+D	T

A: 1. grupta maddeyi doğru olarak cevaplayanların sayısı

B: 1. grupta maddeyi yanlış cevaplayanların ya da boş bırakanların sayısı

C: 2. grupta maddeyi doğru olarak cevaplayanların sayısı

D: 2. grupta maddeyi yanlış cevaplayanların ya da boş bırakanların sayısını gösterir.

$$\Omega = MH = AD/BC$$

Eşitlik 3

$$\Delta\Omega = \Delta MH = -(4/1,7) \ln MH = -2,35 \text{ logit}$$

Eşitlik 4

Eşitlik (3)'den elde edilen sonuçlar:

“ $MH > 1$ ” yanlılık referans grup lehine,

“ $MH < 1$ ” yanlılık ilgilenilen (focal=odak) grup lehine

“ $MH \cong 1$ ” ise yanlılık yoktur

şeklinde yorumlanmaktadır. MH istatistiğinin daha kolay yorumlanabilmesi için logaritmik dönüşüm yapılmaktadır. Logaritmik dönüşüm formülü eşitlik (4)'de görülmektedir. Eşitlik (4)'e göre elde edilen sonuçlar:

“ $\Delta_{MH} > 0$ ” yanlılık ilgilenilen grup lehine

“ $\Delta_{MH} < 0$ ” yanlılık referans grup lehine

“ $\Delta_{MH} \cong 0$ ” ise yanlılık yoktur

şeklinde yorumlanmaktadır. Ayrıca, MH büyüklüğüne göre yanlılığın düzeyi hakkında da yorum yapılabilmektedir.

$I\Delta_{MH}I < 1$ ise A (önemsiz) düzeyde,

$1 \leq I\Delta_{MH}I < 1,5$ ise B (orta) düzeyde,

$I\Delta_{MH}I \geq 1,5$ ise C (yüksek) düzeyde yanlılık

olduğundan söz edilmektedir (Dorans ve Holland, 1992; Holland ve Thayer, 1988). Bu teknik DMF'li maddelerin belirlenmesinde, maddenin hangi grup lehine DMF içerdiğinin saptanmasında ve DMF'nin büyüklüğünü belirlemede oldukça etkin bir teknik olarak kabul edilmektedir (Holland ve Thayer, 1988; Hambleton ve Rogers, 1989). MH yönteminin uygulanmasının kolay olması, küçük gruplar için kullanışlı olması ve istatistiksel olarak yorumlanabilmesi bakımından DMF belirlerken önerilen

ve kullanımı yaygınlaşan bir tekniktir. Ancak bu tekniğin ÇBDMF'nin biçimini belirlenmesinde etkisiz kaldığı belirtilmiştir (Swaminathan ve Rogers, 1990).

2.5.2. Logistik Regresyon Yöntemi

LR, sürekli olmayan (ikili veya kategorik) bir bağımlı değişkeni sürekli bağımsız değişkenler yardımıyla tahmin etmeyi amaçlayan istatistiksel bir modeldir. Lojistik regresyon yardımıyla, bağımsız değişkenlerin bağımlı değişkenin varyansının yüzde kaçını açıklayabildiği, bağımsız değişkenlerin bağımlı değişken üzerindeki etki büyüklüğü ve bağımsız değişkenler arasında etkileşim olup olmadığı belirlenebilir. Lojistik regresyon genellikle bir test veya alt test puanı ölçüt alınarak, aynı yetenek düzeyindeki farklı grup üyelerinin bir maddeye doğru cevap verme olasılıklarının istatistiksel olarak modellenmesine dayanır. LR yöntemi, DMF belirleme çalışmalarında çeşitli yöntemler arasında en etkili ve önerilen yöntemlerden biridir (Camilli ve Shepard, 1985).

Bu teknik kurulan modeller sayesinde maddeye verilen cevaplar ve toplam puan üzerinden DMF'nin sınanmasını sağlar. Regresyon denkleminde grupların etkisinin manidarlığı tek biçimli DMF'yi, grup ve toplam puan etkileşiminin manidarlığı tek biçimli olmayan DMF'yi gösterir (Doğan ve Öğretmen; 2008). DMF analizinde değişkenler modele hiyerarşik olarak sırayla dahil edilirler. Model 1'de sınırlanacak değişken, yani toplam puan "X" modelde yerini alır. Model 2'de, grup değişkeni (cinsiyet, dil, ülke) eklenir ve üçüncü modelde ise önceki değişkenlere ilave olarak etkileşim değişkeni modele alınır. Kurulan bu modeller sayesinde değişen madde fonksiyonunda araştırılan maddenin tek biçimli mi yoksa çok biçimli mi fonksiyonlaşıp fonksiyonlaşmadığı belirlenebilir. DMF'nin düzeyine karar vermek için kullanılan ölçütler vardır. Üzerinde uzlaşılan ve çalışmalarda kullanılan ölçütlerden ikisine ait değerler tablo 10'da bulunmaktadır. Bu çalışmada da DMF düzeyi belirlenirken tabloda bulunan ölçütlerden Gierl ve arkadaşlarının belirlemiş olduğu ölçütler kullanılmıştır.

Tablo 8: Lojistik Regresyon Tekniğinde Değişen Madde Fonksiyonu Düzeyi İçin R^2 Ölçütleri

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

Lojistik regresyonda bu üç aşama için ki-kare değerleri hesaplanır. Hesaplanan bu ki-karelerin manidarlık düzeyi bize DMF ve önemi hakkında bilgi verebilir. Tek biçimli DMF sınaması ikinci ve birinci aşamalar arası farkı gösterir. Tek biçimli olmayan DMF'nin varlığına ilişkin kanıt, üçüncü adım için elde edilen ki-kare değeriyle birinci adımdaki ki-kare değerinin farkını alarak elde edilebilir. Ortaya çıkan bu yeni ki-kare değeri 2 serbestlik dereceli ki-kare dağılım fonksiyonu (tablo değeri) ile karşılaştırılır. İki serbestlik derecesi birinci adımdaki modelin 1 serbestlik derecesi ve 3. adımdaki modelin 3 serbestlik derecesinin farkının bir sonucu olarak ortaya çıkar. İki serbestlik dereceli kay kare testi sonuçları tek biçimli ve tek biçimli olmayan DMF'nin eşzamanlı olarak sınanmasını sağlar (Swaminathan ve Rogers 1990).

LR ve MH tekniklerinin MTK'ya dayalı tekniklere göre daha küçük örneklem gerektirmesi de bu tekniklerin bir avantajı olarak görülebilir (Clauser ve Mazor, 1998; aktaran: Gök; Penfield ve Camilli, 2007).

2.5.3. Madde Tepki Kuramı Alan İndeksleri Tekniği

Alt grupların madde karakteristik eğrileri altında kalan alanların farkları DMF olarak tanımlanabilmektedir (Samejima, 1977). MTK modellerine göre, yetenekler eşitlendikten sonra, iki grubun madde karakteristik eğrilerinde farklılık gözleniyorsa madde için yanlılıktan söz edilebilir. Giriş bölümünde değişen madde fonksiyonu bölümünde maddelerin gruplarda tek biçimli ve çok biçimli fonksiyonlaşmasına ilişkin

bilgiler verilmişti. Bu çalışmada MTK yöntemlerinden madde karakteristik eğrilerinin biçimlerine göre hesaplanan alan indeksleri tekniği kullanılacaktır.

İşaretli ve işaretsiz alan indeksleri genel formülleri eşitlik 4 ve 5’de görünmektedir.

$$\text{İşaretsiz alan indeksi: } B = \int_p^q |P_{1i}(\theta) - P_{2i}(\theta)| d\theta \quad \text{Eşitlik 4}$$

$$\text{İşaretli alan indeksi: } B_i = \int_p^q \{P_{1i}(\theta) - P_{2i}(\theta)\} d\theta \quad \text{Eşitlik 5}$$

Bu eşitliklerde iki farklı grup için elde edilen madde karakteristik eğrileri arasındaki farklar alınmakta ve bu farklar alan indeksi olarak yorumlanmaktadır. Formüllerdeki p ve q değerleri yetenek ranjını göstermekte ve bu da genellikle +3 ve -3 aralığında değer almaktadır. İşaretli alan indeksleri maddenin hangi grup lehine işlediğini de göstermektedir. Başka bir deyişle bu indeks değeri DMF’nin yönü hakkında bilgi vermektedir. Eğer $P_{1i}(\theta)$ değeri $P_{2i}(\theta)$ ’den büyükse alan indeksi artı işaret almakta ve maddenin birinci grubun lehine çalıştığını göstermektedir. Eksi işaret ise maddenin ikinci grubun lehine işlediğine işaret etmektedir (Mellenberg, 1989; Camilli ve Shepard, 1994).

Bu teknikte elde edilen İşaretli ve işaretsiz alan indekslerinin birbirlerine göre durumuna bakılarak DMF’nin biçimi hakkında karar verilebilir. Çok biçimli değişen madde fonksiyonlarında odak ve referans grubun madde karakteristik eğrileri kesişir. Böyle bir durumda işaretli ve işaretsiz alan indeks değerleri arasındaki farkın yüksek olduğu ve işaretsiz alan indekslerinin sıra dışı büyüklüklerde olduğu gözlenir. Alan indekslerinin sıra dışı büyüklüklerde olması yanlılığın güçlü bir kanıtı olarak kabul edilmektedir (Raju, 1988; Gondal, 2001).

MTK’ya bağlı alan indeksleri tekniğine bağlı olan çalışmalarda alan indekslerinin büyüklüğünü belirlemede genel bir ölçüt bulunmamaktadır. Alan indeksleri tekniği ile yapılan çalışmalar incelendiğinde iç ölçütlerden faydalandığı belirlenmiştir. Bu ölçütler, işaretli ve işaretsiz alan ölçümlerinin ortalaması, medyan değerleri, bu değerlerin bir standart kayma veya çeyrek kayma üstü ile alan ölçümlerinin z değerlerinin, anlamlılığı biçiminde listelenebilir (Deveci, 2008; Öğretmen ve Doğan, 2004; Gondal, 2001). Bu çalışmada iç ölçüt olarak maddelerin alan indeksleri

değerlerine ait olan medyan değerlerinin 1 çeyrek kayma üstü kullanılmıştır. İşaretli ve işaretli alan indeksleri hesaplamalarına dayalı olarak maddelerin DMF içerip içirmediğine karar verebilmek için hesaplanan iç ölçütler tablo 11’de yer almaktadır.

Tablo 9: Madde Tepki Kuramına Bağlı Değişen Madde Fonksiyonu Belirleme Yöntemlerinden Alan İndeksleri Tekniğinin Yorumlanmasında Kullanılan İç Ölçütlere İlişkin İstatistikler

	İŞARETLİ ALAN İNDEKSİ			İŞARETSİZ ALAN İNDEKSİ		
	Medyan	Çeyrek Kayma	+1 Çeyrek Kayma +Medyan	Medyan	Çeyrek Kayma	+1 Çeyrek Kayma +Medyan
Aynı Kültür- Farklı Dil (Kanada)	0,283	0,256	0,539	0,594	0,658	1,251
Farklı Kültür Aynı Dil (Avustralya-İngiltere)	-0,305	-0,397	-0,702	0,501	0,486	0,987
Farklı Kültür Farklı Dil (İngiltere- Türkiye)						
1.KİTAPÇIK	1,020	1,311	2,331	2,823	2,832	5,655
5.KİTAPÇIK	0,709	0,606	1,315	1,158	1,508	2,667

Tablo 9’da görüldüğü gibi medyan değeri ölçüt olarak alındığında;

- Aynı Kültür- Farklı Dil (Kanada)örnekleminde işaretli alan indekslerine bağlı yöntem için 0,539, işaretli alan indekslerine bağlı yöntem için 1,251;
- Farklı Kültür Aynı Dil(Avustralya-İngiltere) örnekleminde işaretli alan indekslerine bağlı yöntem için -0,702, işaretli alan indekslerine bağlı yöntem için 0,987;
- Farklı Kültür Farklı Dil (İngiltere- Türkiye) birinci kitapçık örnekleminde işaretli alan indekslerine bağlı yöntem için 2,331, işaretli alan indekslerine bağlı yöntem için 5,655;
- Farklı Kültür Farklı Dil (İngiltere- Türkiye) beşinci kitapçık örnekleminde işaretli alan indekslerine bağlı yöntem için 1,315, İşaretsiz alan indekslerine bağlı yöntem için 2,667 değerleri testte DMF belirlemek için kritik değer olarak kullanılmıştır.

III.BÖLÜM

BULGULAR VE YORUM

Bu bölümde araştırmanın alt problemleri için toplanan verilerden elde edilen bulgular, alt problemlerin sırasına uygun olarak, tablo ve açıklamalarıyla birlikte verilerek yorumlanmıştır. Daha iyi anlaşılması için alt problemler tekrar verilmiştir.

3.1.ALT PROBLEM 1 İÇİN ELDE EDİLEN BULGULAR

A1. PISA 2006 yılında yapılan Uluslararası Öğrenci Değerlendirme Programında (PISA) yer alan fen başarı testi maddeleri; Mantel-Haenszel (MH), Logistik Regresyon (LR) ve Alan İndeksleri Teknikleri ile yapılan analizlerde kültür ve dile göre farklı fonksiyonlaşmakta mıdır?

Bu alt problem 3 başlıkta yürütülmüştür. 3.1.1. Aynı Kültür- Farklı Dil (Kanada); 3.1.2. Farklı Kültür Aynı Dil (Avustralya-İngiltere) ve 3.1.3.Farklı Kültür Farklı Dil (Türkiye- İngiltere) için üç teknik kullanılarak yapılan değişen madde fonksiyonu (DMF) analizlerini içermektedir. İlk alt problemin 2 alt başlığında (3.1.1., 3.1.2.) beşinci kitapçıkta yer alan 21 maddenin; 3. alt başlıkta (3.1.3.) ise birinci ve beşinci kitapçıkta yer alan toplam 40 maddenin DMF içerip içermediği belirlenmeye çalışılmıştır.

3.1.1. Aynı Kültür- Farklı Dil (Kanada Örneği) için Yapılan Değişen Madde Fonksiyonu Analizleri

Kültürün aynı, dilin farklı olduğu alt gruplar için yapılan analizler Kanada örneklemini üzerinden yürütülmüştür. Örnekleme ilişkin yöntem kısmında verilen betimsel istatistiklerden de anlaşılacağı gibi PISA sınavına Kanada'dan katılan öğrencilerin bir kısmı ingilizce form bir kısmı ise fransızca form almaktadır. Aynı kültürün içerisinde farklı dillerde test formları alan öğrencilerden elde edilen verilerden yola çıkılarak Mantel-Haenszel (MH) tekniği ile hesaplanan değişen madde fonksiyonu (DMF) bulguları tablo 10'da yer almaktadır.

Tablo 10: Aynı Kültür- Farklı Dil (Kanada Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Mantel – Haenszel Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Kitapçık-Madde	A	χ^2	P değeri	ΔMH	SE ΔMH	DMF düzeyi
5-3	1.483	7.421	0.006	-0.927	0.332	A
5-6	1.332	4.537	0.033	-0.674	0.310	A
5-11	1.830	16.179	0.000	-1.420	0.353	B
5-17	1.375	5.850	0.016	-0.749	0.302	A
5-18	0.597	9.581	0.002	1.212	0.385	B
5-19	0.570	11.175	0.001	1.320	0.380	B

Tablo 10’da belirtilen MH sonuçlarına göre; 21 maddenin altısı DMF içermektedir. Bu maddelerin üçü önemsenmeyecek düzeyde (A düzeyinde) diğer üçü de orta düzeydedir (B düzeyinde). MH analiz sonuçları incelendiğinde; 3, 6, 11 ve 17. maddelerin referans grubun (dili ingilizce olanlar) lehine, 18. ve 19. maddelerin ise odak grubun (dili fransızca olanlar) lehine DMF içerdiği görülmektedir. Aynı kültür- farklı dil bakımından beşinci kitapçık PISA fen testi maddeleri Lojistik Regresyon (LR) tekniği ile analiz edildiğinde DMF analiz bulguları tablo 13’deki gibidir.

Tablo 11: Aynı Kültür- Farklı Dil (Kanada Örneği) Bakımından PISA 2006 5. Kitapçık Alt Testinde Logistik Regresyon Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Kitapçık-Madde	Grup etkisi (p)	Grup*Toplam puan etkisi (p)	ΔR^2	DMF Türü	DMF düzeyi
5-4	0.042	0.014	0.001	ÇBDMF	A
5-19	0.215	0.021	0.009	ÇBDMF	A
5-20	0.028	0.025	0.000	ÇBDMF	A

Kanada alt gruplarının (ingilizce form ve fransızca form alanlar) beşinci kitapçıkta yer alan 21 fen maddesine ilişkin performans farklılıklarının incelenmesine yönelik yapılan lojistik regresyon analizi sonuçlarına göre üç maddede DMF olduğu görülmektedir. DMF olduğu gözlenen maddelerin tümünde grup madde etkileşimi manidar bulunmuştur. Bu; maddelerin tümünde çok biçimli DMF bulunduğunun bir

göstergesidir. DMF gösteren maddelerin DMF düzeyleri ise Gierl ve arkadaşlarının belirlediği ölçütlere göre A düzeyinde (DMF düzeyi önemli değildir) bulunmuştur.

Bir madde için iki ayrı alt gruptan elde edilen madde karakteristik eğrileri bir düzlem üzerinde aynı anda çizilirse ve madde gruplara göre farklı olarak işliyorsa, söz konusu iki eğri arasında bir alan gözlemlenir. Farklı gruplar için elde edilen eğriler arasındaki alanların hesaplanmasına odaklanan alan indeksleri tekniğinin yorumlanmasında kullanılan ölçütler yöntemler kısmında belirtilmiştir (bkz. Tablo 9) . Bu bölümde tablo 9'daki hesaplanan değerler dikkate alınarak işaretli ve işaretsiz alan indekslerine göre DMF içeren maddeler belirlenmiştir. Kritik değer olarak medyan değerinin +1 çeyrek kayma üstü kullanılmıştır. Tablo 9'a göre aynı kültür farklı dil için yapılan çalışmada kullanılan değerler işaretli alan indeksi için 0,539, işaretsiz alan indekslerine bağlı yöntem için 1,251'dir. Tablo 12'de alan indeksleri tekniği ile alanlar arasındaki farklar hesaplanarak DMF bulunan maddelere ilişkin analiz sonuçları yer almaktadır. Tabloda DMF bulunduran maddeler koyu karakterlerle belirtilmiştir.

Tablo 12: Aynı Kültür- Farklı Dil (Kanada Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Alan İndeksleri Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Kitapçık-	İşaretli alan indeksleri	İşaretsiz alan indeksleri	DMF
5-2	3,30	3,69	ÇBDMF
5-3	0,63	0,63	TBDMF
5-4	0,45	1,68	ÇBDMF
5-6	0,28	2,12	ÇBDMF
5-8	0,54	0,60	ÇBDMF
5-16	0,55	0,56	TBDMF
5-17	1,05	1,67	ÇBDMF
5-20	0,20	2,92	ÇBDMF

Tablo 12 incelendiğinde işaretli alan indeksleri için, 2, 3, 8, 16 ve 17 nolu maddelerin; işaretsiz alan indeksleri için 2, 4, 6, 17 ve 20 nolu maddelerin yanlılık gösterdiği anlaşılmaktadır. İşaretli alan indeksleri yöntemleri ile elde edilen sonuçlara göre; beşinci kitapçıkta yeralan 21 maddeden beşi DMF içermektedir. Hesaplanan işaretsiz alan indekslerine göre beş maddede DMF bulunmuştur. İşaretli alan indekslerinde DMF belirlemek için kullanılan medyan değerine göre DMF gösterdiği belirlenen tüm

maddelerin işareti pozitifdir. Pozitif işaretli olan maddeler referans grubun (dili İngilizce olanlar) lehine DMF içermektedir. Raju (1990) ve Gondal (2001) çalışmalarında işaretsiz alan indeksinin, işaretli alan indeksinden oldukça büyük olduğu durumlarda, DMF ortaya çıktığını ve işaretsiz alan indekslerinin maddelerde DMF'nin hangi grup lehine olduğu hakkında bilgi vermediklerini belirtmişlerdir. Çok Biçimli Değişen Madde Fonksiyonunu (ÇBDMF) belirlemede, işaretsiz alan indekslerinin etkisinin daha büyük olduğunu bildirmişlerdir. DMF'nin araştırıldığı grupların Madde Karakteristik Eğrileri (MKE) çakışır ise bu maddede DMF yoktur. Eğer grupların MKE'leri farklılık gösterirse, madde DMF içerir. Bu maddelerden Gondol ve Raju'nun açıklamaları doğrultusunda alan indekslerinin birbirlerine göre durumları incelendiğinde 3. ve 16. madde dışındaki tüm maddelerde işaretli alan indeksi değerinin işaretsiz alan indeksi değerinden küçük olduğu; dolayısıyla bu maddelerin ÇBDMF'ye sahip olduğu sonucuna ulaşılmıştır.

Kanada'da farklı dil formlarına yönelik benzer bir çalışma Ercikan ve arkadaşları (2002) tarafından yapılmıştır. Araştırmacı uluslararası eğitim değerlendirme programı İngilizce ve Fransızca form verilerini kullanmıştır. Bu iki grup aynı ülke ve eğitim sisteminin parçası iken 70 maddeden İngiltere lehine DMF içeren 10 madde (8'i önemli düzey), Fransa lehine DMF içeren 8 madde (4'ü önemli düzey) olduğu belirlenmiştir. Yapılan yanlılık çalışmalarında ise DMF içeren maddelerdeki olası yanlılığın dile bağlı olduğu sonucuna ulaşılmıştır. Çalışma sonuçları yapılan bu çalışma ile uyum göstermiş benzer biçimde tekniklerle elde edilen bulgularda İngilizce form lehine sayıca daha çok DMF bulunduran maddenin varlığı belirlenmiştir.

3.1.2. Farklı Kültür Aynı Dil (Avustralya-İngiltere) İçin Yapılan Değişen Madde Fonksiyonu Analizleri

Kültürün farklı dilin aynı olduğu (Avustralya-İngiltere) alt gruplar için beşinci kitapçık temel alınarak MH tekniği ile elde edilen DMF bulguları tablo 15'te yer almaktadır.

Tablo 13: Farklı Kültür Aynı Dil (Avustralya-İngiltere Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Mantel – Haenszel Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Kitapçık-Madde	α	χ^2	P değeri	ΔMH	SE ΔMH	DMF düzeyi
5-1	0.733	5.962	0.015	0.729	0.292	A
5-15	1.331	6.474	0.011	-0.672	0.259	A
5-16	1.308	7.545	0.006	-0.630	0.226	A
5-18	0.491	30.629	0.000	1.671	0.302	C
5-19	0.645	13.414	0.000	1.032	0.278	B
5-20	1.467	15.068	0.000	-0.901	0.229	A

Tablo 13'deki MH sonuçlarına göre; beşinci kitapçıkta yeralan 21 maddeden 6 madde DMF içermektedir. Bu maddelerin dördünün A düzeyinde (DMF önemsenmeyecek düzeydedir.), biri B (DMF orta düzeyde), diğeri de C düzeyindedir (DMF önemli düzeydedir). MH analiz sonuçları incelendiğinde; 1., 18. ve 19. maddelerin odak grubun (İngiltere) lehine DMF içerdiği, 15., 16. ve 20. maddelerin ise referans grubun (Avustralya) lehine DMF içerdiği görülmektedir.

Farklı Kültür aynı dil bakımından alt gruplar için (Avustralya-İngiltere) LR tekniği ile elde edilen DMF bulguları tablo 16'da yer almaktadır.

Tablo 14: Farklı Kültür Aynı Dil (Avustralya-İngiltere Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Lojistik Regresyon tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Kitapçık- Madde	Grup etkisi (p)	Grup*Toplam puan etkisi (p)	ΔR^2	DMF Türü	DMF düzeyi(Gierl ve Arkadaşlarının Ölçütleri)
5-1	0.018	0.141	0.005	TBDMF	A
5-3	0.008	0.033	0.001	ÇBDMF	A
5-16	0.004	0.022	0.004	ÇBDMF	A
5-18	0.003	0.282	0.024	TBDMF	A

Tablo 14'de bulunan MH analiz sonuçlarına göre beşinci kitapçıkta yeralan 21 maddeden 4 maddenin farklı kültür aynı dil için DMF içerdiği sonucuna ulaşılmıştır. Tablo 14'de yer alan LR analizi sonuçlarına göre 1. ve 18. madde tek biçimli değişen

madde fonksiyonu (TBDMF); 3. ve 16. maddenin ise çok biçimli değişen madde fonksiyonu (ÇBDMF) içermektedir. DMF gösteren dört maddeninde DMF düzeyleri Gierl ve arkadaşlarının belirlediği ölçütler çerçevesinde önemli bulunmamıştır.

Farklı Kültür aynı dil bakımından alt gruplar için (Avustralya-İngiltere) AİT ile elde edilen DMF bulguları tablo 15’de yer almaktadır.

Tablo 15: Farklı Kültür Aynı Dil (Avustralya-İngiltere Örneği) Bakımından PISA 2006 Beşinci Kitapçık Alt Testinde Alan İndeksleri Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Madde	İşaretli alan indeksleri	İşaretsiz alan indeksleri	DMF
5 -2	-4,24	5,64	ÇBDMF
5-16	-0,37	3,36	ÇBDMF
5-20	-0,80	0,80	TBDMF

Alan indeksleri tekniğinin yorumlanmasında kullanılan iç ölçütlere ilişkin istatistikler incelendiğinde farklı kültür aynı dil örnekleminde işaretli alan indekslerine bağlı yöntem için 0,702, işaretsiz alan indekslerine bağlı yöntem için 0,987 DMF belirlemek için kritik değerdir. Bu değerler temel alındığında tablo 15’de yer alan işaretli alan indeksleri yöntemleri sonuçlarına göre; beşinci kitapçıkta yeralan 21 maddeden iki madde DMF içermektedir. Bu maddelerin her ikisinin de işareti negatiftir. Negatif işaretli olan bu maddeler odak grubun (İngiltere) lehine DMF içermektedir.

İşaretsiz alan indeksleri için hesaplanan 0,987 değeri ölçüt olarak alındığında, 2 ve 16. maddede DMF olduğu sonucuna ulaşılmaktadır. Tabloda yer alan değerler incelendiğinde bu iki maddenin sıra dışı alan indeksleri değerlerine sahip olduğu belirlenmiştir. Bu nedenle, bu maddeler ÇBDMF göstermektedir. Söz konusu olan iki madde için elde edilen madde karakteristik eğrileri kesişmektedir. Tablo 15’de 20. maddenin ise işaretli ve işaretsiz alan indeks değerinin birbirine eşit olduğu görülmektedir. Bu nedenle 20. madde TBDMF göstermektedir. Başka bir ifade ile bu maddenin eğrinin her noktasında İngiltere lehine sonuçlara sahip olduğu söylenebilir.

Abedi 2004 ve 2008 yıllarında ELLs (English Language Learners) sınavında aynı dil formunu alan farklı alt yapılara sahip öğrenciler üzerinde yaptığı çalışmada karmaşık -

zor cümle yapılarında farklı alt gruplardan gelen bireylerin performanslarının karşılaştırılmasında sıkıntılar olabileceğini vurgulamıştır. Bu sonuç; kültürlerin farklılaştığı ancak alınan form dilinin değişmediği durumlarda bu çalışmada farklı tekniklerle belirlenen DMF’li maddeler için bir dayanak olabilir.

3.1.3. Farklı Kültür- Farklı Dil (Türkiye-İngiltere) için Yapılan Değişen Madde Fonksiyonu Analizleri

Türkiye-İngiltere alt grupları için yapılan DMF analizleri birinci ve beşinci olmak üzere iki kitapçık türü için birden yürütülmüştür. Temel amaç; yöntem kısmında da belirtildiği gibi açıklanan maddelerin tamamına ilişkin DMF bulguları elde edebilmektir. Türkiye-İngiltere alt grupları için üç teknikte hem DMF analizleri çalışılmış; hem de alt problem 3’te Türkçe ve İngilizce form setlerinde açıklanan maddelerdeki olası yanlışlık nedenleri araştırılmıştır.

Çalışmanın bu bölümünde farklı kültür farklı dil bakımından alt gruplar için (Türkiye-İngiltere) MH tekniği ile elde edilen DMF bulguları tablo 16’da yer almaktadır.

Tablo 16: Farklı Kültür Farklı Dil (Türkiye-İngiltere Örneği) Bakımından PISA 2006 Birinci Kitapçık ve Beşinci Kitapçık Alt Testinde Mantel – Haenszel Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Kitapçık-Madde	α	χ^2	P değeri	ΔMH	SE ΔMH	DMF düzeyi
1-1	1.504	6.378	0.012	-0.959	0.371	A
1-3	1.384	4.406	0.036	-0.702	0.330	A
1-4	0.578	11.819	0.001	1.289	0.371	B
1-5	3.118	51.650	0.000	-2.673	0.380	C
1-8	2.984	48.883	0.000	-2.569	0.380	C
1-10	1.440	6.151	0.013	-0.857	0.366	A
1-13	1.643	9.110	0.003	-1.167	0.375	A
1-14	3.289	57.521	0.000	-2.798	0.381	C
1-15	0.679	6.349	0.012	0.911	0.350	A
1-18	1.756	15.712	0.562	-1.323	0.327	B
Kitapçık-Madde	α	χ^2	P değeri	ΔMH	SE ΔMH	DMF düzeyi
5-1	1.393	4.274	0.039	-0.778	0.364	A
5-2	3.589	52.415	0.000	-3.003	0.420	C
5-3	1.459	6.482	0.011	-0.888	0.341	A
5-4	1.863	16.191	0.000	-1.463	0.362	B
5-6	0.637	8.530	0.003	1.062	0.353	B
5-9	0.397	35.051	0.000	2.169	0.365	C
5-12	0.490	25.284	0.000	1.675	0.337	C
5-16	0.520	20.342	0.000	1.537	0.340	B
5-17	0.491	24.751	0.000	1.671	0.333	C
5-18	2.487	35.744	0.000	-2.141	0.362	C
5-19	1.537	8.025	0.005	-1.011	0.351	B

Tablo 16 incelendiğinde birinci kitapçıktaki MH sonuçlarına göre 19 maddeden 10'u DMF içermektedir. Maddelerin beşi A düzeyinde, ikisi B, üçü ise C düzeyinde DMF içermektedir. MH analiz sonuçları incelendiğinde; 1., 3., 5., 8., 10., 13., 14. ve 18. maddelerin referans grubun (İngiltere) lehine DMF içerdiği; 4. ve 15. maddelerin ise odak grubun (Türkiye) lehine DMF içerdiği görülmektedir. Beşinci kitapçıkta yer alan 21 maddeden 11 maddenin farklı kültür ve dil için DMF içerdiği bulunmuştur. Bu maddelerin ikisi A, dördü B, beşi de C düzeyinde DMF içermektedir. MH analiz sonuçları incelendiğinde; 1., 2., 3., 4., 18. ve 19. maddelerin referans grubun

(İngiltere) lehine DMF içerdiği; 6., 9., 12., 16. ve 17. maddelerin ise odak grubun (Türkiye) lehine DMF içerdiği görülmektedir.

Zieky (2003) ön deneme uygulaması mümkün olmayan başarı ölçümlerinde “B” düzeyinde bulunan DMF’nin ihmal edilebileceğini belirtmiştir. Bu çerçeveden bakıldığında DMF içeren ve yanlılığın araştırılabileceği madde sayısı MH tekniğine göre sekize inmektedir.

Farklı kültür farklı dil bakımından Türkiye ve İngiltere için LR tekniği ile elde edilen DMF bulguları aşağıdaki tablo 17’de yer almaktadır.

Tablo 17: Farklı Kültür Farklı Dil (Türkiye-İngiltere Örneği) Bakımından PISA 2006 Birinci Kitapçık ve Beşinci Kitapçık Alt Testinde Lojistik Regresyon tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

Kitapçık-Madde	Grup etkisi (p)	Grup*Toplam puan etkisi (p)	ΔR^2	DMF Türü	DMF düzeyi(Gierl ve Arkadaşlarının Ölçütleri)
1-3	0.005	0.0003	0.000	ÇBDMF	A
1-4	0.0000	0.00008	0.032	ÇBDMF	A
1-5	0.861	0.031	0.030	ÇBDMF	A
1-8	0.812	0.006	0.041	ÇBDMF	B
1-15	0.101	0.006	0.022	ÇBDMF	A
1-17	0.003	0.0077	0.007	ÇBDMF	A
Kitapçık-Madde	Grup etkisi (p)	Grup*Toplam puan etkisi (p)	ΔR^2	DMF Türü	DMF düzeyi(Gierl ve Arkadaşlarının Ölçütleri)
5-2	0.004	0.767	0.069	TBDMF	B
5-4	0.210	0.011	0.014	ÇBDMF	A
5-11	0.091	0.039	0.001	ÇBDMF	B
5-12	0.000005	0.008	0.029	ÇBDMF	A
5-14	0.002	0.0003	0.001	ÇBDMF	A
5-16	0.001	0.169	0.022	TBDMF	A

Tablo 17’de yeralan birinci kitapçığa ilişkin Lojistik Regresyon tekniği sonuçlarına göre altı madde ÇBDMF içermektedir. Bu altı maddenin düzeyleri Gierl ve arkadaşlarının ölçütleri çevresinde değerlendirildiğinde; beşinin (3, 4, 5, 15, 17) önemsenmeyecek düzeyde; bir maddenin (8. maddenin) ise orta düzeyde DMF içerdiği

söylenebilir. Tablo 17'deki farklı kültür farklı dil (Türkiye-İngiltere örneği) bakımından PISA 2006 beşinci kitapçık alt testinde lojistik regresyon tekniği ile yapılan analiz sonuçlarına göre altı madde (2, 4, 11, 12, 14 ve 16) DMF içermektedir. Bunlardan 2. ve 16. maddeler TBDMF gösterirken diğer dördü ÇBDMF göstermektedir. Altı maddenin dördünün önemsenmeyecek düzeyde, ikisinin ise orta düzeyde DMF içerdiği saptanmıştır. Bir önceki tablo açıklamasında belirtildiği gibi Ziekly'nin (2003) çalışması baz alındığında LR tekniğine göre DMF'li maddelerin hepsinin ihmal edilebilir düzeyde olduğu söylenebilir.

Farklı kültür farklı dil bakımından alt gruplar için (Türkiye-İngiltere) AİT ile elde edilen DMF bulguları aşağıdaki tablo 18'de yer almaktadır.

Tablo18: Farklı Kültür Farklı Dil (Türkiye-İngiltere Örneği) Bakımından PISA 2006 Birinci ve Beşinci Kitapçık Alt Testinde Alan İndeksleri Tekniği ile Değişen Madde Fonksiyonu Bulunan Maddelere İlişkin Analiz Sonuçları

KİTAPÇIK-MADDE NO	İşaretli alan indeksleri	İşaretsiz alan indeksleri	DMF BİÇİMİ
1-1	-24,00	30,95	ÇBDMF
1-2	34,84	39,89	ÇBDMF
1-3	2,68	3,00	ÇBDMF
1-17	34,69	39,51	ÇBDMF
1-18	35,92	40,59	ÇBDMF
1-19	35,97	40,98	ÇBDMF
KİTAPÇIK-MADDE NO	İşaretli alan indeksleri	İşaretsiz alan indeksleri	DMF BİÇİMİ
5-2	2,82	2,82	TBDMF
5-4	2,02	2,74	ÇBDMF
5-12	0,10	3,57	ÇBDMF
5-14	3,05	8,68	ÇBDMF
5-16	0,22	3,17	ÇBDMF
5-18	1,46	1,52	ÇBDMF

Tablo 18'deki işaretli alan indeksleri yöntemleri sonuçlarına göre; birinci kitapçıkta yeralan 19 maddeden altı madde DMF içermektedir. İşaretli alan indeksleri yorumlanırken kullanılan iç ölçütlere ilişkin değerler temel alındığında madde DMF gösteren tüm maddelerin işareti pozitifdir. İşareti pozitif olan bu maddeler referans grubun (İngiltere) lehine DMF içermektedir. Birinci kitapçık için işaretsiz alan

indeksleri iç ölçütü 5,655 alındığında (bkz tablo 9) 5 maddenin DMF içerdiği sonucuna ulaşılmaktadır. Bu maddelerin hepsi işaretli alan indekslerine göre de yanlış bulunan maddelerdir. İşaretsiz alan indeksleri büyüklükleri 3'ün üstünde ise sıra dışı değerler olarak kabul edilmekte ve bu değerler yanlışlığın en güçlü kanıtı olarak kabul edilmektedir (Raju, 1990). Bu açıklama doğrultusunda 3. madde dışındaki tüm maddelerin bu tanımlamaya oldukça uyduğu ve DMF gösteren tüm maddelerin ÇBDMF gösterdiği belirlenmiştir.

Beşinci kitapçığa ait olan işaretli alan indeksleri değerleri incelendiğinde 21 maddeden dördünde DMF olduğu sonucuna ulaşılmıştır. Tablo 18'de koyu karakterlerle belirtilen bu dört maddenin hepsi İngiltere lehine DMF göstermektedir. İşaretsiz alan indeksleri değerlerine göre beş maddede DMF bulunmaktadır. İşaretli alan indeksleri ve işaretsiz alan indeksleri değerlerinin birbirlerine göre durumu incelendiğinde beşinci kitapçıkta yeralan ve DMF gösteren maddelerden 2. madde dışındaki tüm maddelerin ÇBDMF gösterdiği sonucuna ulaşılmıştır. ÇBDMF gösteren bu maddelerin odak ve referans gruba sağladığı avantaj değişik olmaktadır. Yani bu maddelerin çoğunun araştırılan gruplar için farklı madde karakteristik eğrilerine sahip olduğu ortaya çıkmaktadır.

Elde edilen bulgular incelendiğinde; karşılaştırılan alt gruplar arasındaki farklılıklar arttıkça DMF gösteren madde sayıları ve düzeylerinin arttığı sonucuna ulaşılmıştır. Bu sonuç, Atalay (2010)'ın ve Asil (2010)'in yapmış oldukları çalışmalarda dilsel ve kültürel farklılıklar arttıkça DMF gösteren maddelerin arttığı sonucu ile tutarlılık göstermektedir.

3.2. ALT PROBLEM İKİ İÇİN ELDE EDİLEN BULGULAR

A2. Kültür ve dile göre Mantel-Haenszel (MH), Logistik Regresyon(LR) ve Alan İndeksleri Teknikleri ile elde edilen Değişen Madde Fonksiyonu düzeyleri farklılaşmakta mıdır?

PISA birinci ve beşinci kitapçık fen testindeki maddelerin, farklı yanıtlayıcı grupları için Mantel Haenzel, Lojistik Regresyon, İşaretli Alan İndeksleri ve işaretsiz Alan İndeksleri tekniklerine göre DMF'ye yol açan maddeler tablo 21'de belirtilmiştir. Tabloda belirtilen maddelerin DMF düzeyleri dikkate alınmamıştır. Aşağıda MH ve

LR teknikleri ile ulaşılan tüm düzeylerdeki (A düzeyi, B düzey ve C düzeyi) DMF'li maddelere yer verilmiştir.

Tablo 19: Değişen Madde Fonksiyonu İçeren Maddelerin Tekniklere Göre Dağılımı

	MH tekniği	LR Tekniği	İşaretli alan indeksleri	İşaretsiz alan indeksleri
Aynı Kültür Farklı Dil (Kanada)	3, 6, 11, 17, 18, 19	4, 19, 20	2, 3, 8, 16, 17	2, 4, 6, 17, 20
Farklı Kültür Aynı Dil (İngiltere-Avustralya)	1, 15, 18, 19, 20	1, 3, 16, 18	2, 20	2, 16
Farklı Kültür Farklı Dil (Türkiye – İngiltere) (1. Kitapçık)	1, 3, 4, 5, 8, 10, 13, 14, 15, 18	3, 4, 5, 8, 15, 17	1, 2, 3, 17, 18, 19	1, 2, 17, 18, 19
Farklı Kültür Farklı Dil (Türkiye – İngiltere) (5. Kitapçık)	1, 2, 3, 4, 6, 9, 12, 16, 17, 18, 19	2, 4, 11, 12, 14, 16	2, 4, 14, 18	2, 4, 12, 14, 16

Tablo 19 incelendiğinde, dört tekniğe göre belirlenen DMF veren maddelerin ve sayılarının değiştiği görülmektedir. Bu çalışmada DMF belirlemede ele alınan odak ve referans grup için farklı tekniklerle elde edilen DMF'li maddelerin sayıları bakımından farklılık gösterdikleri tabloya bakılarak çıkarılabilir. Alt problem 1'de MH ve LR için belirlenen düzeylerinde farklılaştığı görülmektedir. Örneğin; aynı kültür farklı dil (Kanada) gruplandırılması için MH tekniği ile altı maddede, LR tekniği ile üç maddede DMF bulunmuştur. Sadece 19. madde her iki teknikte de DMF gösteren madde olarak tespit edilmiştir. Bu maddenin DMF düzeyi ise MH'e göre B iken; LR'a göre A'dır. MH tekniği ile elde edilen DMF analizleri ile işaretli alan indeksine bağlı teknikte elde edilen DMF gösteren maddeler incelendiğinde; madde numarası 3 ve 17 olan maddeler uyuma göstermemektedir. Aynı teknikle işaretsiz alan indekslerine göre DMF gösteren maddeler karşılaştırıldığında madde 17 her iki teknikte de ortaktır. Farklı kültür aynı dil (İngiltere-Avustralya) gruplamasında MH tekniğine göre beş; LR tekniğine göre dört; işaretli alan indeksleri tekniğine göre beş ve İşaretsiz alan indeksleri beş maddede DMF olduğu belirlenmiştir. Farklı kültür farklı dil (Türkiye – İngiltere) için ise MH tekniğine göre 11; LR tekniğine göre altı; İşaretli alan indeksleri ve işaretsiz alan indeksleri tekniklerine göre ikişer maddede DMF olduğu

belirlenmiştir. Tablo 19 her bir teknik için yukarıdan aşağıya doğru incelendiğinde grupların dil ve kültüre bağlı olan farklılıkları arttıkça yöntemlerle elde edilen DMF’li madde sayısı da artmaktadır. Özellikle uluslararası çalışmalarda yürütülen DMF çalışmalarında yöntemlere göre farklılıkların artmasının temel nedeni; üzerinde çalışılan grupların daha heterojen olmasına bağlanmaktadır (Welkenhuysen-Gybels ve Billiet; 2002). Elde edilen sonuçlar bu bilgi ile örtüşmektedir.

Tekniklerle belirlenen DMF’li madde sayılarının değiştiği görülmektedir. Her ne kadar KTK ve MTK yöntemleri kendi araların da uyum gösterse de DMF’li madde sayısı bakımından bir uyum söz konusu değildir. Bu çalışmada karşılaşılan bu bulgu birçok DMF çalışması ile uyum göstermektedir (Gök,2010; Yurdugül ve Aşkar, 2004; Ateşok, 2008; Doğan ve Öğretmen, 2006 ve 2008)

Sayısal anlamda büyüklükler incelendiğinde MTK’na bağlı alan indeksleri yönteminde MH ve LR yöntemlerine göre daha çok DMF gösteren madde olduğu görülmektedir. Bu bulgu Ateşok(2008), Gondal (2001), Doğan ve Öğretmen (2006) yaptığı araştırmaların sonuçları ile uyum göstermektedir.

Yapılan DMF analizleri gibi elde edilen değerler arasındaki uyum da 3 grup için ayrı ayrı incelenmiştir. Bunun için LR için etki büyüklüğü değerleri; MH için hesaplanan ki-kare değerleri, işaretli alan indeksleri ve işaretsiz alan indeksleri ile elde edilen değerler arasındaki sıra farkları korelasyonu incelenmiştir. Her bir grup için tablo 20’de yer alan değerler elde edilmiştir.

Tablo 20: Mantel – Haenszel, Logistik Regresyon ve Alan İndeksleri Tekniği İle Elde Edilen Değişen Madde Fonksiyonu Değerleri Arasındaki Sıra Farkları Korelasyonları

		MH	LR	İAİ	İSAİ
AYNI KÜLTÜR FARKLI DİL (Kanada)	MH	1,00			
	LR	0,681**	1,00		
	İAİ	0,11	-0,10	1,00	
	İSAİ	0,19	0,02	0,642**	1,00
FARKLI KÜLTÜR AYNI DİL (İngiltere-Avustralya)	MH	1,00			
	LR	0,538*	1,00		
	İAİ	-0,02	0,35	1,00	
	İSAİ	0,39	-0,11	-0,65	1,00
FARKLI KÜLTÜR FARKLI DİL (Türkiye – İngiltere) (1. KİTAPÇIK)	MH	1,00			
	LR	0,45	1,00		
	İAİ	-0,33	-0,21	1,00	
	İSAİ	0,01	0,05	0,491*	1,00
FARKLI KÜLTÜR FARKLI DİL (Türkiye – İngiltere) (5. KİTAPÇIK)	MH	1,00			
	LR	0,891**	1,00		
	İAİ	-0,05	-0,07	1,00	
	İSAİ	0,27	0,11	0,43	1,00

** Korelasyon 0.01 düzeyinde anlamlı

* Korelasyon 0.05 düzeyinde anlamlı

İAİ: İşaretli alan indeksleri

İSAİ: İşaretsiz alan indeksleri

MH; LR ve Alan indeksleri ile elde edilen DMF değerleri arasındaki sıra farkları korelasyonları tablo 20’de görülmektedir. Tablo 20’de $\alpha = 0,01$ ve $0,05$ düzeyinde manidar olan korelasyon değerleri koyu olarak gösterilmiştir. Tablo 22’ye göre aynı kültür farklı dil için MH ile LR ve İşaretli alan indeksleri ile İşaretsiz alan indeksleri değerleri arasındaki sıra farkları korelasyon katsayıları $0,01$ düzeyinde; farklı kültür aynı dil için MH ile LR arasındaki sıra farkları korelasyon katsayısı $0,05$ düzeyinde farklı kültür farklı dil (1. kitapçık) için işaretli alan indeksleri ile işaretsiz alan indeksleri arasındaki sıra farkları korelasyon katsayısı $0,05$ düzeyinde ve son olarak

farklı kültür farklı dil (5. kitapçık) için MH ile LR arasındaki sıra farkları korelasyon katsayısı 0,01 düzeyinde anlamlı bulunmuştur.

Elde edilen bulgulardan da anlaşılacağı üzere teknikler birbirlerinden çoğu kez farklılaşmakta ve elde edilen değerler birbiriyle uyumlu gözükmemektedir. Bu sonuç DMF analizlerinde karşılaşılan en temel problemlerden biri olarak karşımıza çıkmaktadır (Karami ve Nodoushan; 2011)

Sonuç KTK'na bağlı tekniklerin kendi arasında; MTK'ya bağlı tekniklerin de kendi aralarında uyumlu olduğu yönündedir. Bu tekniklerin farklılaştıkları noktaların bu farklılığa neden olduğu düşünülmektedir. Bir bakıma KTK yöntemlerinden LR ve MH teknikleri benzer bileşenlere sahiptir, her ikisinde de toplam test puanları üzerinden işlem yapılmasına rağmen kategorilendirmelerde kullanılan ölçütlerin bu uyumu bozduğu düşünülmektedir. Bu çalışmada DMF'li maddelere karar vermede daha etkili bir yöntem olduğu savunulan (Camilli, 2006) etki büyüklüğü üzerinden sonuçlara ulaşılmıştır. Yöntemler kısmında MTK'ya bağlı ölçütler belirlenirken vurgulandığı gibi KTK'ya bağlı yöntemlerde de etki büyüklüğü ölçütü değişebilmektedir. Nitekim Lopez-Pina (2004) yaptıkları çalışma sonucunda, ΔR^2 istatistiğinin yorumlanması için yeni ölçütlerin belirlenmesi gerektiğini vurgulamışlardır. Kalaycıoğlu ve Kelecioğlu (2011) yapmış oldukları çalışmada ΔMH için 0,010 ve 0,020 değerlerini kullanmışlardır. Öte yandan MTK'na dayalı yöntemlerin kullanılmasında sağlanması gereken varsayımların çeşitliliği ve bu varsayımlardan hareketle elde edilen madde parametre kestirimlerinin; MTK ile KTK'na göre DMF belirleme yöntemleri arasındaki uyumsuzluğun veya farklılığın nedeni olabileceği düşünülmektedir. MTK'da elde edilen parametrelerin doğruluğu varsayımların sağlanmasında gösterilen duyarlılığa göre değişecektir. Alan indeksleri tekniğinde yanlış olan maddelerin belirlenmesinde kullanılan iç ölçütlerin sağlam bir dayanağının olmaması ve bu ölçütlerden hareketle elde edilen DMF gösteren maddelerin farklı ölçütler kullanıldığında değişebileceği unutulmamalıdır.

Yapılan bu analizin yanı sıra tek-çift, doğru-yanlış, evet-hayır sonuçlu denemeler için kullanılan “üç alt örnekleme incelenen soruların DMF sonuçları (DMF'li - DMF'siz) tekniklere göre farklılık göstermekte midir?” sorusunun cevabı Cochran q testi ile de

araştırılmıştır. Bu test ile üç tekniğin bir arada uyumlu sonuçlar verip vermediği belirlenmeye çalışılmıştır.

Aynı kültür farklı dil için hesaplanan Cochran q analizi sonucunda $Q=23,5 > \chi^2_{(3,.05)}=7,82$; farklı kültür aynı dil için hesaplanan Cochran q $Q=28 > \chi^2_{(3,.05)}=7,82$; farklı kültür farklı dil birinci kitapçık için hesaplanan Cochran q analizi sonucunda $Q=15 > \chi^2_{(3,.05)}=7,82$; beşinci kitapçık için ise hesaplanan $Q=36 > \chi^2_{(3,.05)}=7,82$ olduğundan tekniklerden elde edilen sonuçların birbirinden farklı olduğu sonucuna ulaşılmıştır.

Alt örneklemelerin farklılaşmasının yanı sıra; ortaya çıkan farklılıkların en temel sebeplerinden birinin; yöntemlerin DMF belirlemede izledikleri istatistiksel yolun ve aşamaların farklılaşması olduğu düşünülmektedir. Yöntemlerle elde edilen sayılar DMF kararı için kullanılmakta; ancak elde edilen bulgulara farklı matematiksel yollarla ulaşılmaktadır. Yöntemlerde kullanılan eşitleme kriterleri, algoritmaları ve kesme puanları farklılıklar göstermektedir. Bu farklılıklarla elde edilen sonuçların da uyumsuzluklar taşıyor olması beklenen bir durumdur.

Farklı tekniklere yönelik yapılan incelemelerde; DMF belirleme yöntemi seçilirken arasında ortaya çıkan uyumsuzluk “hangi yöntemin kullanımı daha uygun” sorusunu akla getirmektedir. Bu soruya, sonuçların uygulanmasının ve seçilen yöntemin pratik olmasının dikkate alınması gerekliliği vurgulanmaktadır. KTK’na dayalı olan yöntemlerin MTK’na göre teorik olarak daha basit, uygulama açısından daha pratik ve kolay yöntemler olduğu belirtilmiştir (Cromwell, 2002). Teknikler bir arada düşünüldüğünde birbirlerine göre avantajlı ve dezavantajlı yönleri vardır. Bu çalışmadaki tekniklerin birbirlerine göre tercih edilirliliği kıyaslanacak olunursa; 100-300 kişilik küçük örneklemelerde tek biçimli değişen madde fonksiyonunu belirlemek amacıyla; yetenek düzeylerinin gruplandırılması gibi ek bir işlemin gerekmediği durumlarda MH tekniği; daha büyük örneklemelerde tek biçimli değişen madde fonksiyonunun yanı sıra çok biçimli değişen madde fonksiyonunu da belirleyebilmek için LR tekniği daha doğru ve güvenilir sonuçlar verebilir. Ancak her iki tekniğinin de I. tip hataya sebep olduğu bilinmektedir. MTK ile yapılan DMF analizlerinden elde edilen sonuçların güvenilirliği ise kuramın sayıltılarının karşılanması ve model - veri uyum iyiliğinin sağlanması koşulundan etkilenmektedir. Model veri uyumu zayıf

olduğu durumlarda odak ve referans gruplardaki madde parametrelerinin kestirimlerinde olağan dışı DMF değerleri elde edilebilir. Bu durum DMF'nin yorumlanması ve gruplar arası karşılaştırılması için yanlışlıklara sebep olacaktır (Hambelton ve Rogers 1989).

3.3. ALT PROBLEM 3 İÇİN ELDE EDİLEN BULGULAR

A3. Uzman kanılarına göre maddelerin kültürler ve diller arasında DMF göstermesinin olası nedenleri nelerdir?

DMF içeren araştırmalarda DMF içeren maddelerden hangilerinin yanlışlık gösterdiğini ve niçin böyle olduğunu belirlemenin güç olduğu; bu nedenle maddelerin yanlış olup olmadığına karar vermek için nitel bir araştırma yapılan analizlerin desteklenmesinin gerektiği giriş ve yöntemler kısmında belirtilmişti. Bu amaçla maddelerin, madde formatları, çeviri, programlar, kültürel yapı gibi durumlar için avantaj sağlayıp sağlamayacağı EK 1 ve EK 2 yer alan sorular aracılığı ile derinlemesine araştırılmıştır. PISA fen başarı testindeki açıklanan maddeler ile iki form hazırlanmıştır. Birinci form (EK 1) ile açıklanan maddelerin Türkçe ve İngilizce formunun yanlışlık içerip içermediği uzman görüşlerine başvurularak değerlendirilmiştir. Bu formu yanıtlayan uzmanların fen bilimleri alanından mezun olmaları (fen bilimleri alanında öğretmenlik yapmaları) ve iyi derecede İngilizce bilmeleri dikkate alınmıştır. İkinci form (EK 2) ile ise açıklanan maddelerin Türkçe formunun yanlışlık içerip içermediği uzman görüşlerine başvurularak değerlendirilmiştir. Bu formu yanıtlayan uzmanların fen bilimleri alanından mezun olmaları ve fen bilimleri alanında öğretmenlik yapmaları yeterli görülmüştür.

3.3.1. Birinci Forma Ait Bulgular

Çalışmanın bu bölümünde 15 uzmana ulaşılmış ve bu uzmanların tamamıyla birebir görüşülmüştür. Forma ait bulguların yanı sıra PISA 2006 raporu incelenerek maddelere ilişkin bilgilere de bu bölümde yer verilmiştir.

Çalışmaya katılan uzmanların kültür ve dil bakımından yanlışlığa ilişkin görüşlerini esas alan sonuçlar tablo 21'de mevcuttur.

Tablo 21: Uzmanların PISA Fen Testi Maddelerinin Kültürler ve Diller Arasında Değişen Madde Fonksiyonu Göstermesinin Olası Nedenlerine İlişkin Görüş Dağılımları (Form 1)

1.KİTAPÇIK	MADDE 1 (S213Q02)		MADDE 8 (S426Q03)		MADDE 9 (S426Q05)		MADDE 16 (S485Q03)		MADDE 17 (S508Q03)	
	Yanılı		Yanılı		Yanılı		Yanılı		Yanılı	
	f	%	f	%	f	%	f	%	f	%
İNGİLİZCE FORM LEHİNE	10	67	10	67	9	60	5	33	6	40
TÜRKÇE FORM LEHİNE									2	13

5.KİTAPÇIK	MADDE 1 (S213Q02)		MADE 12 (S447Q02)		MADDE 13 (S447Q03)		MADDE 14 (S447Q04)		MADDE 18 (S477Q02)		MADDE 19 (S477Q03)	
	Yanlı		Yanlı		Yanlı		Yanlı		Yanlı		Yanlı	
	f	%	f	%	f	%	f	%	%	f	f	%
İNGİLİZCE FORM	1	67	9	60	7	47	6	40	2	13	2	13
TÜRKÇE FORM LEHİNE							1	1	2	13	1	7

Uzmanların vermiş oldukları cevaplar incelendiğinde uzmanların çoğunun, maddelerin İngilizce formunu alan öğrencilerin lehine çalıştığı beklentisinde oldukları gözlenmektedir.

Uzmanların % 67'si birinci ve beşinci kitapçıkta yer alan 1 nolu maddenin; laboratuvar koşullarının Türkiye'de yeterince karşılanmadığı, ilköğretim düzeyindeki okullarda ayrı ayrı fizik, kimya, biyoloji laboratuvarlarının bulunmadığı, bulunsa bile birçok okulda yeterince deney malzemesine ulaşamadığı gibi sosyo ekonomik düzeyle bağlantılı olan sebeplerden dolayı kültür ve dile bağlı olarak İngilizce formu alan öğrenciler lehine çalışmış olabileceği konusunda görüş bildirilmişlerdir. Bu uzmanların bir kısmı bu maddenin sosyo ekonomik nedenlerden dolayı Türkiye'nin farklı türdeki okullarında ya da farklı bölgelerinde yanlı işleyebileceğini belirtmişlerdir. Uzmanlardan biri iletkenliğinin test edilmesine ilişkin ülkemiz müfredatında benzer soru tiplerinin kullanılmayışının Türkçe formun aleyhine bir sonuca neden olabileceğini belirterek; dolaylı olarak kullanılan soru tipinin Türkçe

formu alan öğrencilere göre ağır olduğunu belirtmiştir. Teknolojik farklılıklar, bilim ve teknolojiadaki gelişmeleri takip eden birey sayısının azlığını da bir başka neden olarak belirtilmiştir. Bu maddenin yanlış olduğu konusunda görüş bildiren uzmanlardan bazıları maddeye ait olan okuma parçasının Türk öğrenciler için karmaşık bir anlatıma sahip olduğunu, paragrafta kullanılan dilin ilköğretim 8. Sınıf öğrencileri için ağır olduğunu belirtmiştir. Uzmanların bir kısmı kullanılan soru tipinin Türk öğrencilerin alışık olmadığı bir stil olduğu vurgulamıştır. PISA sınavında yer alan farklı soru tipleri, eğitimleri boyunca genellikle tek soru tipine alışmış öğrencilerde sıkıntı yaratabilmektedir. Aslında çoktan seçmeli soru tipine alışık olan Türk öğrenciler sorunun farklı bir biçimde sorgulanmasından ötürü dezavantaj sahibi olabilir. Yani sorunun tipinin yanı sıra nasıl sorgulandığının da yanlışlık nedeni olabileceği konusunda bir görüş birliği bulunmaktadır.

Madde 1 için görüş bildiren uzmanlar genel olarak Sosyo ekonomik sebepler, metindeki karışık ifadeler ve soru tipinden dolayı maddenin İngilizce formu alan öğrencilerin lehine işleyebileceğini savunmuştur.

PISA raporundaki bilgiler incelendiğinde madde 1 açıklanan ve üzerinde çalışılan maddeler arasındaki en kolay sorudur. Maddeyi alan öğrencilerin %79'u bu maddeyi doğru cevaplamıştır. Maddenin doğru cevaplanması öğrencinin konuyu bilimsel olarak açıklayabilmesini gerektirmektedir.

Yapılan analizler incelendiğinde bu maddenin MH tekniği ile yapılan analiz sonucunda her iki kitapçıkta da A düzeyinde İngiltere lehine DMF içerdiği belirlenmiştir. LR tekniğine göre madde DMF içermemektedir. Madde 1 hem işaretli alan indeksleri hem de işaretli alan indeksleri tekniğine göre DMF içeriyor görünmektedir. Birinci kitapçığa ilişkin verilerle yapılan analizde madde ÇBDMF göstermekte iken; beşinci kitapçık verilerine göre madde İngiltere lehine TBDMF göstermektedir.

Birinci kitapçıkta yer alan 8. madde için uzmanların % 67'si yanlış işleyebileceği konusunda görüş bildirmiştir. Uzmanlar metinde yer alan karışık ifadeler, sorgulanan içeriğe aşina olmama ve müfredat kapsamındaki farklılıklar nedeni ile bu maddenin Türkçe formu alan öğrencilere dezavantaj sağlayacağı yönünde görüş bildirmişlerdir.

Uzmanlar bu maddenin Türkiye fen ve teknoloji dersi müfredatı kapsamında olmadığını, okuma parçasının kısmen coğrafya dersi müfredatında jeoloji konusu içerisinde yer alabileceğini belirtmişlerdir. Ayrıca resimdeki İngilizce formda “shale” kelimesinin karşılığı Türkçe formda “kilin sıkışması ile oluşan tortul kayaç” biçiminde çevrilmiştir. Uzmanlar ilköğretim 8. Sınıf öğrencilerinin bir kelime ile belirtilebilecek bir kavramın uzun bir kelime grubuyla açıklanmasının karışıklığa ve anlaşmazlığa neden olabileceği konusunda hemfikirlerdir. Grisay (2003)’e göre, kelime ve cümlelerin ortalama uzunlukları diller arasında farklılık gösteren bir özellik olduğunu ve ölçme araçlarının uyarlanması aşamasında çevirmenler tarafından tamamen kontrol edilemediğini belirtmiştir. Benzer biçimde soruya ait okuma parçasının başlığının “Grand Kanyon” gibi özel bir isim olması nedeniyle Türk öğrencilerin aleyhine işleyebileceği belirtilmiştir. Özellikle yaşadığı şehirde veya ülkede kanyon parklarına geziye gitme imkanı olan öğrencilerin lehine işleyecek; Türkiye’de yaşayan öğrenciler için ise dezavantaj sağlayabilecek bir madde olduğu savunulmuştur. Soruların cevaplarının Türk öğrencilerin alışık olduğu biçimde okuma parçasının içinde saklı olmadığı, bu okuma parçasındaki soruyu yanıtlayabilmek için bahsi geçen yer hakkında bilgi sahibi olunması gerektiği belirtilmiştir. Aynı zamanda uzmanlardan bir kısmı, soruya ait okuma parçasında soruyu çözmek için gerekli olmayan cümlelerin bulunduğunu belirtmişlerdir. Uzmanlar bu durumun bu uzunluktaki soru tipine henüz alışık olmayan 8. sınıf Türk öğrencilere dezavantaj sağlayabileceğini düşünmektedir. Bu maddede yanlışlığın temel nedeni olarak daha çok “kültüre aşina olmama” üzerinde durulmuştur.

Yapılan DMF analizleri incelendiğinde bu maddenin MH tekniğine göre C düzeyinde İngiltere lehine DMF içerdiği belirlenmiştir. LR tekniği analiz sonuçlarına göre ise madde B düzeyinde ÇBDMF göstermektedir. Alan indeksleri tekniği sonuçlarına göre madde DMF göstermemektedir.

Bir diğer taraftan bu soruyu doğru cevaplamak, PISA 2006 fen ölçeğinde 451 puanının zorluğuna karşılık gelmektedir. Soruyu sınava katılan ve bu kitapçığı alan öğrencilerin %68 ‘i doğru cevaplamıştır. Soruda öğrencilerden konuyu bilimsel olarak açıklayabilmeleri beklenmektedir.

Açıklanan maddelerden birinci kitapçıkta yer alan 9. madde için uzmanların %60'ı yanlış işleyebileceği konusunda görüş bildirmiştir. Uzmanların verdiği cevaplar incelendiğinde aynı okuma parçasına ait olan 8. maddenin yanlışlık nedenlerine benzer nedenlerle maddenin İngilizce formu alan öğrencilerin lehine işleyebileceği belirtilmiştir. “Grand Canyon” un daha bilindik bir bölge olması sebebiyle İngilizce formunu alan öğrenciler lehine çalışmasının beklendiği ve bu maddenin yanlış olarak kabul edilebileceği konusunda uzmanlar ortak görüş bildirmişlerdir. Uzmanların verdiği cevaplar incelendiğinde kültürel farklılıkların temel yanlışlık nedeni olarak belirtildiği görülmüştür. Türkçe formu alan öğrencilere “Büyük Canyon’daki oluşum” yerine “Peri bacalarının oluşumu” sorulmuş olsaydı maddenin Türk öğrencilerin lehine işleyebileceği belirtilmiştir. Maddenin yanlış olabileceği hakkında görüş bildiren uzmanlardan bir kısmı maddenin cevabı için kesin bir yargı bulunmadığını ve değişik kültürlerde maddenin farklı biçimlerde cevaplanabileceğini belirtmiştir.

Yapılan DMF analizleri incelendiğinde bu maddenin MH ve LR tekniklerine göre DMF göstermediği belirlenmiştir. alan indeksleri ile yapılan analizlerde maddenin elde edilen alan indeksi değeri 1,82 olup DMF içermediği belirlenmiştir.

Bu maddede PISA sınavında 411 puanın zorluğuna karşılık gelmekle birlikte soruyu öğrencilerin % 76’sı doğru cevaplamıştır. Madde de öğrencilerden konuyu bilimsel olarak açıklayabilmeleri beklenmektedir.

Birinci kitapçıkta yer alan bu iki madde için uzmanlar; maddelerin gezi alışkanlığı olan, kültürel açıdan zengin bir çevreden gelen öğrencilerin lehine çalışmasını beklediklerini belirtmişlerdir.

Uzmanların % 33’ü açıklanan maddelerden birinci kitapçıkta yer alan 16. maddenin İngilizce formu alan öğrenciler lehine yanlış olmasını beklediklerini belirtmişlerdir. Uzmanlar çeviri sonucu kullanılan sözcük, kavramları Türkçe formunu alan öğrenciler bakımından anlaşılır bulmamışlardır. Örnek olarak; “bir mermer parçasının gece boyunca sirke içine konmadan önceki kütlesi 2,0 gramdır” ifadesinin anlaşılır olmadığını ve kafa karıştırıcı bir ifade olduğunu savunmuşlardır. Uzmanlar bu cümlelerin İngilizce’sinin basit ve net olduğunu Türkçe ifadesinin anlaşılmasının daha zor olduğunu dile getirmişlerdir. Okuma parçasında yer alan görsel objenin açıklayıcı

olmadığını bildirmişlerdir. Aynı zamanda kireç tozunun asidik ortamda aşınıp aşınmayacağını ve sirkenin asetik asit (etanoik asit) olduğunu bu yaş grubunun bilemeyeceğini; bir çözeltinin asidik mi yoksa bazik mi olduğunun lise müfredat konularında yer aldığını belirtmişlerdir.

Yapılan DMF analizleri incelendiğinde bu maddenin MH, LR ve alan indeksleri tekniğine göre DMF içermemektedir.

Bu maddeyi doğru cevaplamak PISA testinde 460 puanlık bir zorluğa karşılık gelmektedir. OECD ülkelerinde sınava giren ve bu soru ile karşılaşan öğrencilerin % 67 si bu soruyu doğru cevaplamıştır. Soruda öğrenciden bilimsel kanıtları kullanarak cevabı işaretlemesi beklenmektedir.

Birinci kitapçıkta yer alan 17. madde için uzmanların % 40'ı İngilizce formu alan öğrencilerin lehine yanlı işleyebileceği konusunda görüş bildirmiştir. Uzmanların %13'ü (iki) ise Türkçe formu alan öğrencilerin lehine yanlı işleyebileceği konusunda görüş bildirmiştir. Uzmanlar; genetik yapısı değiştirilmiş gıdaların oldukça bilgi verici ve sorgulayıcı bir içeriğe sahip, güncel bir konu olmasına rağmen bu konuya ilişkin temel bilgilerin lise 11. sınıf müfredatında yer aldığını belirtmişlerdir. Müfredat kapsamı dışında bir konu olduğu için bu maddenin Türkçe formu alan öğrenciler için uygun olmayacağı “madde yanlıdır” görüşünde olan uzmanların ortak fikridir. Uzmanlar; müfredat dışındaki farklılıkların göz ardı edildiği taktirde bile Türk öğrencilerin yeterince belgesel izleme alışkanlıklarının olmayışı; çevre bilinci ve biyoteknolojik eğitimlerin erken yaşlarda verilmeyişini de bu maddede yanlılığa sebep olabilecek etkenler arasında göstermişlerdir. Maddenin yanıtlanmasının güncel konulara hakimiyetin yanı sıra deneysel zeka gibi üst düzey davranışların gerekliliği de vurgulanmıştır. Buna rağmen uzmanlardan biri, bir tarım bölgesi olan Türkiye'deki öğrencilerin ülkemizin konumu ve tarımsal olanakları nedeni ile bu soruyu İngiltere'de yaşayan öğrencilere göre daha kolay yanıtlayabileceği görüşündedir. Türkçe formu alan öğrencilerin lehine yanlı işleyebileceği konusunda görüş bildiren uzmanlardan biri de İngilizce formun eksik bilgi içerdiğini; Türkçe formda daha fazla açıklayıcı bilgi yer aldığını savunmuştur.

Yapılan DMF analizleri incelendiğinde bu maddenin MH tekniğine göre DMF içermediği LR tekniğine göre ise A düzeyinde ÇBDMF gösterdiği sonucuna ulaşılmıştır. Alan indeksleri tekniği çerçevesinde madde hem İşaretli alan indeksleri hem de İşaretsiz alan indeksleri sonucuna göre DMF içermektedir. Maddeye ilişkin gruptan elde edilen madde karakteristik eğrileri bu analiz doğrultusunda kesişmekte ve ÇBDMF göstermektedir.

Madde 17 OECD ülkelerinde bu kitapçığı alan öğrencilerin % 74 ‘ü tarafından doğru cevaplandırılmıştır. Bu maddenin cevaplanması için öğrenciden bilimsel konuları belirleyebilme davranışı beklenmektedir.

Uzmanlar % 60’ı açıklanan maddelerden beşinci kitapçıkta yer alan 12. maddenin İngilizce formu alan öğrenciler lehine yanlı olmasını beklediklerini belirtmişlerdir. Yanlı işleyebileceğini belirten uzmanlardan biri maddenin müfredat kapsamında ve kendi okullarında deneyle yapılabileceğini belirtmiş ancak esas sıkıntının koşullar ve malzeme temininde yaşanan/yaşanacak sıkıntılar olduğunu, okullardaki imkansızlıkların madde yanlılığını etkileyebileceğini vurgulamıştır. Örnek olarak; madde ile ilişkili okuma parçasında anlatılan deneyde malzemeler arasında yer alan “mineral yağın” bulunmasının çok sıkıntılı olduğunu vermiştir. Aynı uzman bu maddenin ülkemize has sınıfsal farklılıklardan dolayı Türkiye’nin farklı bölgelerinde bile farklı işleyebileceği hususunda görüş bildirmiştir. Benzer olarak bir diğer uzman bu sorunun yanıtlanmasında öğrencilerin “kontrollü deney” hakkında bilgi sahibi olması gerektiğini vurgulamış ve ülkemizde yürütülen Fen bilimleri derslerinin ağırlık olarak çok azının laboratuvar koşullarında yürütüldüğünü bildirmiştir. Maddenin Türkçe formu alan öğrenciler için dezavantaja neden olacağını savunan uzmanlar aynı zamanda müfredattaki farklılıkların ve mevcut metindeki çeviri yetersizliğinin yanlılığa sebep olabileceğini belirtmişlerdir. Bir başka uzman bu tip soruların uluslararası tatil kültürüne sahip olan ülke öğrencilerinin cevaplayabileceği nitelikte bir soru olduğunu belirtmiştir. Türk öğrencilerin çok azının böyle bir şansa sahip olması nedeniyle bu uzman Türkçe formu alan öğrencilerin çoğunluğunun bu maddeyi yanıtlamakta zorlanacağını belirtmiştir.

Yapılan DMF analizleri incelendiğinde bu maddenin MH tekniğine göre yapılan analizlerde Türkçe formu alan öğrenciler lehine C düzeyinde DMF içerdiği

belirlenirken LR tekniğine göre yapılan analizlerde madde A düzeyinde DMF içermektedir. Alan indeksleri tekniği sonuçlarına göre madde hesaplanan İşaretli alan indekslerine göre DMF içermezken; işaretli alan indeksleri değerine göre DMF göstermemektedir. DMF'nin biçimi ÇBDMF olup gruplardan elde edilen madde karakteristik eğrileri kesişmektedir.

Bu maddenin cevaplanması PISA ölçeğinde 499 puanlık bir zorluğa denk gelmektedir. Bu kitapçığı alan öğrencilerin % 58 'i bu soruyu doğru cevaplandırmıştır. Soru orta güçlük düzeyinde bir sorudur. Sorunun cevaplanması için öğrencinin bilimsel konuları belirleyebilmesi gerekmektedir.

Beşinci kitapçıkta yer alan 13. madde için uzmanların % 47'si İngilizce formu alan öğrencilerin lehine yanlı işleyebileceği konusunda görüş bildirmiştir. Ülkemizde teori oluştururken bir bilim insanının izlediği temel aşamaların lise müfredatında olması, laboratuvar ortamında böyle bir deneyin yapılmasının zorluğu, çeviri yetersizliği... gibi nedenlerle bu maddenin Türkçe formu alan öğrenciler için dezavantaj sağlayacağı görüşündedirler. Uzmanlar ZnO ve Mineral yağ kullanılmasını anlayabilmek için bu maddelerin güneş ışınlarına karşı davranışının bilinmesi gerektiğini, bunun da söz konusu düzey için müfredat dışı olduğu, metinde kullanılan sözcüğün (UV ışınları) Türkiye'de o yaş grubuna hitap etmediğini belirtmişlerdir.

Yapılan DMF analizleri incelendiğinde bu maddenin MH ve LR tekniklerine göre DMF göstermediği belirlenmiştir. Madde Tepki kuramına göre hesaplanan İşaretli alan indekslerine göre madde Türkçe formu alan öğrencilerin lehine DMF içerirken; işaretli alan indekslerine göre madde DMF göstermektedir. Madde Türkçe formu alan öğrencilerin lehine işlemesine rağmen; alan indeksleri değerleri birlikte değerlendirildiğinde; iki gruptan elde edilen madde karakteristik eğrileri yine de kesişmekte ve ÇBDMF göstermektedir.

Beşinci kitapçık madde 13 ile karşılaşan tüm öğrencilerden % 43'ü bu soruyu doğru cevaplandırmıştır. Tıpkı 12. maddede olduğu gibi öğrencinin maddeyi doğru cevaplandırması için bilimsel konuları belirleyebilmesi gerekmektedir. Aynı köke ait olan 12. maddeye göre daha zor bir soru olduğu belirlenmiştir.

Uzmanlar % 40'ı açıklanan maddelerden beşinci kitapçıkta yer alan 14. maddenin İngilizce formu alan öğrenciler lehine yanlı olmasını beklediklerini belirtmişlerdir. Uzmanlardan yalnızca biri (%7) bu maddenin Türkçe formu alan öğrenciler lehine olduğunu belirtmiştir. Uzmanlardan alınan yanıtlar incelendiğinde yanlılık nedenleri olarak; çeviri yetersizliği, Türkiye’de verilen eğitimin daha çok teorik olması ve teorik eğitimin pratik eğitime yansımaması, uygun laboratuvar koşullarının sağlanamaması sıralanmıştır. Uzmanlardan biri bu sorunun mantıksal akıl yürütme gerektiren bir madde olduğunu ve bu maddede kültürel ve kişisel farklılıkların ön plana çıkabileceğini belirtmektedir.

Yapılan DMF analizleri incelendiğinde bu maddenin MH tekniğine göre DMF içermediği LR tekniğine göre ise A düzeyinde ÇBDMF gösterdiği sonucuna ulaşılmıştır. Alan indeksleri tekniği sonuçlarına göre madde her iki indekse göre de DMF içermektedir. DMF’nin biçimi ÇBDMF olup madde daha çok İngilizce formu alan öğrenciler lehine işlemiştir.

Bu madde de diğerlerinden farklı olarak öğrencinin bilimsel kanıtları kullanabilmesi beklenmektedir. OECD ülkeleri çapında soruyu doğru cevaplama yüzdesi yalnızca % 27’dir.

Aynı okuma parçasına ait 12. 13. ve 14. maddelere ilişkin verilen yanıtlar göz önüne alındığında verilen pratik eğitimlerin fazlalığı ve sosyo ekonomik koşullar göz önüne alındığında İngilizce formu alan öğrenciler lehine çalışmasının beklendiği konusunda uzmanlar ortak görüş bildirmişlerdir. Aynı zamanda 8. sınıf fen bilimleri müfredatı düşünüldüğünde paragrafta anlatılan deneyin anlaşılmasının bu yaş grubu için oldukça zor olduğu belirtilmiştir.

Beşinci kitapçıkta yer alan 18. madde için uzmanların % 13’ü İngilizce formu alan öğrencilerin lehine; uzmanların %13’ü (iki) de Türkçe formu alan öğrencilerin lehine yanlı işleyebileceği konusunda görüş bildirmiştir. Uzmanların verdiği yanıtlar göz önüne alındığında yanlılık sebebi olarak müfredat kapsamındaki farklılıklar ve Türkiye’deki öğrencilerin öğrenme alışkanlıklarına uygunluk ön plana çıkmıştır. Bu madde için uzmanların çoğunluğu (%74) yanlı olmadığı yönünde görüş bildirmiştir. Bu durum MH dışındaki tekniklerle yapılan analizler ile uyum göstermektedir.

Yapılan DMF analizleri incelendiğinde bu maddenin MH tekniğine göre C düzeyinde İngilizce formu alan öğrenciler lehine DMF içermektedir. Maddenin LR tekniğine göre DMF göstermediği belirlenmiştir. Alan indeksleri tekniği sonuçlarına göre madde hem İşaretli alan indeksleri hem de İşaretsiz alan indekslerine göre DMF göstermektedir. İşaretli alan indeksleri sonuçlarına göre madde İngiltere lehine işlemiş; DMF'nin biçimi ise ÇBDMF şeklinde belirlenmiştir.

Beşinci kitapçıkta yer alan 19. madde için uzmanların % 13'ü İngilizce formu alan öğrencilerin lehine; uzmanların %7'si ise Türkçe formu alan öğrencilerin lehine yanlı işleyebileceği konusunda görüş bildirmiştir. Uzmanların verdiği yanıtlar göz önüne alındığında İngilizce formu alan öğrencilerin lehine yanlılık bildiren uzmanlar metin içerisindeki karışık ifadelerin ve genel kültür düzeyini etkileyen faktörlerin bu maddeyi yanıtlamada yanlılığa sebep olduğunu bildirirken; Türkçe formu alan öğrencilerin lehine yanlı işleyebileceği konusunda görüş bildiren uzman ise Türk öğrencilerin öğrenme alışkanlıklarının yanlılık sebebi olarak bildirmiştir.

Yapılan DMF analizleri incelendiğinde bu maddenin sadece MH tekniğine göre DMF içermediği belirlenmiştir. MH tekniğine göre maddenin DMF düzeyi B'dir. Maddenin çalışmada kullanılan diğer tekniklere göre DMF içermediği belirlenmiştir.

Uzmanların büyük bir kısmı beşincikitapçıkta yer alan 18. ve 19. madde için 8. sınıf öğrencilerinin mevcut bilgileri ile kolayca cevaplanabileceğinden dil ve kültür açısından yanlılığa sebep olmasının beklenmediği belirtilmiştir.

PISA raporu incelendiğinde her iki maddenin de benzer zorluk düzeyine karşılık geldiği söylenmektedir. Her iki soru da OECD ülkelerinde %75 düzeyinde doğru cevaplanmıştır. Her iki maddeyi cevaplamak için öğrenciden bilimsel konuları belirleyebilme davranışına sahip olması beklenmektedir.

Bu çalışmanın yanı sıra çalışma dilden soyutlanarak ikinci bir form hazırlanmış ve Türkçe formu alan öğrencilere yöneltilen açıklanan sorular fen bilimleri alanlarından mezun ilköğretimde öğretmenlik yapan öğretmenlerle paylaşılmış ve kültürel bakımdan yanlılık araştırılmıştır.

3.3.2. İkinci Forma Ait Bulgular:

Çalışmanın bu bölümünde fen bilimleri alanında öğretmenlik yapan uzmanlara telekomünikasyon ve internet yolu ile ulaşılmıştır; ulaşılan uzmanlardan 30'undan yanıt alınabilmektedir. Görüşlerine başvurulmuş uzmanlara öncelikle yanlılık açıklanmış; daha sonra PISA'nın uluslararası bir sınav olduğu, soruların farklı dillere çevrildiği, hazırlanan ve çevrilen soruların bazı ülkeler – kültürler için uygun olamayabileceği ve Türkiye'de bu sınava katılanların ilköğretim 8. Sınıf öğrencileri olduğu hatırlatılmıştır. Tüm bunları göz önünde bulunduran uzmanların soruda Türk öğrencilere avantaj ya da dezavantaj sağlayacak durumları değerlendirmeleri ve bunları belirtmeleri istenmiştir. Çalışmaya katılan uzmanların kültür bakımından yanlılığa ilişkin görüşlerini esas alan tablo 22 aşağıda mevcuttur.

Tablo 22: Uzmanların PISA Fen Testi Maddelerinin Kültürler ve Diller Arasında DMF Göstermesinin Olası Nedenlerine İlişkin Görüş Dağılımları (Form 2)

1.KİTAPÇIK	MADDE 1		MADDE 8		MADDE 9		MADDE 16		MADDE 17	
	(S213Q02)		(S426Q03)		(S426Q05)		(S485Q03)		(S508Q03)	
	f	%	f	%	f	%	f	%	f	%
YANLI	22	73	15	50	13	43	10	33	11	37
YANLI DEĞİL	8	27	15	50	17	57	20	67	19	63

5.KİTAPÇIK	MADDE 1		MADDE 12		MADDE 13		MADDE 14		MADDE 18		MADDE 19	
	(S213Q02)		(S447Q02)		(S447Q03)		(S447Q04)		(S477Q02)		(S477Q03)	
	f	%	f	%	f	%	f	%	f	%	f	%
YANLI	22	73	16	53	14	47	13	43	11	37	7	23
YANLI DEĞİL	8	27	14	47	16	53	17	57	19	63	23	77

Tablo 22'de yer alan birinci ve beşinci kitapçıkta açıklanan maddelere ilişkin yanlılık görüşleri incelendiğinde her iki kitapçıkta da yer alan 1. maddenin yanlı olabileceği konusunda yanıt alınan uzmanların birçoğu (22 uzman (%73)) görüş bildirmiştir. Uzmanların verdiği yanıtlar incelendiğinde okuma parçası olarak kullanılan metnin karmaşıklığı; kullanılan sözcük ve kavramların Türk öğrenciler için anlaşılmasının

güçlüğü ve sorgulanan içeriğe öğrencilerin aşına olmaması; Türkiye’de farklı bir müfredat kapsamında inceleniyor olmasının; Türk eğitim sisteminde iletkenliğin bu biçimde sorgulanmayacağı belirtilerek bir anlamda kullanılan soru tipinin yanlışlığa neden olabileceği belirtilmiştir. Maddenin yanlış olduğunu belirten tüm uzmanlar özellikle anlatımın ağır olması ve iletkenlik sorgulanmasının bizim kültürümüzde böyle yapılmadığı konusunda ortak görüş bildirmişlerdir.

Birinci kitapçıkta yer alan diğer dört sorunun yanlış olup olmadığına ilişkin uzmanlar şu biçimde görüş bildirmişlerdir.

Aynı okuma parçasına ait olan 8. madde için uzmanların % 50’si; 9. madde için ise uzmanların % 43’ü yanlış olacağını ve Türk öğrenciler için dezavantaj sağlayacağını savunmuştur. Her iki madde için benzer nedenlerin öne sürüldüğü dikkat çekmektedir. Olası yanlışlık nedenleri olarak; sorgulanan içeriğe aşına olunmaması, okuma parçasında kullanılan kelimelerin 8. sınıf öğrencileri için ağır olduğu vurgulanmıştır. Bildirilen yanlışlık nedenleri göz önüne alındığında bu durumun bizim kültürümüzde yetişen öğrenciler için sorun yaratabilecek nedenler olduğu söylenebilir. Anlatımdaki zorluğa çeviri sonucu kullanılan sözcüklerin zorluk derecelerinin farklılıkları, çeviriden sonra meydana gelen metindeki anlamsal farklılıklar neden olmuş olabilir. Bu da doğrudan bu okuma parçasına ait olan 2 sorunun öğrencilerimize dezavantaj sağlamasına neden olabilmektedir. Soru yanlıdır biçiminde görüş bildiren uzmanların çoğu bu okuma parçasına ait olan 8. ve 9. Maddenin bizim kültürümüzde sorgulanmasının uygun olmayacağını belirtmişlerdir. Uzmanlardan biri 9. Maddede yeralan çeldiricilerin 8. Sınıf öğrencilerine göre oldukça zor ve kuvvetli olabileceğini belirtmiştir. Yine aynı maddede yer alan okyanus kavramının müfredatta geçmemesi nedeni ile bilim ve kültürdeki gelişmeleri takip etmeyen öğrencilere zorluk yaratacağı 1 uzman tarafından belirtilmiş; uzman Türkiye’de 8. Sınıf öğrencilerinin böyle bir bilinçle büyümediklerini vurgulamıştır. Bir bakıma bu kültürümüze yönelik öğrenme-öğretme yaklaşımlarının farklılığı olarak algılanabilir.

16. Madde için uzmanların %33’ü “yanlıdır” demiştir. Maddenin Türkiye’de sınava girecek öğrencilere dezavantaj sağlayacağı yönünde görüş bildirilmiştir. Bunun nedenleri olarak sorgulanan içeriğe aşına olmama; kullanılan sözcük ve kavramların metnin ve metine bağlı soruların anlaşılabilirliğini engellemesi, sorgulanan içeriğin

müfredat kapsamında yer almaması ve maddenin kültürümüze uygun nitelikte olmaması belirtilmiştir.

17. maddeyi uzmanların %37'si yanlış bulmuştur. Yanlılığın olası nedenleri olarak genel olarak; müfredat kapsamındaki farklılıklar ön plana çıkmaktadır. Uzmanlardan biri soruda ne karşılaştırıldığının anlaşılmadığını belirtmiştir. Başka bir uzman paragrafta anlatılan ile sorgulanan şey arasında bağlantı bulunmadığını belirtmiştir. Aslında aynı neden bir başka uzman tarafından şöyle açıklanmıştır: “Çevirisi yapılan okuma parçası ve sorunun kültürümüze yeterince uyarlanmadığını düşünüyorum. Bu Türk öğrenciler için dezavantaj sağlayacaktır.”

Beşinci kitapçıkta yer alan açıklanan sorulara ilişkin yanıtlarda da birinci kitapçıkta olduğu gibi maddelerin yanlış olduğunu düşünen uzmanların hepsi maddelerin Türkiye aleyhine olacağı hususunda ortak görüş bildirmiştir.

Beşinci kitapçığın ortak okuma parçasına ait 12. 13. ve 14. maddeleri için uzmanların sırayla %53'ü, %47'si, % 43'ü “ yanlışdır” demiştir. 3 soru için de benzer yanlışlık nedenleri belirtildiği görülmüştür. Verdikleri yanıtlar incelendiğinde temel neden olarak kullanılan sözcük ve kavramların metnin ve metine bağlı soruların anlaşılabilirliğini engellemesi, sorgulanan içeriğin müfredat kapsamında yer almaması ve maddenin kültürümüze uygun nitelikte olmaması belirtilmiştir. Bu soruların yanlış olduğunu düşünen uzmanların 3'ü sorgulanan içeriğin Türkiye’de eğitim gören 8. Sınıf öğrencilerinin seviye ve ihtiyaçlarına uygun olmadığını özellikle vurgulamıştır. 13. Soru için bir uzman soruyla karşılaşan öğrencinin neyi karşılaştıracağını anlamasının oldukça zor olduğunu belirtmiştir.

Uzmanların %.37'si 18. maddenin yanlış olduğunu düşünmektedir. Verilen yanıtlar incelendiğinde okuma metni ve soru arasında ilişkinin kurulamayışı, kullanılan sözcük ve kavram nedeniyle maddenin yanlış işlediğinin düşünüldüğü belirtilmiştir. Bunun ise yüksek ihtimalle çeviriden kaynaklandığını ve çevirinin bu anlamsal farklılıklara neden olduğu görüşündedirler. Belirgin örneğini şu biçimde vermişlerdir. Uzmanlar okuma parçasında virüslerden bahsedildiğini fakat soruda bakterilerle ilgili soru yöneltildiğini belirtmiştir. Uzmanlar bilgisi verilmeyen konudan soru

sorulamayacağını ifade etmişlerdir. 3 uzman verilen okuma parçasında örnek olmayışının öğrenciler için dezavantaj sağlayacağını belirtmişlerdir.

Tıpkı madde 18. maddede olduğu gibi 19. maddede de direkt bakterilere dönük bir bilgi sorusu sorulduğu, ancak verilen metinde bu bilgiye ulaşılacak bir anlatım sergilemediğinden uzmanların %23'ü maddenin yanlış olduğu görüşündedir. Aynı okuma parçasına ilişkin bu iki sorunun öğrencilere dezavantaj sağlamasına neden olanın metin ve sorulardaki anlamsal farklılığın neden olduğu görüşü hakimdir.

Araştırmada uzmanlar tarafından öne sürülen yanlışlık nedenlerinin diğer çeviri ve uyarılama çalışmalarına dönük yapılan çalışmalardaki nedenler ile uyum içinde olduğunu söylenebilir (Ercikan ve Koh, 2005; Çet, 2006)

Çalışmanın bu bölümünde farklı formlara yanıt veren uzmanların yanlışlık kararlarının birbirine benzer olup olmadığı belirlenmeye çalışılmış bunun içinde her bir soru için sınıflayıcı ölçekle elde edilmiş 2x2 lik tablolar oluşturulmuştur. Bu tablolar ile her bir soru için kısmi mertebe katsayısı olarakta bilinen Phi katsayısı hesaplanmıştır. Aşağıdaki tabloda her bir soru için elde edilen Phi katsayısı değerleri bulunmaktadır.

Tablo 23: Farklı Formlara Yanıt Veren Uzmanlardan Elde Edilen Phi Katsayısı Değerleri

1. KİTAPÇIK					5. KİTAPÇIK					
MADDE 1 (S213Q02)	MADDE 8 (S426Q03)	MADDE 9 (S426Q05)	MADDE 16 (S485Q03)	MADDE 17 (S508Q03)	MADDE 1 (S213Q02)	MADDE 12 (S447Q02)	MADDE 13 (S447Q03)	MADDE 14 (S447Q04)	MADDE 18 (S477Q02)	MADDE 19 (S477Q03)
-0,07	0,16	0,16	0	0,16	-0,07	0	0	0,03	-0,1	-0

Phi katsayısının önem testi için dağılım ki kare dağılımına uyduğundan bu dağılım değerleri kullanılmıştır. Karşılaştırma sonucunda tablo 23'de elde edilen tüm phi katsayıları ($\phi_{1,8,9...19} < \chi^2_{(1,0,05)} = 3,841$) manidar olmadığından farklı formları alan uzmanların kararları benzerlik göstermektedir.

İlgili çalışmalar incelendiğinde DMF çalışmalarının Dil ve Kültür çerçevesinde yapılandırıldığı bir gerçektir. Dolayısıyla hem dil hem de kültür DMF'nin temel nedenleri arasındadır. Bu çalışmada üzerinde durulan dil ve kültürden kaynaklanan yanlılığı, hem kültür farklılıklarının hem de dil farklılıklarının farklı biçimde etkileyeceği ve etkilediği düşünülmektedir. Kültürün aynı dilin farklı olduğu Kanada örneğinde olduğu gibi alan indeksleri tekniği ile bulunan DMF'li madde sayısı kültürün farklı dilin aynı olduğu Avustralya-İngiltere örneğine göre göreceli olarak fazla bulunması çalışılan bu örneklerde DMF de dilin daha belirleyici bir etken olduğunu düşündürebilir. Ancak Dil, hem bütün insan nüfusları için ortak bir yetenek, hem de kültür topluluklarını birbirinden ayıran farklılaşmış bir uygulamadır. Kültür ise, insanın toplumsal çevreye genel uyarlama yöntemi olduğu kadar, aynı zamanda topluluktan topluluğa farklılaşma gösteren bir yaşam ve düşünce tarzıdır. Nitekim aynı alt grup için MH tekniğiyle bulunan DMF'li madde sayısı eşitken; LR tekniği ile DMF olduğu belirlenen madde sayısı Avustralya-İngiltere grubunda alan indeksleri tekniğinde belirlenen madde sayısının üstünlüğü kadar olmasa da fazla bulunmuştur. Bu tanımlardan ve açıklamalardan dilin ve kültürün birbiri içerisinde kavramlar olduğu anlaşılmaktadır. Ayrı ayrı düşünülmesi yanlış yorumlamalara neden olabilir. Bu çalışma içinde kültürün, dilin hem de dil ve kültürün ortak etkileşiminin DMF'li madde sayısı üzerinde etkili olduğu, birinin diğerine baskın olmayacağı düşünülmektedir.

Yapılan DMF çalışmalarında çoklukla çeviri sorunları üzerinde durulmaktadır. Son yıllarda bilgi, yetenek ve tutum gibi özellikleri ölçmek amacıyla farklı dilleri konuşan bireylere çeviri testlerin uygulamasında artış görülmektedir. Bu artış farklı dillere çevrilen ölçme araçlarındaki olası DMF nedenlerini araştırmaya dönük bir çabayı da yanında getirmiştir. Bu çabalardan birinde sekiz üyeden bir komite oluşturulmuş ve çeviri ve uyarlama çalışmalarında ki dört temel DMF nedeni şu biçimde belirtilmiştir. Bunlar dört ana başlık altında toparlanmıştır.

- Ekleme ve silmelerin anlamı değiştirmesi,
- Kelime, ifade ve cümle yapısındaki farklılıklar (kültürün doğasında olan),
- Kelime, ifade ve cümle yapısındaki farklılıklar (kültürün doğasında olmayan),
- Madde yapısındaki farklılıklar (Gierl ve Khaliq, 2001).

Bu arařtırmada üçüncü alt problem bulgularının ilk bölümünde uzmanlardan elde edilen yanıtlar yukarıda belirtilen başlıklar ile uyum göstermektedir. Bunun yanı sıra özellikle kültürler arası çalışmalarda yanlılık kaynakları olarak çeviri yetersizlikleri, kavramların farklı kültürlerde farklı anlamlara gelebilmesi ya da bazı kavramlara farklı kültüre ait grupları farklı düzeylerde aşına olması, farklı eğitim programı kapsamaları görülmektedir (Ercikan, 1998; Grisay ve arkadaşları, 2007; Yıldırım ve Berberoğlu, 2009; Asil, 2011). Yine de řu nokta unutulmamalıdır. DMF'nin olası nedenleri araştırılırken řu iki sınırlık göz ardı edilmemelidir. Birincisi belirlenen yanlılık sebepleri belirgin olmamasına rağmen gerçek yanlılık sebebi olarak adledilmemelidir. İkincisi ise bu nedenler o maddenin bir grubun lehine olma durumunu açıkça gösteremez. Uzman görüşleri madde içeriğı ya da kapsam alanı ile ilgili DMF'ye bağılı sebepler hakkında genel bir bakış sağlayabilir ancak söz konusu olan bu bakış açısının farklı alt gruplarda “niçin” veya “nasıl” farklı fonksiyonlaştığını açıklamak için yeterli olmayacaktır. Bu soruların yanıtını verebilmek için test maddeleri ile yanıtlayıcı grup arasındaki etkileşime, madde yanıtlanırken öğrencilerin düşünme süreçlerine odaklanmak fayda sağlayabilir.

Genel olarak; hem dil hem de kültür temel alındığında uzmanlardan PISA fen sorularına ilişkin yanlılık ile ilgili alınan görüşlerde bu temel nedenler etrafında toplanmaktadır.

IV.BÖLÜM

SONUÇ VE ÖNERİLER

Araştırmanın bu bölümünde, ulaşılan bulguların ışığında sonuç ve önerilere yer verilmiştir.

4.1 SONUÇLAR

Bu çalışmada 2006 yılı PISA Fen testinde yer alan çoktan seçmeli maddelerde DMF araştırılmıştır. Birinci ve beşinci kitapçıkta yer alan çoktan seçmeli maddeler için incelenen DMF, kültür ve dil değişkenleri temel alınarak kültürler arası bir biçimde farklı örneklerde incelenmiştir. Bu çalışmada DMF belirleme yöntemlerinden MH, LR ve AİT kullanılmıştır. Elde edilen bulgular ışığında her bir alt probleme ait elde edilen sonuçlar şu biçimde özetlenebilir.

4.1.1 Alt Problem 1'e İlişkin Sonuçlar

PISA 2006 başarı testinde yer alan fen maddeleri, üç farklı DMF belirleme yöntemine göre, testteki madde kapsamı bakımından yanıtlayıcıların dili ve kültürü bakımından DMF'ye neden olmaktadır. Elde edilen bulgular, PISA 2006 Fen başarı testinin madde yanlılığı bakımından çeşitli açılardan tartışılır olduğunu göstermektedir.

Tablo 24: Kültür ve Dil Farklılıklarına Göre DMF Gösteren Maddelerin Tekniklerdeki Sayısal Dağılımı

	Madde sayısı	DMF'li Madde Sayısı			
		MH	LR	İAİ	İSAİ
Aynı Kültür- Farklı Dil (Kanada)	21	6	3	5	5
Farklı Kültür- Aynı Dil (Avustralya- İngiltere)	21	6	4	2	2
Farklı Kültür – Farklı Dil (1.Kitapçık) (Türkiye- İngiltere)	19	10	6	6	5
Farklı Kültür – Farklı Dil (5. Kitapçık) (Türkiye- İngiltere)	21	11	6	4	5

Tablo 24'deki kültür ve dil farklılıklarına göre DMF gösteren maddelerin dağılımı incelendiğinde; kültürel ve dilsel farklılıklar arttıkça DMF gösteren maddelerin sayısının ve içerdikleri DMF miktarının arttığı gözlenmiştir. Bu sayısal dağılımda DMF gösteren madde sayısının en fazla Maentel Hanzel (MH) tekniğinde bulunduğu anlaşılmaktadır.

Aynı Kültür- Farklı Dil (Kanada) için MH tekniği ile DMF li bulunan maddelerin üçü önemsenmeyecek düzeyde, üçü de orta düzeyde DMF içermektedir. MH analiz sonuçlarına göre DMF gösteren maddelerin dördü; test dili Fransızca olanlar lehinedir. Kanada alt gruplarının beşinci kitapçıkta yer alan 21 fen maddesine ilişkin performans farklılıklarının incelenmesine yönelik yapılan lojistik regresyon analizi sonuçlarına göre üç maddede DMF olduğu görülmektedir. DMF olduğu gözlenen soruların tümünde grup madde etkileşimi manidar bulunmuştur. Bu nedenle maddelerin tümünde çok biçimli DMF bulunmaktadır. DMF gösteren üç madde de DMF düzeyi önemli bulunmamıştır. İşaretli alan indeksleri yöntemleri ile elde edilen sonuçlara göre; beşinci kitapçıkta yeralan 21 maddeden beşi DMF içermektedir. Hesaplanan işaretli alan indekslerine göre ise beş madde de DMF bulunmuştur. İşaretli alan indekslerinde DMF belirlemek için kullanılan medyan değerine göre DMF gösterdiği belirlenen dört madde test dili Fransızca olanlar lehine DMF içermektedir; işareti negatif olan 18. Madde ise test dili İngilizce olanlar lehine DMF içermektedir. Alan indekslerinin birbirlerine göre durumları incelendiğinde iki madde dışında DMF gösterdiği belirlenen tüm maddelerde DMF'nin biçimi çok biçimli değişen madde fonksiyonu (ÇBDMF) olarak belirlenmiştir.

Farklı kültür- aynı dil (Avustralya- İngiltere) gruplarında yapılan MH analiz sonuçlarına göre; 21 maddenin altısı DMF içermektedir. Bu maddelerin dördünde bulunan DMF önemsenmeyecek düzeyde, biri orta düzeyde, diğeri de önemli düzeydedir. MH analiz sonuçları incelendiğinde; 3 madde İngiltere lehine; 3 madde ise Avustralya lehine işlemektedir. DMF gösteren 4 maddeden 2'si tek biçimli değişen madde fonksiyonu (TBDMF); 2'si ise ÇBDMF içermektedir. DMF gösteren dört maddenin de DMF düzeyleri Gierl ve arkadaşlarının belirlediği ölçütler çerçevesinde önemli bulunmamıştır. İşaretli alan indeksleri yöntemleri sonuçlarına göre; beşinci kitapçıkta yeralan 21 maddeden 2 madde DMF içermektedir. Bu 2 madde de İngiltere

lehine DMF içermektedir. İşaretsiz alan indeksleri ile elde edilen sonuçlara göre de 2 madde DMF içermekte olup, söz konusu maddelerin sıradışı değerlere sahip olması nedeniyle DMF biçimi ÇBDMF'dir. İşaretili alan indeksleri ile DMF gösterdiği belirlenen maddelerden biri ise TBDMF göstermektedir.

Türkiye İngiltere için yapılan analizlerde birinci kitapçık için elde edilen sonuçlar şu biçimde özetlenebilir. MH sonuçlarına göre 19 maddeden 10'u DMF içermektedir. Maddelerin beşi A düzeyinde, ikisi B, üçü ise C düzeyinde DMF içermektedir. MH analiz sonuçlarına göre sekiz madde referans grubun (İngiltere) lehine, iki madde ise odak grubun (Türkiye) lehine DMF içermektedir. birinci kitapçığa ilişkin Lojistik Regresyon tekniği sonuçlarına göre altı madde çok biçimli değişen madde fonksiyonu içermektedir. Bu altı maddenin beşinin önemsenmeyecek düzeyde; bir maddenin ise orta düzeyde DMF içerdiği bulunmuştur. Birinci kitapçığa ilişkin lojistik regresyon tekniği sonuçlarına göre altı madde ÇBDMF içermektedir. Bu altı maddenin beşinin (3, 4, 5, 15, 17) önemsenmeyecek düzeyde; bir maddenin (8. maddenin) ise orta düzeyde DMF içerdiği belirlenmiştir. Birinci kitapçıkta yeralan 19 maddeden altı maddenin alan indekleri tekniği ile yapılan analizler sonucu DMF içerdiği saptanmıştır. DMF gösteren altı maddenin de pozitifdir. İşareti pozitif olan bu maddeler referans grubun (İngiltere) lehine DMF içermektedir. Birinci kitapçık için İşaretsiz alan indeksleri iç ölçütü 5,655 olarak belirlenmiş ve bu iç ölçüte göre 5 maddenin DMF içerdiği sonucuna ulaşılmıştır. Bu maddelerin hepsi İşaretili alan indeksleri göre de yanlı bulunan maddelerdir. İşaretsiz alan indeksleri büyüklükleri 3'ün üstünde ise sıra dışı değerler olarak kabul edilmiştir ve bu açıklama doğrultusunda 3. Madde dışındaki tüm maddelerin bu tanımlamaya oldukça uyduğu ve DMF gösteren tüm maddelerin ÇBDMF gösterdiği belirlenmiştir.

Beşinci kitapçıkta ise 21 maddeden 11'i DMF içermektedir. Bu maddelerin ikisi A düzeyinde, dördü B, dördü ise C düzeyindedir. MH analiz sonuçları incelendiğinde; 6 maddenin referans grubun (İngiltere) lehine; 5 maddenin ise odak grubun (Türkiye) lehine DMF içerdiği belirlenmiştir. Beşinci kitapçık alt testinde lojistik regresyon tekniği ile yapılan analiz sonuçlarına göre altı madde DMF içermektedir. Bu maddelerden iki madde (2 ve 16) TBDMF gösterirken diğer dördü ÇBDMF

göstermektedir. Altı maddenin dördü önemsenmeyecek düzeyde, ikisi ise orta düzeyde DMF içermektedir.

Beşinci kitapçığa ait olan işaretli alan indeksleri değerleri incelendiğinde 21 maddeden dördünde DMF olduğu sonucuna ulaşılmıştır. Bu dört maddenin hepsi İngiltere lehine DMF göstermektedir. İşaretsiz alan indeksleri değerlerine göre beş maddede DMF bulunmaktadır. İki indeks değerinin birbirlerine göre durumu incelendiğinde beşinci kitapçıkta yeralan ve DMF gösteren maddelerden 2. madde dışındaki tüm maddelerin ÇBDMF gösterdiği sonucuna ulaşılmıştır.

Elde edilen bulgular incelendiğinde; karşılaştırılan alt gruplar arasındaki farklılıklar arttıkça DMF gösteren madde sayıları ve düzeylerinin arttığı sonucuna ulaşılmıştır.

Araştırma sonuçlarına bağlı olarak DMF içeren maddelerin sayısının yüksek olması, bu testten alınan puanların karşılaştırılması durumunda bizleri doğru yorumlardan uzaklaştıracaktır. Dolayısı ile PISA testinin, dil ve kültür bakımından bazı gruplara avantaj sağlıyor görünmesi bu test ile alınan kararların, geçerliliğini ve güvenilirliğini de tartışılır kılmaktadır. Bu durum, bu sınava giren tüm ülkeler için doğru karşılaştırmalar yapılmadığını, olası yanlılık nedenleri ortaya konulup gerekli değişmezlik koşulları sağlanmadığı sürece böyle uluslararası sınav sonuçlarında elde edilen bilgilerin doğru olmadığını düşündürür niteliktedir.

4.1.2 Alt Problem 2'ye İlişkin Sonuçlar

PISA çalışmasında, madde yanlılığı tespitinde kullanılan farklı metotların sonuçları arasında yüksek derecede uyum tespit edilememiştir. Çalışmada kullanılan üç teknik birarada düşünüldüğünde uyumlu sonuçlar alınamamıştır. Bunun yanı sıra KTK'ya ait tekniklerle MTK'ya bağlı tekniklerin kendi aralarında uyumlu olduğu sonucuna ulaşılmıştır.

4.1.3 Alt Problem 3'e İlişkin Sonuçlar

Maddelerin kültürler ve diller arasında farklı fonksiyon göstermesinin olası nedenlerinin neler olduğu konusunda uzman görüşlerinden ortaya çıkan ortak sonuç şu

biçimde özetlenebilir. Uzmanlar PISA maddelerinin yanlılık göstermesinin temel nedeninin çeviri, müfredat farklılıkları ve uyarlama yanlışlıkları sonucu metnin ağır bir dile dönüşmesi, sosyo ekonomik düzeyden kaynaklanan farklılıklar olduğunu belirtmişlerdir. Maddelerin karmaşık ve uzun metinlerle ilişkilendirilmesinin özellikle Türk öğrencilerin alışkın olmadığı bir formata dönüşmesinin bazı maddelerin yanıtlanmasını zorlaştırdığı fikrinin hakim olduğu sonucuna ulaşılmıştır. Uzmanlar PISA fen maddelerinin çoğunluklu olarak İngiltere lehine işleyeceği görüşündedirler.

Uzman görüşleri incelendiğinde; birden fazla dile çevrilen ölçme araçlarında çeviriden kaynaklanan farklılıkların ve kültürel farklılıkların yanlılığa neden olabileceği sonucuna ulaşılmıştır. Bu sonuçtan yola çıkarak uzmanların inceledikleri sorularla ilgili temel önerileri şu biçimdedir: Kısa, basit ve aktif cümleler kullanılması; deyim atasözü gibi kelime grupları ile; kültüre özgü sözcük ve terimlerin kullanılmasından kaçınılması gerekmektedir.

4.2. ÖNERİLER

PISA anketlerinden elde edilen veriler ile çok sayıda ülkeler arası karşılaştırmalar yapılmakta ve belirli kararlar alınmaktadır. PISA 2006 fen bilimlerine karşı fen testi maddelerinde kültür ve dile göre DMF gösteren maddeler yer alması nedeniyle bu testlerde düzenli olarak DMF analizleri yapılması önerilmektedir. Yapılan DMF analizleri sonrasında DMF gösteren maddelerin neden DMF gösterdiği nitel araştırmalarla desteklenmeli ve söz konusu maddelerde yanlılığa neden olan özellik ortadan kaldırılamıyorsa o madde ya da maddeler testlerden çıkarılmalıdır.

Dile ve kültüre göre DMF gösteren maddelerin uzmanlar tarafından değerlendirilmesi bunların olası nedenlerinin tespit edilmesi ve benzer maddelerin ileriki yıllarda hazırlanırken dikkat edilmesi önerilmektedir. Özellikle toplumların kenti yaşam koşulları ve yaşantılarına göre kavramlar ürettiği; ortaya konulan bir olgunun o toplumda görülüyorsa ona karşı gelen bir kavramın da bulunmayacağı unutulmamalıdır.

Olası DMF nedenleri araştırılırken öğrencilerin düşünme süreçleri ve maddeler arası etkileşime odaklanılan çalışmalar yapılabilir. Madde yanıtlanırken öğrenciyi etkileyen

etkenlerin düşünme süreçlerine odaklanılarak çalışılması ve bu sürecin farklı alt gruplarda karşılaştırması DMF'nin olası nedenlerini ortaya çıkarmada işe koşulabilir. Örneğin; bu çalışmanın kapsamı içerisinde yer almayan ancak olası bir yanlışlık sebebi olarak gösterilen soruların test formundaki dizilimi ve form sayısı ile yanlışlık arasındaki ilişki başka bir çalışmanın konusu olabilir.

Maddeler yazılırken kültürler arası farklılıklar dikkate alınmalıdır. Toplumlar mevcut kavramları ile görüş geliştireceğinden özellikle madde yazımında kavramların yüzeysel anlamları ve derin anlamları üzerine çalışmalar yapmak yararlı olacaktır. Yapılan DMF belirleme ve yanlışlık çalışmaları dikkate alınarak belli türde soruların yanlışlık eğilimi taşıyıp taşımayacağı üzerine çalışmalar yapılabilir.

Bu çalışma PISA testi fenle ilgili olan çoktan seçmeli sorularla sınırlı tutulmuştur. PISA da çok çeşitli soru türleri olup; farklı soru türlerini kapsayan çalışmalar yapılabilir.

PISA 2006 ulusal raporunun son kısmında yapılan karşılaştırmalı çalışmalardan elde edilen sonuçlar incelendiğinde; fen bilimleri performansı için çeşitli değişkenler bakımından düzenlenerek hesaplanmış puanların farklılaştığı görülmektedir. Puanlar üzerinde belirli değişkenler etrafında yapılan düzeltmeler sonrası, ülkelerin belirlenen puanları ve belirlenen sıralamalarında önemli farklılıklar ortaya çıkmıştır. Bu nedenle farklı özellikler (değişkenler) göz önüne alınarak DMF çalışmaları bu değişkenler etrafında yapılabilir. Böylelikle geniş ölçekli ölçme araçlarına yönelik birden çok DMF çalışmasından elde edilen bulgular daha doğru bir biçimde genellenerek DMF kaynakları daha net ortaya çıkarılabilir. Bu da uluslararası karşılaştırmalarda kullanılacak testlerin geliştirilmesinde test geliştiricilere yol gösterecektir.

Lojistik regresyon ve alan indeksleri tekniklerinde DMF yorumlanırken farklı ölçütlerin, farklı etki büyüklüğü sınıflandırmalarının kullanılabileceği bilinmektedir. DMF çalışmalarında DMF yorumlanırken kullanılan ölçütlerin belirlenmesine ya da karşılaştırılmasına yönelik yapılacak çalışmaların elde edilen istatistiksel sonuçların yorumlanmasını kolaylaştıracağı düşünüldüğünden farklı örneklemelerde, simülasyon verilerinde ölçüt çalışmaları yapılabilir. DMF analiz yöntemleri için genel olarak

kullanılabilecek etki büyüklüğü sınıflamalarının oluşturulması yöntemlerin karşılaştırılmasındaki hataları azaltacaktır.

KAYNAKÇA

- Abedalaziz N.,2011, Detecting DIF using Item Characteristic Curve Approaches.The International Journal of Educational and Psychological Assessment , Vol. 8(2)
- Abedi J. (2008). Classification System for English Language Learners: Issues and Recommendations Educational Measurement: Issues and Practice; 27(3), 17-31.
- Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, 33, 4-14.
- Acar, T.(2008) Maddenin Farklı Fonksiyonlaşmasını Belirlemede Kullanılan Genelleştirilmiş Aşamalı Doğrusal Modelleme, Lojistik Regresyon ve Olabilirlik Oranı Tekniklerinin Karşılaştırılması, Doktora Tezi, Eğitim Bilimleri Bölümü.
- Adams, R.J., ve Rowe, K.J., (1988). Test Bias In Educational Teearch, Methodology, and measurement: an International Handbook (Eds. J. Keeves), Oxford: Pergamon Pres p. 398-403.
- Adedoyin O.O., 2010. Using IRT approach to detect gender biased items in public examinations: A case study from the Botswana junior certificate examination in Mathematics. *Educational Research and Reviews* Vol. 5 (7), pp. 385-399.
- Angoff, W.H., ve Sharon, A.T. (1974). The evaluation of differences in test performance of two or more groups. *Educational and Psychological Measurement* , 34, 807-816.
- Asil, M. (2010). Uluslararası Öğrenci Degerlendirme Programı (PISA) 2006 öğrenci anketinin kültürler arası esdeğerliğinin incelenmesi. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi, Ankara.
- Atalay K. (2010) PISA 2006 Öğrenci Anketinde Yer Alan Tutum Maddelerinin Değişen Madde Fonksiyonu Açısından İncelenmesi. Yüksek Lisans Tezi Eğitim Bilimleri Bölümü, Hacettepe Üniversitesi, Ankara.

- Atalay K. (2010) PISA 2006 Öğrenci Anketinde Yer Alan Tutum Maddelerinin Değişen Madde Fonksiyonu Açısından İncelenmesi. Yüksek Lisans Tezi Eğitim Bilimleri Bölümü, Hacettepe Üniversitesi, Ankara.
- Atalay K., Gök B., Kelecioğlu H., Arsan N. (2012). Değişen madde fonksiyonunun belirlenmesinde kullanılan farkı yöntemlerin karşılaştırılması: bir simülasyon çalışması, Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 43, 270-281.
- Ayala, R. J. (2009). *The theory and practice of item response theory*. New York: The Guilford Press.
- Bakan Kalaycıoğlu, D. ve Kelecioğlu, H. (2011). Öğrenci Seçme Sınavı'nın Madde Yanlılığı Açısından incelenmesi. *Eğitim ve Bilim*, 36, 3-13.
- Bakan Kalaycıoğlu, D.(2008) Öğrenci Seçme Sınavı'nın Madde Yanlılığı Açısından İncelenmesi, Doktora Tezi, Eğitim Bilimleri Bölümü, Hacettepe Üniversitesi.
- Baker, Frank. (2001). *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation
- Beaton, A.E. (1998). Comparing Cross-National Student Performance On TIMSS Using Different Test Items. *International Journal of Educational Research* 29, 529-542
- Berberoğlu, G. (1995). Differential Item Functioning Analysis of Computation, Word Problem and Geometry Questions Across Gender and SES Groups. *Studies in Educational Evaluation*, 21, 439-456.
- Berberoğlu, G. (1996). The University Entrance Examinations in Turkey. *Studies in Educational Evaluation* Vol. 22, No. 4, pp. 363-373.
- Brislin, R. W., Lonner, W., ve Thorndike, R. M. (1973). *Cross-cultural research methods.*, New York: Wiley.

- Camilli, G. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 221-256). Westport: American Council on Education ve Praeger Publishers.
- Camilli, G., ve Shepard, L.A. (1994). *Methods for identifying biased test items*. Thousand Oaks: Sage.
- Cathy Huaqing Qi, Scott C. Marley.(2009). Differential Item Functioning Analysis of the Preschool Language Scale-4 Between English-Speaking Hispanic and European American Children From Low-Income Families Families, *Topics in Early Childhood. Special Education*, Volume 29 Number 3,171-180
- Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate behavioral Research*, 1, 245-267.
- Cattell, R.B. (1978). *The scientific use of factor analysis*. New York: Plenum.
- Childs,R.A.,(1990). Gender Bias and Fairness. Washington, Clearinghouse on Tests Measurement and Evaluation., ERIC Identifier: ED328610
- Cole, N.S., and Moss, P.A. (1989). Bias in Test Use. In *Educational Measurement* (3rd. eds)(Ed.. Robert L. Linn), 201-218. Collier MacMillian: London
- Crocker, L. ve Algina J. (1986). *Introduction To Classical and Modern Test Theory*. Orlando: Harcourt Brace Jovanovich Inc.
- Cronbach, L. J. (1990). *Essentials of psychological testing* (5. Edition). New York: HarperveRow Publishers.
- Çet, S. (2006). *A Multivariate Analysis In Detecting Differentially Functioning Items Through The Use Of Programme For International Student Assessment (PISA) 2003 Mathematics Literacy Items*, Doktora Tezi
- Deveci Ateşok, N. (2008) *Üniversitelerarası Kurul Yabancı Dil Sınavının Madde Yanlılığı Bakımından İncelenmesi; Doktora Tezi, Ölçme ve Değerlendirme Anabilim Dalı, Ankara Üniversitesi*

- Doğan N., Öğretmen T. (2008). Değişen Madde Fonksiyonunu Belirlemede Mantel-Haenszel, Ki-Kare ve Lojistik Regresyon Tekniklerinin Karşılaştırılması. *Eğitim ve Bilim*, Cilt 33, Sayı 148
- Doğan, N. ve Öğretmen, T. (2006). Madde Yanlılığını Belirleme Teknikleri Arasında Bir Karşılaştırma. *Eurasian Journal of Educational Research*, 23, 94-105.
- Dorans, N. J., ve Holland, P. W. (1992). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland, ve H. Wainer, (Eds.), *Differential Item Functioning* (pp. 35-66). New Jersey: USA.
- Embretson, S. E. (1996). Item response theory models and inferential bias in multiple group comparisons. *Applied Psychological Measurement*, 20, 201-212.
- Engelhard, G. (1990). Gender Differences In Performance On Mathematics Items: Evidence From The United States And Thailand. *Contemporary Educational Psychology*, 15, 13-26.
- Ercikan, K., (1998), Translation Effects In International Assessments, *International Journal of Educational Research*, 29, 543-553.
- Ercikan, K., Arim, R., Law, D., Domene, J., ve Lacroix, S. (2010). Application of Think Aloud Protocols for examining and confirming sources of Differential Item Functioning identified by expert reviews. *Educational Measurement: Issues and Practice*, 29, 24-35.
- Ercikan, K., Gierl, M., McCreith, T., Puhon, G., ve Koh, K. (June, 2002). Comparability of English and French versions of SAIP for reading, mathematics and science items. Paper presented at the annual meeting of the Canadian Society for Studies in Education, Toronto, ON.
- Giri M.J. ve Khaliq S.N.. (2001) Identifying Sources of Differential Item and Bundle Functioning on Translated Achievement Tests: A Confirmatory Analysis *Journal of Educational Measurement*; Vol. 38, No. 2,; pp. 164-187

- Gondal, M., B.(2001). *Pakistan`daki 4. Sınıf Ve Urdu (Ulusal Dil) Başarı Test Maddelerinin Madde Yanlılığı Açısından İncelenmesi*. Yayınlanmamış Doktora Tezi, Orta Doğu Teknik Üniversitesi, Eğitim Bilim Bölümü, Ankara.
- Gök B., Kelecioğlu H., Doğan N. (2010), Değişen Madde Fonksiyonunu Belirlemede Mantel–Haenszel Ve Lojistik Regresyon Tekniklerinin Karşılaştırılması, Eğitim Ve Bilim, Cilt 35, Sayı 156, S3-16.
- Grisay, A. (2003). Translation Procedures in OECD/PISA 2000 International Assessment. *Language Testing*, 20(2), 225-240
- Hambleton, R K., Swaminathan, H. ve Rogers, H. J. (1991). Fundamentals of item response theory. London: Sage Publication.
- Hambleton, R. K. ve Swaminathan, H. Item Response Theory: Principles and Applications. Boston: Kluwer-Nijhoff Publishing, 1985.
- Hambleton, R. K., ve Jones, R. W. (1994). Comparison of empirical and judgmental procedures for detecting differential item functioning. *Educational Research Quarterly*, 18, 21-36.
- Hambleton, R.K. and Rogers, H., J.(1996). Developing an Item Bias Review Form. Which is Available Through ERIC/AE.
- Hambleton, R.K. ve Jones, R.W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*.12(3), 38-47.
- Hambleton, R.K. ve Swaminathan, H. (1985). Item Response Theory: Principles and Application. Boston: Kluwer Academic Publishers Group.
- Hambleton, R.K., Swaminathan, H. ve Rogers, H.J. (1991). *Fundamentals of Item Response theory*. California: Sage Publications Inc.
- Harold J. (1989). The determination of cultural item bias in the California Achievement Tests Arizona Üniversitesi; Doktora tezi.

- Harris, A. M. ve Carlton, S. T. (1993). Patterns of Gender Differences on Mathematics Items on the Scholastic Aptitude Test. *Applied Measurement in Education*, 6, 137 – 151.
- Hidalgo M. D., Lopez-Pina J. (2004). Differential Item Functioning Detection And Effect Size A Comparison Between Logistic Regression And Mantel-Haenszel Procedures. *Educational and Psychological Measurement* 64: 903
- Holland, P. W., ve Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer ve H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ: Erlbaum.
- Idaszak, J.R., Bottom, W.P. ve Drasgow, F. (1987). ‘A test of the measurement equivalence of the revised Job Diagnostic Survey’, *Journal of Applied Psychology*, 73(4): 69-74.
- Kan A.. (2007) Test Yansızlığı: H.Ü. Yabancı Dil Muafiyet Sınavının Cinsiyete ve Bölümlere Göre DMF Analizi, *Eurasian Journal of Educational Research*, 29, pp, 45-
- Karakaya İ., Kutlu Ö. (2012). Seviye Belirleme Sınavındaki Türkçe Alt Testlerinin Madde Yanlılığının İncelenmesi; *Eğitim ve Bilim*; Cilt 37; sayı 165
- Karami H., Nodoushan M.A.(2011). Differential Item Functioning (DIF): Current problems and future directions *International Journal of Language Studies (IJLS)*, Vol. 5(3), s. 141
- Klime, E. ve Baumert, J. (2001). Identifying National Cultures of Mathematics Education: Analysis of Cognitive Demands and Differential Item Functioning In TIMSS. *European Journal of Psychology of Education*, 15(3), 385 – 402.
- Koh, K., ve Ercikan, K. (June, 2002). Construct comparability in French and English versions of SAIP. Paper presented at the annual meeting of the Canadian Society for Studies in Education Meeting, Toronto, ON.

- Kristjansson , E., Aylesworth,R., Mcdowell,I., and Zumbo,D A (1997). Comparison of Four Methods for Detecting Differential Item Functioning in Ordered Response Items, ERIC/AE.
- Kurnaz F. B. (2006) Peabody Resim Kelime Testinin Madde Yanlılığı Açısından İncelenmesi, Yüksek Lisans Tezi, Hacettepe Üniversitesi; Eğitim Bilimleri Bölümü; Ankara.
- Loehlin, J.C. (1987). Latent variable models. New Jersey: Lawrence Erlbaum Associates.
- Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillsdale, NJ: Lawrence Erlbaum Publishers.
- Lord, M.F., ve Novic, R. M. (1968). Statistical theories of mental test scores. New York: Addison-Wesley Publishing Company.
- Marin, G., ve VanOss Marin, B. (1991). Research with Hispanic populations. Newbury Park: Sage Publications
- Mark J. Gierl, Shameem NyJa Khaliq, 2001 Identifying Sources of Differential Item and Bundle Functioning on Translated Achievement Tests: A Confirmatory Analysis *Journal of Educational Measurement* , Vol. 38, No. 2, pp. 164-187
- Mellenberg, G. J. (1983). Conditional Item Bias Methods. In S. H. Irvine and W. J. Barry (Eds), Human Assesment and Cultural Factors (pp. 293-302). Newyork: Plenum Press.
- Mellenberg, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research: Applications of Item Response Theory*.13, 123–14
- Millsap, R.E. (2005). Four unresolved problems in studies of factorial invariance. In A. Maydeu-Olivares ve J.J. McArdle (Eds.) Contemporary Psychometrics. (pp. 153-171) Mahwah, NJ: Lawrence Erlbaum.

- Millsap, R.E. ve Meredith, W. (1992). Component analysis in multivariate aging research. *Experimental Aging Research*, 18, 203-212
- Öğretmen T. (1995). DIF Analysis Of The Verbal Ability Section of the First Stage of the University Entrance Examination in Turkey. Yüksek Lisans Tezi. Eğitim Bilimleri Bölümü, Orta Doğu Teknik Üniversitesi, Ankara.
- Öğretmen, T. (2006). Uluslararası Okuma Becerilerinde Gelişim Projesi (PIRLS) Testinin psikometrik Özelliklerinin İncelenmesi: Türkiye- Amerika Birleşik 2001 Devletleri Örneği. Doktora Tezi. Eğitim Bilimleri Bölümü, Hacettepe Üniversitesi.
- Öğretmen, T., ve Doğan, N. (2004). OKÖSYS Matematik alt testine ait maddelerin yanlışlık analizi. İnönü Üniversitesi Eğitim Fakültesi Dergisi. 8, 61-76.
- Pae, T. II (2004). DIF For Examinees With Different Academic Backgrounds. *Language Testing*, Vol. 21, No. 1, 53-73
- Pae, T. II. and Park, P.G.(2006). Examining The Relationship Between Differential İtem Functioning and Differential Test Functioning. web: <http://ltj.sagepub.com/cgi/content/abstract/23/4/475>
- PISA 2006 Uluslararası Öğrenci Başarılarını Değerlendirme Programı Ulusal Ön Rapor (2007) <http://earged.meb.gov.tr/dosyalar> adresinden indirilmiştir.
- Raju, N. S. (1988). The area between two item characteristic curves. *Psychometrika*, 53, 495–502.
- Raju, N. S., Laffitte, L. J., ve Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology*, 87, 517–529.
- Raju, N. S., ve Ellis, B. B. (2002). Differential item and test functioning. In F. Drasgow ve N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations* (p. 156-188). San Francisco, CA: Jossey-Bass.

- Ross, J. S. and Okabe, J. (2006). The Subjective and Objective Interface of Bias Detection on Language Tests. *International Journal of Testing*, Vol. 6, No. 3, Pages 229-253
- Samejima, F. (1977). A use of the information function in tailored testing. *Applied Psychological Measurement*, 1, 233-247.
- Shepard, L.A., Camilli, G. ve Williams, D.M. (1985). Validity of approximation techniques for detecting item bias. *Journal of Educational Measurement*, 22, 77-105
- Solano-Flores, G., ve Li, M. (2006). The use of generalizability (G) theory in the testing of linguistic minorities. *Educational Measurement: Issues and Practice*, 25(1), 13-22.
- Swaminathan, H., ve Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- Tittle, C.K. (1988). Test Bias In Educational Teearch, Methodology, and measurement: an International Handbook (Eds. J. Keeves), 392-398.Pergamon Press: UK.
- Vijver, F.J.R. van de (1997). Bias research: Introduction. *European Review of Applied Psychology*, 47(4), 261-262.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Welkenhuysen-Gybels .J. ve Billiet J. (2002). A Comparison of Techniques for Detecting Cross-Cultural Inequivalence at the Item Level; Quality ve Quantity 36: 197–218.

- Yenal, E. (1995). Differential Item Functioning Analysis of the Quantitative Ability Section of the First Stage of the University Entrance Examination in Turkey. Yayınlanmamış Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi, Ankara.
- Yıldırım, H.(2006). The Differential Item Functioning (DIF) Analysis Of Mathematics Items in The International Assessment Programs. Yayınlanmamış Doktora Tezi. Orta Doğu Teknik Üniversitesi, Matematik Eğitim Bölümü. Ankara.
- Yurdugül, H. ve Aşkar, P. (2004) Ortaöğretim Kurumları Öğrenci Seçme Ve Yerleştirme Sınavının Öğrencilerin Yerleşim Yerlerine Göre Diferansiyel Madde Fonksiyonu Açısından İncelenmesi. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 27, 268-275.
- Zieky, M. (2003). *A DIF Primer*. Princeton, NJ: Educational Testing Service.
http://www.ets.org/Media/Tests/PRAXIS/pdf/DIF_primer.pdf

EKLER

EK 1: Uzman Kanısı Formu 1

Sayın Hocam,

PISA 2006 sınavında yer alan fen bilimleri çoktan seçmeli soruların kùltùrlere ve dile göre yanlı olup olmadıklarını belirlemek amacıyla bir doktora tezi araştırması yürütmekteyim. Bu amaçla 2006 yılında uygulanan PISA 1. ve 5. Kitapçıklarda yer alan (AÇIKLANAN) maddelerle çalışmaktayım. Bu maddelerin yanlılığa sahip olup olmadığını belirlemek ve bunun olası nedenlerine ilişkin çıkarımlar yapabilmek için görüşlerinize ihtiyaç duymaktayım.

Yanlılık, bir maddenin farklı gruplardan gelen, benzer yetenek düzeyindeki bireyler tarafından doğru yanıtlanma olasılığının farklılık göstermesidir. Başka bir ifade ile testte yer alan maddelerin aynı yetenek grubundaki öğrenciler için avantaj ya da dezavantaj sağlama durumudur. Madde yanlılığı; farklı dillerdeki ölçme araçlarının çevirisinden, maddelerin kùltüre ve cinsiyete özgü içeriklerinden...vs kaynaklanabilir.

Aşağıda 2006 PISA sınavına ait açıklanan okuma parçaları ve onlara ait çoktan seçmeli maddelerin Türkçe ve İngilizce formları yer almaktadır. Öncelikle her iki forma ilişkin okuma parçaları ve onlara ait soruları okumanız gerekmektedir. Formu doldururken ise ilk aşamada maddenin yanlı olup olmadığına ilişkin yanıtınızı vermeniz beklenmektedir. Eğer yanıtınız evet ise; yanlılığın nedenlerinin uzman görüşleri ile açıklanması çalışma için oldukça önemlidir. Bu nedenleri her bir soru için belirttiniz.

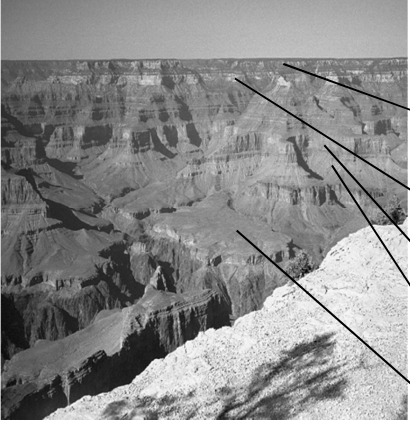
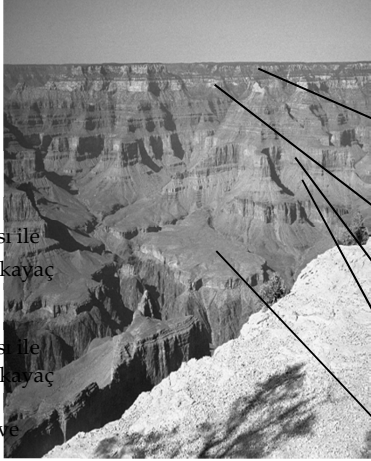
Araştırmaya vermiş olduğunuz değerli katkılarınızdan dolayı çok teşekkür ederim.

N. Bilge Başusta
Eğitimde Ölçme ve Değerlendirme A.D
Doktora Öğrencisi

S O R U 1	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 2	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 3	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 4	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 5	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?

S O R U 6	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 7	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 8	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 9	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?
S O R U 10	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; soru hangi grup için daha kolaydır? Türkçe Form () İngilizce Form () ▪ Bunun nedenleri hakkındaki görüşleriniz nelerdir?

GİYSİLER Bir grup İngiliz bilim adamı, konuşma engelli çocuklara ‘konuşma’ gücü verecek ‘akıllı’ giysiler üretiyor. Benzeri olmayan bir elektro tekstil ürününden yapılan ve ses üreten bir aygıtla bağlanmış yelek giyen çocuklar, dokunmaya duyarlı kumaşa hafifçe vurarak konuşmalarının başkaları tarafından anlaşılabilir duruma gelmesini sağlamaktadırlar. Bu kumaş, normal kumaş ve içine kusursuz bir şekilde yerleştirilmiş karbon iplikçikler sayesinde elektriği iletebilen bir fileden yapılmıştır. Kumaş üzerine basınç uygulandığında, iletken iplikçiklerden geçen sinyaller değiştirilir ve bir bilgisayar devresi kumaşa nerede dokunulduğunu belirler. Daha sonra, bu devre kendisine bağlı olan ve iki kibrit kutusundan daha büyük olmayan bir elektronik aracın tetiklemektedir. Bilim adamlarından birisi şöyle söylemektedir: “İşin en çarpıcı kısmı, kumaşı nasıl dokuduğumuz ve sinyalleri onun içinden nasıl gönderdiğimizdir - onu normal bir kumaşta var olan dokunuş şekli içerisine, kimsenin göremeyeceği şekilde yerleştirebiliriz.” Bu kumaş, zarar görmeksizin yıkanabilir, nesnelerin etrafına sarılabilir ya da sıkılıp top durumuna getirilebilir. Bilim adamları, onun toptan üretiminin ucuz olacağını da ileri sürmektedirler.	CLOTHES A team of British scientists is developing “intelligent” clothes that will give disabled children the power of “speech”. Children wearing waistcoats made of a unique electrotexile, linked to a speech synthesiser, will be able to make themselves understood simply by tapping on the touch-sensitive material. The material is made up of normal cloth and an ingenious mesh of carbonimpregnated fibres that can conduct electricity. When pressure is applied to the fabric, the pattern of signals that passes through the conducting fibres is altered and a computer chip can work out where the cloth has been touched. It then can trigger whatever electronic device is attached to it, which could be no bigger than two boxes of matches. “The smart bit is in how we weave the fabric and how we send signals through it – and we can weave it into existing fabric designs so you cannot see it’s in there,” says one of the scientists. Without being damaged, the material can be washed, wrapped around objects or scrunched up. The scientist also claims it can be mass-produced cheaply.
Soru 1: GİYSİLER Aşağıdaki laboratuvar araçlarından hangisi kumaşın elektriği ilettiğini deneyebilmemiz için gerekecek araçlar arasında yer alabilir? A. Voltmetre B. Işık kutusu C. Mikrometre D. Ses ölçer	Question1: CLOTHES Which piece of laboratory equipment would be among the equipment you would need to check that the fabric is conducting electricity? A. Voltmeter B. Light box C. Micrometer D. Sound meter

GRAND KANYON (BÜYÜK KANYON) Grand Canyon (Büyük Kanyon) Amerika Birleşik Devletleri'ndeki bir çöldedir. Burası, birçok kaya katmanını içeren çok geniş ve derin bir kanyondur. Geçmiş bir zaman diliminde yerkabuğunda meydana gelen hareketler bu katmanları yukarıya doğru itmiştir. Günümüzde bu kanyonun bazı bölümleri 1.6 km derinliğindedir. Kanyonun dibinde Colorado Nehri akmaktadır. Aşağıda Büyük Kanyon' un güney kenarından çekilmiş bir resmi görülmektedir. Kanyon 'un bu resminde birkaç değişik kaya tabakası görülebilmektedir.  Kireçtaşı (A) Kilin sıkışması ile oluşan tortul kayaç Kireçtaşı (B) Kilin sıkışması ile oluşan tortul kayaç Tortul kayaç ve granit	THE GRAND CANYON The Grand Canyon is located in a desert in the USA. It is a very large and deep canyon containing many layers of rock. Sometime in the past, movements in the Earth's crust lifted these layers up. The Grand Canyon is now 1.6 km deep in parts. The Colorado River runs through the bottom of the canyon. See the picture below of the Grand Canyon taken from its south rim. Several different layers of rock can be seen in the walls of the canyon.  Limestone (A) Shale (A) Limestone (B) Shale (B) Schists and granite
Soru 2: Büyük Kanyon' da hava sıcaklığı 0 °C 'ın altındaki sıcaklıklardan 40 °C'ın üstündeki sıcaklıklara kadar değişebilmektedir. Burası bir çöl alanı olmasına karşın, kayalardaki çatlaklarda bazen su bulunabilmektedir. Bu sıcaklık değişimleri ve çatlaklardaki su kayaların parçalanmasını nasıl hızlandırabilmektedir? A. Donan su, sıcak kayaları eritir. B. Su, kayaları birbirine yapıştırır. C. Buz kayaların yüzeyini düzleştirir. D. Kaya çatlaklarında donan su genişir.	Question 2 The temperature in the Grand Canyon ranges from below 0 oC to over 40 oC. Although it is a desert area, cracks in the rocks sometimes contain water. How do these temperature changes and the water in rock cracks help to speed up the breakdown of rocks? A. Freezing water dissolves warm rocks. B. Water cements rocks together. C. Ice smoothes the surface of rocks. D. Freezing water expands in the rock cracks.
Soru 3: Büyük Kanyon'un "Kireçtaşı (A)" olarak belirtilen tabakasında deniz tarağı, balık ve mercan gibi birçok deniz hayvanının fosilleri bulunmaktadır. Bu fosillerin orada bulunabilmeleri için milyonlarca yıl önce ne olmuştur? A. Eski zamanlarda insanlar okyanustan oraya su ürünleri getirmişlerdir. B. Bir zamanlar okyanuslarda büyük dalgalar oluştu ve bunlar deniz yaşamını karalara sürükledi. C. O zamanlarda okyanus buraları kaplamıştı, sonra sular eski yerine çekildi. D. Bazı deniz hayvanları, denize göç etmeden önce bir süre karada yaşadılar.	Question3 There are many fossils of marine animals, such as clams, fish and corals, in the Limestone A layer of the Grand Canyon. What happened millions of years ago that explains why such fossils are found there? A. In ancient times, people brought seafood to the area from the ocean. B. Oceans were once much rougher and sea life washed inland on giant waves. C. An ocean covered this area at that time and then receded later. D. Some sea animals once lived on land before migrating to the sea

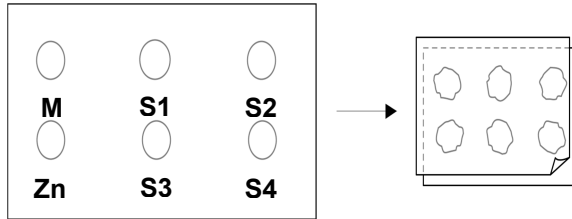
GÜNEŞTEN KORUYUCULAR

Jale ve Osman, güneşten koruma ürünlerinden hangisinin ciltleri için en iyi korumayı sağladığını merak ettiler. Güneşten koruma ürünleri için, her ürünün güneş ışığındaki ultraviyole ışınlarını ne derecede emdiğini gösteren bir *Güneşten Koruma Faktörü (GKF)* tanımlanmıştır. GKF'si yüksek olan bir güneşten koruyucu, GKF'si düşük olan bir güneşten koruyucuya göre cildi daha uzun süre korur.

Jale, bazı güneşten koruma ürünlerini birbiriyle karşılaştırmak için bir yol düşündü. Osman ile birlikte aşağıdaki malzemeleri topladılar:

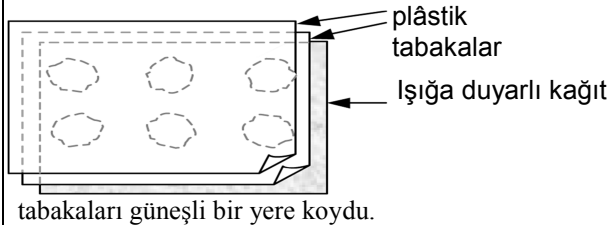
- güneş ışığını emmeyen (geçiren) iki temiz plastik tabaka;
- bir adet ışığa duyarlı kağıt;
- mineral yağ (M) ve çinko oksit (ZnO) içeren bir krem
- S1, S2, S3 ve S4 adını verdikleri dört farklı güneşten koruma ürünü.

Jale ve Osman, mineral yağı güneş ışınlarının çok büyük bir kısmını geçirdiği için, çinko oksidi de güneş ışınlarının tamamına yakınına geçirmediği için seçtiler. Osman, bir plastik tabaka üzerinde yuvarlak içine alınmış yerlerin her birine her maddeden birer damla koydu sonra bunların üzerini ikinci bir plastik tabaka ile kapattı. Bu plastik tabakaların üzerine büyük bir



kitap yerleştirerek üstten iyice bastırdı.

Daha sonra, Jale hazırladıkları plastik tabakaları ışığa duyarlı kâğıdın üzerine koydu. Işığa duyarlı kâğıt, güneş ışığında tutulduğu süreye göre koyu griden beyaza (ya da çok açık griye) doğru renk değiştiren bir kâğıttır. En sonunda da, Osman hazırladıkları bu

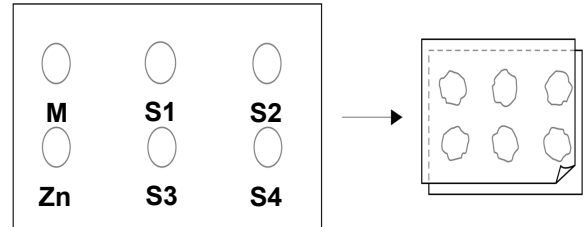


SUNSCREENS

Mimi and Dean wondered which sunscreen product provides the best protection for their skin. Sunscreen products have a *Sun Protection Factor (SPF)* that shows how well each product absorbs the ultraviolet radiation component of sunlight. A high SPF sunscreen protects skin for longer than a low SPF sunscreen.

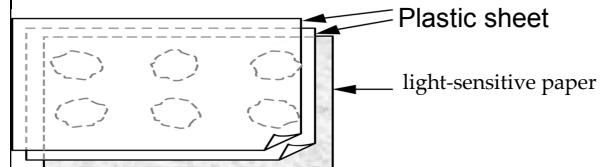
Mimi thought of a way to compare some different sunscreen products. She and Dean collected the following:

- two sheets of clear plastic that do not absorb sunlight;
- one sheet of light-sensitive paper;
- mineral oil (M) and a cream containing zinc oxide (ZnO); and
- four different sunscreens that they called S1, S2, S3, and S4.



Mimi and Dean included mineral oil because it lets most of the sunlight through, and zinc oxide because it almost completely blocks sunlight.

Dean placed a drop of each substance inside a circle marked on one sheet of plastic, then put the second



plastic sheet over the top. He placed a large book on top of both sheets and pressed down.

Mimi then put the plastic sheets on top of the sheet of light-sensitive paper. Light-sensitive paper changes from dark grey to white (or very light grey), depending on how long it is exposed to sunlight. Finally, Dean placed the sheets in a sunny place.

Soru 4: Aşağıdaki ifadelerden hangisi, güneşten koruyucuların etkililiğini karşılaştırma amacıyla yapılan bir çalışmada mineral yağ ve çinko oksidin rolünün bilimsel tanımıdır? A. Mineral yağ ve çinko oksidin ikisi de etkisi araştırılan birer etkidir. B. Mineral yağ test edilen bir etken, çinko oksit ise karşılaştırma için kullanılan bir maddedir. C. Mineral yağ karşılaştırma için kullanılan bir madde, çinko oksit ise test edilen bir etkidir. D. Mineral yağ ve çinko oksidin ikisi de karşılaştırma için kullanılan birer maddedir.	Question4 Which one of these statements is a scientific description of the role of the mineral oil and the zinc oxide in comparing the effectiveness of the sunscreens? A. Mineral oil and zinc oxide are both factors being tested. B. Mineral oil is a factor being tested and zinc oxide is a reference substance. C. Mineral oil is a reference substance and zinc oxide is a factor being tested. D. Mineral oil and zinc oxide are both reference substances.
Soru 5: Jale ve Osman'ın yanıtlamaya çalıştığı soru aşağıdakilerden hangisidir? A. Güneşten koruyucu maddelerden her birinin koruma gücü diğerlerine kıyasla nasıldır? B. Güneşten koruyucular cildi ultraviyole ışınlarından nasıl korur? C. Mineral yağdan daha az koruma sağlayan bir güneşten koruyucu var mıdır? D. Çinko oksitten daha çok koruma sağlayan bir güneşten koruyucu var mıdır?	Question 5 Which one of these questions were Mimi and Dean trying to answer? A. How does the protection for each sunscreen compare with the others? B. How do sunscreens protect your skin from ultraviolet radiation? C. Is there any sunscreen that gives less protection than mineral oil? D. Is there any sunscreen that gives more protection than zinc oxide?
Soru 6: İkinci plastik tabakanın üzerine neden iyice bastırılmıştır? A. Damlaların kurummasını önlemek için B. Damlaları mümkün olduğunca yaymak için C. Damlaları yuvarlaklar içinde tutmak için D. Damlalara eşit kalınlık vermek için	Question 6 Why was the second sheet of plastic pressed down? A. To stop the drops from drying out. B. To spread the drops out as far as possible. C. To keep the drops inside the marked circles. D. To make the drops the same thickness.

AŞININ TARİHÇESİ Mary Montagu güzel bir kadındı. 1715 yılında çiçek hastalığına yakalandı. Hastalığı geçirdi; fakat izleri kaldı. 1717 yılında Türkiye'de yaşarken, bu ülkede yaygınca kullanılmakta olan ve adına aşılama denen bir tedaviyi gördü. Bu tedavide sağlıklı gencin derisi çizilerek ona zayıflatılmış çiçek virüsü veriliyordu. Kişi kısa bir süre için hasta oluyor, ancak hastalığı genellikle çok hafif bir şekilde geçiyordu. Mary, bu aşılama yönteminin güvenli olduğuna inandı ve kendi oğlu ile kızının da bu şekilde aşılmasına izin verdi. 1796 yılında Edward Jenner çiçek hastalığına karşı antikor geliştirmek için insandaki çiçek hastalığı virüsünü değil, ineklerde görülen çiçek hastalığı virüsünü kullanarak aşılama yöntemini geliştirdi. Jenner'in bulduğu bu aşılama yönteminin, çiçek hastalığı virüsü verilmesine kıyasla, yan etkileri daha azdır ve tedavi gören kişi virüsü başka insanlara bulaştıramaz. Bu tedâvi biçimi aşılama adıyla tanındı.	The History of Vaccination Mary Montagu was a beautiful woman. She survived an attack of smallpox in 1715 but she was left covered with scars. While living in Turkey in 1717, she observed a method called inoculation that was commonly used there. This treatment involved scratching a weak type of smallpox virus into the skin of healthy young people who then became sick, but in most cases only with a mild form of the disease. Mary Montagu was so convinced of the safety of these inoculations that she allowed her son and daughter to be inoculated. In 1796, Edward Jenner used inoculations of a related disease, cowpox, to produce antibodies against smallpox. Compared with the inoculation of smallpox, this treatment had less side effects and the treated person could not infect others. The treatment became known as vaccination.
Soru 7 İnsanlar hangi çeşit hastalıklara karşı aşılanabilir? A. Hemofili gibi kalıtsal hastalıklar B. Çocuk felci gibi virüslerin neden olduğu hastalıklar C. Şeker hastalığı gibi vücudun işlevsel bozukluklarından kaynaklanan hastalıklar D. Tedavisi olmayan her çeşit hastalık	Question 7 What kinds of diseases can people be vaccinated against? A. Inherited diseases like haemophilia. B. Diseases that are caused by viruses, like polio. C. Diseases from the malfunctioning of the body, like diabetes. D. Any sort of disease that has no cure.
Soru 8 Hayvanlar ya da insanlar bakterilerin neden olduğu bulaşıcı bir hastalığa yakalanır ve iyileşirse, hastalığa neden olan bakteriler genellikle onlarda tekrar hastalık oluşturmaz. Bunun nedeni aşağıdakilerden hangisidir? A. Vücudun, aynı çeşitten bir hastalığa neden olabilecek bütün bakterileri öldürmüş olması B. Vücudun, bu tür bakterileri çoğalmadan önce öldürecek antikorlar yapmış olması C. Alyuvarların, aynı çeşit hastalığa neden olabilecek bütün bakterileri öldürmesi D. Alyuvarların, vücuttaki bu tip bakterileri yakalayarak vücuttan atması.	Question 8 If animals or humans become sick with an infectious bacterial disease and then recover, the type of bacteria that caused the disease does not usually make them sick again. What is the reason for this? A. The body has killed all bacteria that may cause the same kind of disease. B. The body has made antibodies that kill this type of bacteria before they multiply. C. The red blood cells kill all bacteria that may cause the same kind of disease. D. The red blood cells capture and get rid of this type of bacteria from the body.

ASİT YAĞMURU

Aşağıda, Caryatids adı verilen ve Atina Akropolünde 2500 yıl önce inşa edilmiş olan heykellerin fotoğrafı görülmektedir. Heykeller, mermer adı verilen bir cins kayadan yapılmıştır. Mermer kireçtaşından (kalsiyum karbonattan) oluşmaktadır.

Orijinal heykeller 1980 yılında kopyalarıyla değiştirilerek Akropol müzesinin içine alındı. Bu



heykeller asit yağmurundan zarar görmüşlerdi.

ACİD RAIN

Below is a photo of statues called Caryatids that were built on the Acropolis in Athens more than 2500 years ago. The statues are made of a type of rock called marble. Marble is composed of calcium carbonate.

In 1980, the original statues were transferred inside the museum of the Acropolis and were replaced by replicas. The original statues were being eaten away by acid rain.

**Soru 9: ASİT YAĞMURU**

Bir mermer parçasının gece boyunca sirke içine konmadan önceki kütlesi 2,0 gramdır. Sonraki gün bu parça sirkeden çıkarılarak kurutulmuştur. Kurutulmuş olan bu mermer parçasının kütlesi ne kadar olabilir?

- A. 2,0 gramdan daha az
- B. Tam olarak 2,0 gram
- C. 2,0 ile 2,4 gram arasında
- D. 2,4 gramdan fazla

Question 9

A marble chip has a mass of 2.0 grams before being immersed in vinegar overnight. The chip is removed and dried the next day. What will the mass of the dried marble chip be?

- A. Less than 2.0 grams
- B. Exactly 2.0 grams
- C. Between 2.0 and 2.4 grams
- D. More than 2.4 grams

GENETİK YAPISI DEĞİŞTİRİLEN (GYD) MISIR YASAKLANMALIDIR Doğayı koruma grupları, yeni ortaya çıkan genetik yapısı değiştirilmiş (GYD) mısırın yasaklanmasını istemektedirler. GYD mısır, geleneksel mısır bitkilerini öldüren yeni ve güçlü bir zararlı ot ilacından etkilenmeyecek şekilde geliştirilmiştir. Bu yeni zararlı ot ilacı, mısır tarlalarında kullanıldığında büyüyen zararlı otların pek çoğunu öldürecektir. Doğayı koruma yanlısı olanlar, yeni ilacın öldüreceği zararlı otlar küçük hayvanların ve özellikle böceklerin beslenmesine yaradığından, bu yeni zararlı ot ilacının GYD mısır ile birlikte kullanılmasının çevre için kötü olacağını söylemektedirler. GYD mısırın kullanılmasını destekleyenler buna cevap olarak bilimsel bir incelemenin, sonucun bu şekilde olmayacağını gösterdiğini söylemektedirler. Yukarıdaki yazıda sözü edilen bilimsel incelemenin bazı ayrıntıları şunlardır: • Mısır, ülkenin değişik yerlerindeki 200 tarlaya ekilmiştir. • Her tarla önce iki eşit parçaya ayrılmıştır. Tarlanın bir parçasında yeni güçlü zararlı ot ilacı ile ilaçlanmış olan genetik yapısı değiştirilmiş (GYD) mısır yetiştirilmiştir. Tarlanın diğer parçasında da geleneksel zararlı ot ilacı ile ilaçlanmış geleneksel mısır yetiştirilmiştir. • Yeni zararlı ot ilacı ile ilaçlanan GYD mısır içinde bulunan böceklerin sayısı, geleneksel zararlı ot ilacı ile ilaçlanmış olan geleneksel mısır içinde bulunan böceklerin sayısı ile hemen hemen aynıdır.	GM CORN SHOULD BE BANNED Wildlife conservation groups are demanding that a new genetically modified (GM) corn be banned. This GM corn is designed to be unaffected by a powerful new herbicide that kills conventional corn plants. This new herbicide will kill most of the weeds that grow in cornfields. The conservationists say that because these weeds are feed for small animals, especially insects, the use of the new herbicide with the GM corn will be bad for the environment. Supporters of the use of the GM corn say that a scientific study has shown that this will not happen. Here are details of the scientific study mentioned in the above article: • Corn was planted in 200 fields across the country. • Each field was divided into two. The genetically modified (GM) corn treated with the powerful new herbicide was grown in one half, and the conventional corn treated with a conventional herbicide was grown in the other half. • The number of insects found in the GM corn, treated with the new herbicide, was about the same as the number of insects in the conventional corn, treated with the conventional herbicide.
Soru 10 Mısır ülkenin değişik yerlerindeki 200 tarlaya ekilmişti. Bilim adamları niçin birden fazla yerde ekim yapmışlardır? A. Yeni GYD mısırı, birçok çiftçinin deneme fırsatı bulması için B. Ne kadar GYD mısır yetiştirebileceklerini görmeleri için C. GYD mısır ekimini olabildiğince geniş bir alana yaymak için D. Mısırın değişik yetiştirme koşullarda nasıl büyüyeceğini görmek için	Question 10 Corn was planted in 200 fields across the country. Why did the scientists use more than one site? A. So that many farmers could try the new GM corn. B. To see how much GM corn they could grow. C. To cover as much land as possible with the GM crop. D. To include various growth conditions for corn.

EK 2: Uzman Kanısı Formu 2

Sayın Hocam,

PISA 2006 sınavında yer alan fen bilimleri çoktan seçmeli soruların kültürlere ve dile göre yanlı olup olmadıklarını belirlemek amacıyla bir doktora tezi araştırması yürütmekteyim. Bu amaçla 2006 yılında uygulanan PISA 1. ve 5. Kitapçıklarda yer alan (AÇIKLANAN) maddelerle çalışmaktayım. Bu maddelerin yanlılığa sahip olup olmadığını belirlemek ve bunun olası nedenlerine ilişkin çıkarımlar yapabilmek için görüşlerinize ihtiyaç duymaktayım.

Yanlılık, bir maddenin farklı gruplardan gelen, **benzer yetenek düzeyindeki** bireyler tarafından doğru yanıtlanma olasılığının farklılık göstermesidir. Başka bir ifade ile testte yer alan maddelerin **aynı yetenek grubundaki öğrenciler için avantaj ya da dezavantaj sağlama durumudur**. Madde yanlılığı; farklı dillerdeki ölçme araçlarının çevirisinden, maddelerin kültüre ve cinsiyete özgü içeriklerinden...vs kaynaklanabilir.

Aşağıda 2006 PISA sınavına ait açıklanan okuma parçaları ve onlara ait çoktan seçmeli madde/maddeler (sorular) yer almaktadır. Sayfa 9’ da maddelere ilişkin olası yanlılık nedenlerini ortaya koymanıza yardımcı olabilecek bir form bulunmaktadır.

- Formu doldururken ilk aşamada maddenin yanlı olup olmadığına ilişkin yanıtınızı vermeniz beklenmektedir. (Bu aşamada PISA’ nın uluslar arası bir sınav olduğunu, soruların farklı dillere çevrildiğini, hazırlanan ve çevrilen soruların bazı ülkeler – kültürler için uygun olamayabileceği ve Türkiye’ de bu sınava katılanların **ilköğretim 8. Sınıf** öğrencileri olduğunu düşünerek sorunun Türk öğrencilerine avantaj ya da dezavantaj sağlayacağı durumları değerlendirmeniz beklenmektedir.) Bazı okuma parçalarına ait 1’ den fazla madde (soru) bulunmaktadır. Bu maddelere ilişkin değerlendirmenizi belirttiğiniz taktirde birlikte yapabilirsiniz.
- Eğer yanıtınız evet ise (YANLILIK OLABİLİR); yanlılığın nedenlerinin uzman görüşleri ile açıklanması çalışma için oldukça önemlidir.
- Bu nedenleri her bir soru için ayrı ayrı belirttiniz.

Araştırmaya vermiş olduğunuz değerli katkılarınızdan dolayı çok teşekkür ederim.

Bir sonraki sayfada bulunan formu bilgisayar ortamında doldurmanız halinde formu n.bilgeuzun@gmail.com adresine yollayabilirsiniz.

SORU1	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU2	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU3	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU4	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU5	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU6	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU7	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU8	- Yanlılık içermekte midir? EVET () HAYIR () - Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?

SORU9	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?
SORU10	▪ Yanlılık içermekte midir? EVET () HAYIR () ▪ Yanıtınız EVET ise; olası yanlılık sebepleri neler olabilir?

GIYSİLER

Bir grup İngiliz bilim adamı, konuşma engelli çocuklara ‘konuşma’ gücü verecek ‘akıllı’ giysiler üretiyor. Benzeri olmayan bir elektro tekstil ürününden yapılan ve ses üreten bir aygıtla bağlanmış yelek giyen çocuklar, dokunmaya duyarlı kumaşa hafifçe vurarak konuşmalarının başkaları tarafından anlaşılabilir duruma gelmesini sağlamaktadırlar. Bu kumaş, normal kumaş ve içine kusursuz bir şekilde yerleştirilmiş karbon iplikçikler sayesinde elektriği iletebilen bir fileden yapılmıştır. Kumaş üzerine basınç uygulandığında, iletken iplikçiklerden geçen sinyaller değiştirilir ve bir bilgisayar devresi kumaşa nerede dokunulduğunu belirler. Daha sonra, bu devre kendisine bağlı olan ve iki kibrit kutusundan daha büyük olmayan bir elektronik aracın tetiklemektir. Bilim adamlarından birisi şöyle söylemektedir: “İşin en çarpıcı kısmı, kumaşı nasıl dokuduğumuz ve sinyalleri onun içinden nasıl gönderdiğimizdir - onu normal bir kumaşta var olan dokunuş şekli içerisine, kimsenin göremeyeceği şekilde yerleştirebiliriz.” Bu kumaş, zarar görmeksizin yıkanabilir, nesnelerin etrafına sarılabilir ya da sıkılıp top durumuna getirilebilir. Bilim adamları, onun toptan üretiminin ucuz olacağını da ileri sürmektedirler.

Soru 1: GIYSİLER

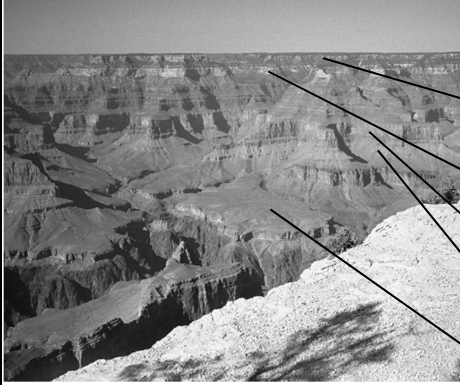
Aşağıdaki laboratuvar araçlarından hangisi kumaşın elektriği ilettiğini deneyebilmemiz için gerekecek araçlar arasında yer alabilir?

- E. Voltmetre
- F. Işık kutusu
- G. Mikrometre
- H. Ses ölçer

GRAND KANYON (BÜYÜK KANYON)

Grand Canyon (Büyük Kanyon) Amerika Birleşik Devletleri'ndeki bir çöldedir. Burası, birçok kaya katmanını içeren çok geniş ve derin bir kanyondur. Geçmiş bir zaman diliminde yer kabuğunda meydana gelen hareketler bu katmanları yukarıya doğru itmiştir. Günümüzde bu kanyonun bazı bölümleri 1.6 km derinliğindedir. Kanyonun dibinde Colorado Nehri akmaktadır.

Aşağıda Büyük Kanyon' un güney kenarından çekilmiş bir resmi görülmektedir. Kanyon 'un bu resminde birkaç değişik kaya tabakası görülebilmektedir.



Kireçtaşı (A)

Kilin sıkışması ile
oluşan tortul kavaç

Kireçtaşı(B)

Kilin sıkışması ile
oluşan tortul kayaç (B)

Tortul kayaç ve granit

Soru 2:

Büyük Kanyon' da hava sıcaklığı 0 °C 'ın altındaki sıcaklıklardan 40 °C'ın üstündeki sıcaklıklara kadar değişebilmektedir. Burası bir çöl alanı olmasına karşın, kayalardaki çatlaklarda bazen su bulunabilmektedir. Bu sıcaklık değişimleri ve çatlaklardaki su kayaların parçalanmasını nasıl hızlandırabilmektedir?

- E. Donan su, sıcak kayaları eritir.
- F. Su, kayaları birbirine yapıştırır.
- G. Buz kayaların yüzeyini düzleştirir.
- H. Kaya çatlaklarında donan su genleşir.

Soru 3:

Büyük Kanyon'un "Kireçtaşı (A)" olarak belirtilen tabakasında deniz tarağı, balık ve mercan gibi birçok deniz hayvanının fosilleri bulunmaktadır. Bu fosillerin orada bulunabilmeleri için milyonlarca yıl önce ne olmuştur?

- E. Eski zamanlarda insanlar okyanustan oraya su ürünleri getirmişlerdir.
- F. Bir zamanlar okyanuslarda büyük dalgalar oluştu ve bunlar deniz yaşamını karalara sürükledi.
- G. O zamanlarda okyanus buraları kaplamıştı, sonra sular eski yerine çekildi.
- H. Bazı deniz hayvanları, denize göç etmeden önce bir süre karada yaşadılar.

GÜNEŞTEN KORUYUCULAR

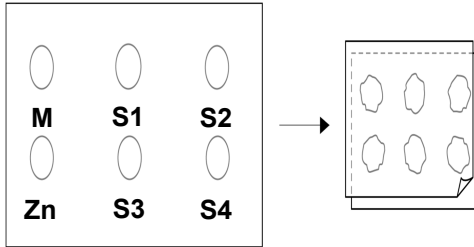
Jale ve Osman, güneşten koruma ürünlerinden hangisinin ciltleri için en iyi korumayı sağladığını merak ettiler. Güneşten koruma ürünleri için, her ürünün güneş ışığındaki ultraviyole ışınlarını ne derecede emdiğini gösteren bir *Güneşten Koruma Faktörü (GKF)* tanımlanmıştır. GKF'si yüksek olan bir güneşten koruyucu, GKF'si düşük olan bir güneşten koruyucuya göre cildi daha uzun süre korur.

Jale, bazı güneşten koruma ürünlerini birbiriyle karşılaştırmak için bir yol düşündü. Osman ile birlikte aşağıdaki malzemeleri topladılar:

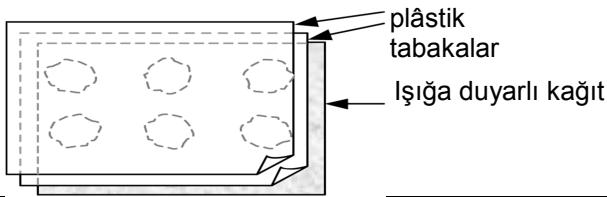
- güneş ışığını emmeyen (geçiren) iki temiz plastik tabaka;
- bir adet ışığa duyarlı kağıt;
- mineral yağ (M) ve çinko oksit (ZnO) içeren bir krem
- S1, S2, S3 ve S4 adını verdikleri dört farklı güneşten koruma ürünü.

Jale ve Osman, mineral yağı güneş ışınlarının çok büyük bir kısmını geçirdiği için, çinko oksidi de güneş ışınlarının tamamına yakını geçirmediği için seçtiler.

Osman, bir plastik tabaka üzerinde yuvarlak içine alınmış yerlerin her birine her maddeden birer damla koydu sonra bunların üzerini ikinci bir plastik tabaka ile kapattı. Bu plastik tabakaların üzerine büyük bir kitap yerleştirerek üstten iyice bastırdı.



Daha sonra, Jale hazırladıkları plastik tabakaları ışığa duyarlı kâğıdın üzerine koydu. Işığa duyarlı kâğıt, güneş ışığında tutulduğu süreye göre koyu griden beyaza (ya da çok açık griye) doğru renk değiştiren bir kâğıttır. En sonunda da, Osman hazırladıkları bu tabakaları güneşli bir yere koydu.



Soru 4:

Aşağıdaki ifadelerden hangisi, güneşten koruyucuların etkililiğini karşılaştırma amacıyla yapılan bir çalışmada mineral yağ ve çinko oksidin rolünün bilimsel tanımıdır?

- E. Mineral yağ ve çinko oksidin ikisi de etkisi araştırılan birer etkidir.
- F. Mineral yağ test edilen bir etken, çinko oksit ise karşılaştırma için kullanılan bir maddedir.
- G. Mineral yağ karşılaştırma için kullanılan bir madde, çinko oksit ise test edilen bir etkidir.
- H. Mineral yağ ve çinko oksidin ikisi de karşılaştırma için kullanılan birer maddedir.

Soru 5:

Jale ve Osman'ın yanıtlamaya çalıştığı soru aşağıdakilerden hangisidir?

- E. Güneşten koruyucu maddelerden her birinin koruma gücü diğerlerine kıyasla nasıldır?
- F. Güneşten koruyucular cildi ultraviyole ışınlarından nasıl korur?
- G. Mineral yağdan daha az koruma sağlayan bir güneşten koruyucu var mıdır?
- H. Çinko oksitten daha çok koruma sağlayan bir güneşten koruyucu var mıdır?

Soru 6:

İkinci plastik tabakanın üzerine neden iyice bastırılmıştır?

- E. Damlaların kurummasını önlemek için
- F. Damlaları mümkün olduğunca yaymak için
- G. Damlaları yuvarlaklar içinde tutmak için
- H. Damlalara eşit kalınlık vermek için

AŞININ TARİHÇESİ

Mary Montagu güzel bir kadındı. 1715 yılında çiçek hastalığına yakalandı. Hastalığı geçirdi; fakat izleri kaldı. 1717 yılında Türkiye'de yaşarken, bu ülkede yaygınca kullanılmakta olan ve adına aşılama denen bir tedaviyi gördü. Bu tedavide sağlıklı gencin derisi çizilerek ona zayıflatılmış çiçek virüsü veriliyordu. Kişi kısa bir süre için hasta oluyor, ancak hastalığı genellikle çok hafif bir şekilde geçiyordu.

Mary, bu aşılama yönteminin güvenli olduğuna inandı ve kendi oğlu ile kızının da bu şekilde aşılmasına izin verdi.

1796 yılında Edward Jenner çiçek hastalığına karşı antikor geliştirmek için insandaki çiçek hastalığı virüsünü değil, ineklerde görülen çiçek hastalığı virüsünü kullanarak aşılama yöntemini geliştirdi. Jenner'in bulduğu bu aşılama yönteminin, çiçek hastalığı virüsü verilmesine kıyasla, yan etkileri daha azdır ve tedavi gören kişi virüsü başka insanlara bulaştıramaz. Bu tedavi biçimi aşılama adıyla tanındı.

Soru 7

İnsanlar hangi çeşit hastalıklara karşı aşılanabilir?

- E. Hemofili gibi kalıtsal hastalıklar
- F. Çocuk felci gibi virüslerin neden olduğu hastalıklar
- G. Şeker hastalığı gibi vücudun işlevsel bozukluklarından kaynaklanan hastalıklar
- H. Tedavisi olmayan her çeşit hastalık

Soru 8

Hayvanlar ya da insanlar bakterilerin neden olduğu bulaşıcı bir hastalığa yakalanır ve iyileşirse, hastalığa neden olan bakteriler genellikle onlarda tekrar hastalık oluşturmaz.

Bunun nedeni aşağıdakilerden hangisidir?

- E. Vücudun, aynı çeşitten bir hastalığa neden olabilecek bütün bakterileri öldürmüş olması
- F. Vücudun, bu tür bakterileri çoğalmadan önce öldürecek antikorlar yapmış olması
- G. Alyuvarların, aynı çeşit hastalığa neden olabilecek bütün bakterileri öldürmesi
- H. Alyuvarların, vücuttaki bu tip bakterileri yakalayarak vücuttan atması.

ASİT YAĞMURU

Aşağıda, Caryatids adı verilen ve Atina Akropolünde 2500 yıl önce inşa edilmiş olan heykellerin fotoğrafı görülmektedir. Heykeller, mermer adı verilen bir cins kayadan yapılmıştır. Mermer kireçtaşı (kalsiyum karbonattan) oluşmaktadır.

Orijinal heykeller 1980 yılında kopyalarıyla değiştirilerek Akropol müzesinin içine alındı. Bu heykeller asit yağmurundan zarar görmüşlerdi.

**Soru 9: ASİT YAĞMURU**

Bir mermer parçasının gece boyunca sirke içine konmadan önceki kütlesi 2,0 gramdır. Sonraki gün bu parça sirkeden çıkarılarak kurutulmuştur. Kurutulmuş olan bu mermer parçasının kütlesi ne kadar olabilir?

- E. 2,0 gramdan daha az
- F. Tam olarak 2,0 gram
- G. 2,0 ile 2,4 gram arasında
- H. 2,4 gramdan fazla

GENETİK YAPISI DEĞİŞTİRİLEN (GYD) MISIR YASAKLANMALIDIR

Doğayı koruma grupları, yeni ortaya çıkan genetik yapısı değiştirilmiş (GYD) mısırın yasaklanmasını istemektedirler.

GYD mısır, geleneksel mısır bitkilerini öldüren yeni ve güçlü bir zararlı ot ilacından etkilenmeyecek şekilde geliştirilmiştir. Bu yeni zararlı ot ilacı, mısır tarlalarında kullanıldığında büyüyen zararlı otların pek çoğunu öldürecektir.

Doğayı koruma yanlısı olanlar, yeni ilacın öldüreceği zararlı otlar küçük hayvanların ve özellikle böceklerin beslenmesine yaradığından, bu yeni zararlı ot ilacının GYD mısır ile birlikte kullanılmasının çevre için kötü olacağını söylemektedirler. GYD mısırın kullanılmasını destekleyenler buna cevap olarak bilimsel bir incelemenin, sonucun bu şekilde olmayacağını gösterdiğini söylemektedirler.

Yukarıdaki yazıda sözü edilen bilimsel incelemenin bazı ayrıntıları şunlardır:

- Mısır, ülkenin değişik yerlerindeki 200 tarlaya ekilmiştir.
- Her tarla önce iki eşit parçaya ayrılmıştır. Tarlanın bir parçasında yeni güçlü zararlı ot ilacı ile ilaçlanmış olan genetik yapısı değiştirilmiş (GYD) mısır yetiştirilmiştir. Tarlanın diğer parçasında da geleneksel zararlı ot ilacı ile ilaçlanmış geleneksel mısır yetiştirilmiştir.
- Yeni zararlı ot ilacı ile ilaçlanan GYD mısır içinde bulunan böceklerin sayısı, geleneksel zararlı ot ilacı ile ilaçlanmış olan geleneksel mısır içinde bulunan böceklerin sayısı ile hemen hemen aynıdır.

Soru 10

Mısır ülkenin değişik yerlerindeki 200 tarlaya ekilmiştir. Bilim adamları niçin birden fazla yerde ekim yapmışlardır?

- E. Yeni GYD mısır, birçok çiftçinin deneme fırsatı bulması için
- F. Ne kadar GYD mısır yetiştirebileceklerini görmeleri için
- G. GYD mısır ekimini olabildiğince geniş bir alana yaymak için
- H. Mısırın değişik yetiştirme koşullarda nasıl büyüyeceğini görmek için

EK 3
EZDIFF MH VE LR ANALİZ ÇIKTILARI

KANADA MH ve LR ANALİZ ÇIKTISI

Mantel-Haenszel and Logistic Regression Analysis of
DIFFERENTIAL ITEM FUNCTIONING

Programmed by Niels G. Waller

EZDIF example input file

Reference Group: E:\ezdiff\KANADA-ING.prn

Focal Group: E:\ezdiff\KANADA-FR.prn

Number of Cases in Reference Group: 1288

Number of Cases in Focal Group: 359

Conditioning Levels

0 6 11 16
5 10 15 20

Note:

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

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

Results for Pass Number: 1

ITEM	Alpha	X ²	P-Value	SE	MH D-DIF (MH D-DIF)
A 1	0.781	2.013	0.156	0.580	0.387
A 2	1.418	2.903	0.088	-0.820	0.454
A 3**	1.483	7.421	0.006	-0.927	0.332
A 4	1.214	2.125	0.145	-0.456	0.307
A 5	0.820	1.540	0.215	0.467	0.353
A 6*	1.332	4.537	0.033	-0.674	0.310
A 7	0.927	0.238	0.626	0.177	0.321
A 8	1.194	1.301	0.254	-0.417	0.341
A 9	0.827	1.488	0.223	0.447	0.347
A 10	0.831	0.852	0.356	0.434	0.425
B 11***	1.830	16.179	0.000	-1.420	0.353
A 12	1.083	0.293	0.588	-0.188	0.311
A 13	1.319	3.748	0.053	-0.650	0.328
A 14	1.042	0.061	0.805	-0.096	0.306
A 15	1.136	0.645	0.422	-0.300	0.344
A 16	1.176	1.443	0.230	-0.381	0.300
A 17*	1.375	5.850	0.016	-0.749	0.302
B 18***	0.597	9.581	0.002	1.212	0.385
B 19***	0.570	11.175	0.001	1.320	0.380
A 20	1.013	0.001	0.970	-0.030	0.297

A 21 0.904 0.521 0.470 0.236 0.300

Number of Items Purged in Pass 1: 0

Item Numbers:

Results for Pass Number: 2

ITEM	Alpha	X ²	P-Value	SE	MH D-DIF (MH D-DIF)
A 1	0.781	2.013	0.156	0.580	0.387
A 2	1.418	2.903	0.088	-0.820	0.454
A 3**	1.483	7.421	0.006	-0.927	0.332
A 4	1.214	2.125	0.145	-0.456	0.307
A 5	0.820	1.540	0.215	0.467	0.353
A 6*	1.332	4.537	0.033	-0.674	0.310
A 7	0.927	0.238	0.626	0.177	0.321
A 8	1.194	1.301	0.254	-0.417	0.341
A 9	0.827	1.488	0.223	0.447	0.347
A 10	0.831	0.852	0.356	0.434	0.425
B 11***	1.830	16.179	0.000	-1.420	0.353
A 12	1.083	0.293	0.588	-0.188	0.311
A 13	1.319	3.748	0.053	-0.650	0.328
A 14	1.042	0.061	0.805	-0.096	0.306
A 15	1.136	0.645	0.422	-0.300	0.344
A 16	1.176	1.443	0.230	-0.381	0.300
A 17*	1.375	5.850	0.016	-0.749	0.302
B 18***	0.597	9.581	0.002	1.212	0.385
B 19***	0.570	11.175	0.001	1.320	0.380
A 20	1.013	0.001	0.970	-0.030	0.297
A 21	0.904	0.521	0.470	0.236	0.300

Logistic Regression Procedure. Note: Items purged from previous step

Item Number: 1	Estimate	SE Estimate	Z	p-value
Intercept	-2.565526	0.286374	-8.958649	0.000000
Trait	0.318916	0.023190	13.752481	0.000000
Group	-0.325457	0.530168	-0.613875	0.539438
Trait x Group	0.003280	0.046379	0.070718	0.943637

Item Number: 2	Estimate	SE Estimate	Z	p-value
Intercept	0.346597	0.297515	1.164971	0.244308
Trait	0.148843	0.023249	6.402242	0.000000
Group	0.783806	0.555418	1.411200	0.158497
Trait x Group	-0.040438	0.046497	-0.869686	0.384681

Item Number: 3	Estimate	SE Estimate	Z	p-value
Intercept	-2.735612	0.260488	-10.501876	0.000000

Trait	0.275361	0.019643	14.018320	0.000000
Group	0.619858	0.502578	1.233356	0.217733
Trait x Group	-0.019753	0.039286	-0.502808	0.615210

Item Number: 4	Estimate	SE Estimate	Z	p-value
Intercept	-3.304093	0.249637	-13.235606	0.000000
Trait	0.251579	0.017919	14.039878	0.000000
Group	-1.013990	0.499095	-2.031658	0.042452
Trait x Group	0.087794	0.035838	2.449750	0.014466

Item Number: 5	Estimate	SE Estimate	Z	p-value
Intercept	-3.292777	0.279538	-11.779360	0.000000
Trait	0.328202	0.021845	15.024078	0.000000
Group	-0.082128	0.535891	-0.153256	0.878227
Trait x Group	-0.013805	0.043690	-0.315977	0.752086

Item Number: 6	Estimate	SE Estimate	Z	p-value
Intercept	-1.926516	0.225446	-8.545349	0.000000
Trait	0.190648	0.016562	11.511111	0.000000
Group	-0.440527	0.441102	-0.998696	0.318184
Trait x Group	0.055331	0.033124	1.670417	0.095150

Item Number: 7	Estimate	SE Estimate	Z	p-value
Intercept	-3.804997	0.268788	-14.156113	0.000000
Trait	0.290691	0.019480	14.922597	0.000000
Group	-0.442314	0.538590	-0.821244	0.411703
Trait x Group	0.024494	0.038960	0.628687	0.529698

Item Number: 8	Estimate	SE Estimate	Z	p-value
Intercept	-2.277403	0.257625	-8.839987	0.000000
Trait	0.260754	0.019953	13.068280	0.000000
Group	0.620040	0.492200	1.259731	0.208060
Trait x Group	-0.040198	0.039906	-1.007316	0.314027

Item Number: 9	Estimate	SE Estimate	Z	p-value
Intercept	-3.554117	0.279820	-12.701423	0.000000
Trait	0.333698	0.021547	15.486886	0.000000
Group	-0.265005	0.541133	-0.489722	0.624438
Trait x Group	0.001580	0.043094	0.036671	0.970754

Item Number: 10	Estimate	SE Estimate	Z	p-value
Intercept	-3.592101	0.357867	-10.037519	0.000000
Trait	0.435609	0.030305	14.374281	0.000000
Group	0.154675	0.640583	0.241460	0.809248
Trait x Group	-0.038050	0.060610	-0.627789	0.530286

Item Number: 11	Estimate	SE Estimate	Z	p-value
Intercept	-2.858051	0.271993	-10.507823	0.000000

Trait	0.308145	0.020442	15.074000	0.000000
Group	-0.265591	0.507159	-0.523684	0.600615
Trait x Group	0.072756	0.040884	1.779561	0.075451

Item Number: 12	Estimate	SE Estimate	Z	p-value
Intercept	-3.637927	0.261511	-13.911167	0.000000
Trait	0.265357	0.018689	14.198383	0.000000
Group	-0.628584	0.526649	-1.193555	0.232935
Trait x Group	0.048546	0.037378	1.298775	0.194321

Item Number: 13	Estimate	SE Estimate	Z	p-value
Intercept	-2.418322	0.250003	-9.673182	0.000000
Trait	0.256841	0.018878	13.605562	0.000000
Group	-0.013509	0.479492	-0.028173	0.977530
Trait x Group	0.022195	0.037755	0.587870	0.556752

Item Number: 14	Estimate	SE Estimate	Z	p-value
Intercept	-3.007874	0.250493	-12.007799	0.000000
Trait	0.230904	0.018122	12.741317	0.000000
Group	0.405935	0.502147	0.808399	0.419053
Trait x Group	-0.030063	0.036245	-0.829445	0.407050

Item Number: 15	Estimate	SE Estimate	Z	p-value
Intercept	-2.420681	0.256568	-9.434838	0.000000
Trait	0.272139	0.019661	13.841327	0.000000
Group	-0.343922	0.486172	-0.707408	0.479478
Trait x Group	0.037763	0.039323	0.960344	0.337114

Item Number: 16	Estimate	SE Estimate	Z	p-value
Intercept	-2.843368	0.247809	-11.474051	0.000000
Trait	0.190224	0.017523	10.855391	0.000000
Group	0.418164	0.501643	0.833589	0.404711
Trait x Group	-0.020454	0.035047	-0.583608	0.559616

Item Number: 17	Estimate	SE Estimate	Z	p-value
Intercept	-2.695853	0.241146	-11.179318	0.000000
Trait	0.180379	0.017050	10.579649	0.000000
Group	-0.007793	0.487342	-0.015990	0.987246
Trait x Group	0.022206	0.034099	0.651217	0.515056

Item Number: 18	Estimate	SE Estimate	Z	p-value
Intercept	-2.540665	0.275588	-9.219078	0.000000
Trait	0.303279	0.022252	13.629475	0.000000
Group	-0.731176	0.517221	-1.413661	0.157773
Trait x Group	0.014682	0.044503	0.329910	0.741537

Item Number: 19	Estimate	SE Estimate	Z	p-value
Intercept	-3.728815	0.319058	-11.686947	0.000000

Trait	0.380220	0.026361	14.423366	0.000000
Group	0.753930	0.607216	1.241618	0.214669
Trait x Group	-0.121588	0.052723	-2.306166	0.021305

Item Number: 20	Estimate	SE Estimate	Z	p-value
Intercept	-2.426885	0.226025	-10.737225	0.000000
Trait	0.187893	0.016320	11.513200	0.000000
Group	-0.993628	0.451700	-2.199753	0.028053
Trait x Group	0.072989	0.032640	2.236209	0.025558

Item Number: 21	Estimate	SE Estimate	Z	p-value
Intercept	-2.895570	0.243081	-11.911948	0.000000
Trait	0.192821	0.017267	11.166960	0.000000
Group	-0.173666	0.492243	-0.352805	0.724309
Trait x Group	0.003265	0.034534	0.094552	0.924690

AVUSTRALYA-İNGİLTERE MH ve LR ANALİZ ÇIKTISI

Mantel-Haenszel and Logistic Regression Analysis of DIFFERENTIAL ITEM FUNCTIONING

Programmed by Niels G. Waller

EZDIF example input file

Reference Group: E:\ezdiff\AVUSTRALYA5.prn

Focal Group: E:\ezdiff\İNGİLTERE5.prn

Number of Cases in Reference Group: 1067

Number of Cases in Focal Group: 968

Conditioning Levels

0 6 11 16

5 10 15 20

Note:

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

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

Results for Pass Number: 1

SE

ITEM	Alpha	X ²	P-Value	MH D-DIF	(MH D-DIF)
A 1*	0.737	5.754	0.016	0.718	0.294
A 2	0.818	1.143	0.285	0.472	0.404
A 3	0.864	1.637	0.201	0.345	0.258
A 4	0.992	0.001	0.976	0.020	0.242
A 5	1.043	0.108	0.743	-0.100	0.260
A 6	1.104	0.830	0.362	-0.233	0.242
A 7	1.019	0.016	0.900	-0.043	0.243
A 8	1.207	2.931	0.087	-0.443	0.251
A 9	1.059	0.238	0.626	-0.134	0.248
A 10	0.960	0.060	0.806	0.095	0.306
A 11	0.881	0.913	0.339	0.298	0.292
A 12	1.030	0.055	0.814	-0.069	0.239
A 13	1.186	2.187	0.139	-0.400	0.260
A 14*	1.228	4.141	0.042	-0.483	0.232
A 15**	1.351	7.172	0.007	-0.707	0.259
A 16***	1.328	8.382	0.004	-0.666	0.227
A 17	1.000	0.002	0.965	0.001	0.227
CF 18***	0.491	30.629	0.000	1.671	0.302
A 19***	0.663	11.847	0.001	0.964	0.276
A 20***	1.498	16.548	0.000	-0.949	0.231
A 21	0.978	0.032	0.857	0.051	0.224

Number of Items Purged in Pass 1: 1

Item Numbers:

18

Results for Pass Number: 2

SE

ITEM	Alpha	X ²	P-Value	MH D-DIF	(MH D-DIF)
A 1*	0.733	5.962	0.015	0.729	0.292
A 2	0.789	1.640	0.200	0.556	0.403
A 3	0.851	2.015	0.156	0.378	0.257
A 4	0.987	0.005	0.942	0.030	0.240
A 5	1.021	0.017	0.896	-0.049	0.262
A 6	1.077	0.440	0.507	-0.173	0.242
A 7	0.998	0.001	0.973	0.004	0.242
A 8	1.192	2.567	0.109	-0.414	0.250
A 9	1.053	0.189	0.664	-0.120	0.247
A 10	0.936	0.192	0.661	0.155	0.307
A 11	0.852	1.512	0.219	0.377	0.291
A 12	1.024	0.035	0.852	-0.057	0.239
A 13	1.185	2.193	0.139	-0.398	0.259
A 14	1.206	3.389	0.066	-0.440	0.233
A 15*	1.331	6.474	0.011	-0.672	0.259
A 16**	1.308	7.545	0.006	-0.630	0.226
A 17	0.987	0.008	0.927	0.032	0.226
CF 18***	0.491	30.629	0.000	1.671	0.302
B 19***	0.645	13.414	0.000	1.032	0.278
A 20***	1.467	15.068	0.000	-0.901	0.229
A 21	0.950	0.236	0.627	0.120	0.225

Logistic Regression Procedure. Note: Items purged from previous step

Item Number: 1	Estimate	SE Estimate	Z	p-value
Intercept	-1.971990	0.213565	-9.233676	0.000000
Trait	0.298318	0.017509	17.037590	0.000000
Group	-0.905744	0.381558	-2.373808	0.017794
Trait x Group	0.051518	0.035019	1.471149	0.141565

Item Number: 2	Estimate	SE Estimate	Z	p-value
Intercept	0.468651	0.259995	1.802536	0.071762
Trait	0.178111	0.021405	8.320892	0.000000
Group	0.547666	0.462956	1.182977	0.237099
Trait x Group	-0.081193	0.042811	-1.896565	0.058173

Item Number: 3	Estimate	SE Estimate	Z	p-value
----------------	----------	-------------	---	---------

Intercept	-3.007006	0.202530	-14.847207	0.000000
Trait	0.316795	0.016073	19.709551	0.000000
Group	-1.018931	0.383315	-2.658211	0.007981
Trait x Group	0.068691	0.032146	2.136831	0.032854

Item Number: 4	Estimate	SE Estimate	Z	p-value
Intercept	-3.580718	0.195735	-18.293734	0.000000
Trait	0.299189	0.014984	19.967821	0.000000
Group	-0.101909	0.389956	-0.261335	0.793888
Trait x Group	0.003083	0.029967	0.102896	0.918066

Item Number: 5	Estimate	SE Estimate	Z	p-value
Intercept	-3.241922	0.206526	-15.697374	0.000000
Trait	0.335000	0.016483	20.323412	0.000000
Group	0.053022	0.391276	0.135511	0.892236
Trait x Group	-0.007932	0.032967	-0.240600	0.809914

Item Number: 6	Estimate	SE Estimate	Z	p-value
Intercept	-2.138688	0.175699	-12.172449	0.000000
Trait	0.231292	0.013742	16.831284	0.000000
Group	0.086255	0.338708	0.254659	0.799038
Trait x Group	-0.003634	0.027484	-0.132226	0.894832

Item Number: 7	Estimate	SE Estimate	Z	p-value
Intercept	-3.963723	0.202572	-19.566940	0.000000
Trait	0.307600	0.015291	20.117039	0.000000
Group	-0.200950	0.409485	-0.490738	0.623719
Trait x Group	0.010336	0.030581	0.338001	0.735434

Item Number: 8	Estimate	SE Estimate	Z	p-value
Intercept	-2.055911	0.181892	-11.302891	0.000000
Trait	0.243271	0.014348	16.955498	0.000000
Group	-0.188313	0.344763	-0.546210	0.585044
Trait x Group	0.028874	0.028695	1.006245	0.314541

Item Number: 9	Estimate	SE Estimate	Z	p-value
Intercept	-3.456836	0.200191	-17.267668	0.000000
Trait	0.319560	0.015641	20.430933	0.000000
Group	-0.018677	0.389894	-0.047904	0.961802
Trait x Group	0.000352	0.031282	0.011245	0.991030

Item Number: 10	Estimate	SE Estimate	Z	p-value
Intercept	-3.290625	0.254501	-12.929735	0.000000
Trait	0.417810	0.021195	19.712973	0.000000
Group	-0.536215	0.443787	-1.208271	0.227229
Trait x Group	0.039810	0.042389	0.939149	0.347881

Item Number: 11	Estimate	SE Estimate	Z	p-value
-----------------	----------	-------------	---	---------

Intercept	-2.318263	0.214132	-10.826324	0.000000
Trait	0.318311	0.017588	18.097933	0.000000
Group	-0.208940	0.384542	-0.543347	0.587012
Trait x Group	-0.000027	0.035177	-0.000756	0.999397

Item Number: 12	Estimate	SE Estimate	Z	p-value
Intercept	-3.948794	0.201051	-19.640762	0.000000
Trait	0.286655	0.014989	19.123867	0.000000
Group	0.220858	0.411263	0.537024	0.591370
Trait x Group	-0.018438	0.029979	-0.615028	0.538676

Item Number: 13	Estimate	SE Estimate	Z	p-value
Intercept	-2.636677	0.197085	-13.378346	0.000000
Trait	0.298832	0.015797	18.916698	0.000000
Group	0.571753	0.370173	1.544556	0.122770
Trait x Group	-0.040816	0.031595	-1.291855	0.196706

Item Number: 14	Estimate	SE Estimate	Z	p-value
Intercept	-3.068444	0.183204	-16.748772	0.000000
Trait	0.257407	0.013979	18.413773	0.000000
Group	0.273073	0.365230	0.747674	0.454833
Trait x Group	-0.010707	0.027958	-0.382983	0.701814

Item Number: 15	Estimate	SE Estimate	Z	p-value
Intercept	-2.744260	0.199835	-13.732613	0.000000
Trait	0.306825	0.015980	19.200363	0.000000
Group	0.375764	0.374790	1.002597	0.316298
Trait x Group	-0.012637	0.031960	-0.395385	0.692643

Item Number: 16	Estimate	SE Estimate	Z	p-value
Intercept	-2.778466	0.176506	-15.741447	0.000000
Trait	0.181336	0.013072	13.872612	0.000000
Group	1.033448	0.361770	2.856645	0.004370
Trait x Group	-0.060077	0.026143	-2.298033	0.021766

Item Number: 17	Estimate	SE Estimate	Z	p-value
Intercept	-2.645109	0.172293	-15.352399	0.000000
Trait	0.205692	0.012960	15.871796	0.000000
Group	-0.411748	0.346821	-1.187204	0.235429
Trait x Group	0.027715	0.025919	1.069286	0.285199

Item Number: 18	Estimate	SE Estimate	Z	p-value
Intercept	-2.696414	0.230078	-11.719577	0.000000
Trait	0.334192	0.017939	18.629597	0.000000
Group	-1.207779	0.411117	-2.937796	0.003381
Trait x Group	0.038590	0.035877	1.075617	0.282358

Item Number: 19	Estimate	SE Estimate	Z	p-value
-----------------	----------	-------------	---	---------

Intercept	-3.607398	0.227221	-15.876160	0.000000
Trait	0.386853	0.018521	20.887245	0.000000
Group	-0.374518	0.421484	-0.888570	0.374448
Trait x Group	-0.015157	0.037042	-0.409196	0.682484

Item Number: 20	Estimate	SE Estimate	Z	p-value
Intercept	-2.937491	0.179349	-16.378628	0.000000
Trait	0.236321	0.013568	17.417448	0.000000
Group	0.255641	0.359628	0.710849	0.477344
Trait x Group	0.007545	0.027136	0.278046	0.781034

Item Number: 21	Estimate	SE Estimate	Z	p-value
Intercept	-2.680343	0.172612	-15.528128	0.000000
Trait	0.199523	0.012920	15.442884	0.000000
Group	-0.408069	0.349115	-1.168866	0.242736
Trait x Group	0.024808	0.025840	0.960077	0.337249

İNGİLTERE - TÜRKİYE 1. KİTAPÇIK MH ve LR ANALİZ ÇIKTISI

Mantel-Haenszel and Logistic Regression Analysis of DIFFERENTIAL ITEM FUNCTIONING

Programmed by Niels G. Waller

EZDIF example input file

Reference Group: E:\ezdiff\İNGİLTERE1.prn

Focal Group: E:\ezdiff\TÜRKİYE1.prn

Number of Cases in Reference Group: 790

Number of Cases in Focal Group: 365

Conditioning Levels

0 6 11 16

5 10 15 19

Note:

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

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

Results for Pass Number: 1

ITEM	Alpha	X ²	P-Value	SE	MH D-DIF (MH D-DIF)
A 1	1.194	1.047	0.306	-0.416	0.379
A 2	1.053	0.064	0.800	-0.121	0.365
A 3	1.146	0.806	0.369	-0.320	0.337
CF 4***	0.435	24.322	0.000	1.958	0.398
CR 5***	2.819	44.570	0.000	-2.435	0.372
A 6	0.792	2.106	0.147	0.548	0.359
A 7	0.829	1.369	0.242	0.440	0.352
CR 8***	2.739	39.282	0.000	-2.368	0.386
A 9	1.060	0.077	0.781	-0.136	0.380
A 10	1.157	0.784	0.376	-0.342	0.356
A 11*	0.689	5.738	0.017	0.875	0.355
A 12	0.982	0.002	0.961	0.044	0.355
A 13	1.344	2.948	0.086	-0.694	0.383
CR 14***	3.135	51.225	0.000	-2.685	0.382
CF 15***	0.475	20.905	0.000	1.752	0.380
A 16	0.810	1.587	0.208	0.496	0.370
A 17*	0.720	3.850	0.050	0.773	0.386
B 18***	1.587	10.122	0.001	-1.086	0.333
B 19*	0.575	6.568	0.010	1.302	0.488

Results for Pass Number: 2

WARNING: Insufficient Data Found in Level: 16 - 19

		SE				
ITEM		Alpha	X ²	P-Value	MH D-DIF	(MH D-DIF)
A	1*	1.504	6.378	0.012	-0.959	0.371
A	2	1.287	2.380	0.123	-0.593	0.365
A	3*	1.348	4.406	0.036	-0.702	0.330
B	4***	0.578	11.819	0.001	1.289	0.371
CR	5***	3.118	51.650	0.000	-2.673	0.380
A	6	0.871	0.707	0.400	0.324	0.354
A	7	0.933	0.150	0.698	0.162	0.349
CR	8***	2.984	48.883	0.000	-2.569	0.380
A	9	1.356	3.600	0.058	-0.715	0.366
A	10*	1.440	6.151	0.013	-0.857	0.340
A	11	0.755	3.247	0.072	0.660	0.352
A	12	1.169	0.989	0.320	-0.367	0.347
B	13***	1.643	9.110	0.003	-1.167	0.375
CR	14***	3.289	57.521	0.000	-2.798	0.381
A	15*	0.679	6.349	0.012	0.911	0.350
A	16	0.938	0.109	0.741	0.151	0.368
A	17	0.900	0.337	0.562	0.248	0.377
B	18***	1.756	15.712	0.000	-1.323	0.327
A	19	0.806	0.968	0.325	0.508	0.468

Logistic Regression Procedure. Note: Items purged from previous step

Item Number: 1	Estimate	SE Estimate	Z	p-value
Intercept	-2.450775	0.267852	-9.149726	0.000000
Trait	0.398546	0.029110	13.691135	0.000000
Group	-0.433832	0.477529	-0.908494	0.363836
Trait x Group	0.102405	0.058219	1.758943	0.078893

Item Number: 2	Estimate	SE Estimate	Z	p-value
Intercept	-4.366141	0.287974	-15.161573	0.000000
Trait	0.459205	0.030005	15.304444	0.000000
Group	0.116067	0.578205	0.200737	0.840945
Trait x Group	0.007450	0.060009	0.124151	0.901221

Item Number: 3	Estimate	SE Estimate	Z	p-value
Intercept	-2.120064	0.221267	-9.581486	0.000000
Trait	0.281764	0.023174	12.158553	0.000000
Group	-1.181525	0.422707	-2.795138	0.005287
Trait x Group	0.167425	0.046348	3.612319	0.000319

Item Number: 4	Estimate	SE Estimate	Z	p-value
Intercept	-4.368420	0.277736	-15.728704	0.000000
Trait	0.334663	0.026737	12.516780	0.000000

Group	-2.977790	0.587728	-5.066610	0.000000
Trait x Group	0.212049	0.053474	3.965429	0.000078

Item Number: 5	Estimate	SE Estimate	Z	p-value
Intercept	-3.513857	0.301592	-11.651048	0.000000
Trait	0.444216	0.029539	15.038191	0.000000
Group	0.095088	0.541433	0.175623	0.860626
Trait x Group	0.127173	0.059078	2.152610	0.031588

Item Number: 6	Estimate	SE Estimate	Z	p-value
Intercept	-3.447184	0.241471	-14.275776	0.000000
Trait	0.295657	0.024247	12.193558	0.000000
Group	-0.849877	0.498129	-1.706139	0.088293
Trait x Group	0.067923	0.048494	1.400654	0.161628

Item Number: 7	Estimate	SE Estimate	Z	p-value
Intercept	-2.184084	0.229529	-9.515518	0.000000
Trait	0.315881	0.025387	12.442601	0.000000
Group	0.134520	0.439608	0.305999	0.759669
Trait x Group	-0.032931	0.050774	-0.648574	0.516763

Item Number: 8	Estimate	SE Estimate	Z	p-value
Intercept	-1.555676	0.276555	-5.625195	0.000000
Trait	0.299868	0.026728	11.219165	0.000000
Group	-0.112233	0.473236	-0.237160	0.812581
Trait x Group	0.147100	0.053456	2.751766	0.006035

Item Number: 9	Estimate	SE Estimate	Z	p-value
Intercept	-2.561421	0.265517	-9.646931	0.000000
Trait	0.404816	0.029401	13.768612	0.000000
Group	0.036689	0.482846	0.075985	0.939446
Trait x Group	0.026127	0.058803	0.444312	0.656913

Item Number: 10	Estimate	SE Estimate	Z	p-value
Intercept	-3.317353	0.252474	-13.139391	0.000000
Trait	0.387143	0.026503	14.607267	0.000000
Group	-0.592651	0.490994	-1.207044	0.227701
Trait x Group	0.100216	0.053007	1.890624	0.058963

Item Number: 11	Estimate	SE Estimate	Z	p-value
Intercept	-3.403830	0.239054	-14.238777	0.000000
Trait	0.304105	0.024214	12.559255	0.000000
Group	-0.879582	0.491139	-1.790903	0.073611
Trait x Group	0.056628	0.048427	1.169334	0.242548

Item Number: 12	Estimate	SE Estimate	Z	p-value
Intercept	-3.538320	0.248689	-14.227901	0.000000
Trait	0.363563	0.025614	14.193988	0.000000

Group	-0.748975	0.498994	-1.500970	0.133679
Trait x Group	0.092376	0.051228	1.803249	0.071650

Item Number: 13	Estimate	SE Estimate	Z	p-value
Intercept	-2.003706	0.259617	-7.717940	0.000000
Trait	0.363901	0.029038	12.531764	0.000000
Group	0.801419	0.470157	1.704575	0.088584
Trait x Group	-0.046183	0.058077	-0.795217	0.426676

Item Number: 14	Estimate	SE Estimate	Z	p-value
Intercept	-1.175050	0.262230	-4.480995	0.000008
Trait	0.255243	0.025365	10.062736	0.000000
Group	0.378307	0.461892	0.819038	0.412960
Trait x Group	0.096306	0.050730	1.898404	0.057931

Item Number: 15	Estimate	SE Estimate	Z	p-value
Intercept	-3.832184	0.272188	-14.079192	0.000000
Trait	0.414166	0.029624	13.980893	0.000000
Group	0.895828	0.546843	1.638183	0.101698
Trait x Group	-0.162993	0.059247	-2.751063	0.006047

Item Number: 16	Estimate	SE Estimate	Z	p-value
Intercept	-2.539534	0.256372	-9.905665	0.000000
Trait	0.390709	0.029136	13.409911	0.000000
Group	0.263064	0.478028	0.550311	0.582229
Trait x Group	-0.052703	0.058272	-0.904438	0.365981

Item Number: 17	Estimate	SE Estimate	Z	p-value
Intercept	-2.183131	0.268057	-8.144287	0.000000
Trait	0.387892	0.029532	13.134811	0.000000
Group	-1.395408	0.473870	-2.944706	0.003307
Trait x Group	0.157702	0.059063	2.670054	0.007707

Item Number: 18	Estimate	SE Estimate	Z	p-value
Intercept	-1.874139	0.210896	-8.886562	0.000000
Trait	0.234439	0.022172	10.573457	0.000000
Group	0.618227	0.414774	1.490514	0.136404
Trait x Group	-0.010314	0.044345	-0.232584	0.816132

Item Number: 19	Estimate	SE Estimate	Z	p-value
Intercept	-2.073513	0.354917	-5.842244	0.000000
Trait	0.501956	0.042222	11.888514	0.000000
Group	-0.512005	0.567566	-0.902107	0.367217
Trait x Group	0.024974	0.084444	0.295753	0.767480

İNGİLTERE - TÜRKİYE 5. KİTAPÇIK MH ve LR ANALİZ ÇIKTISI

Mantel-Haenszel and Logistic Regression Analysis of DIFFERENTIAL ITEM FUNCTIONING

Programmed by Niels G. Waller

EZDIF example input file

Reference Group: E:\ezdiff\İNGİLTERE5.prn

Focal Group: E:\ezdiff\TÜRKİYE5.prn

Number of Cases in Reference Group: 968

Number of Cases in Focal Group: 333

Conditioning Levels

0 6 11 16
5 10 15 20

Note:

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

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

Results for Pass Number: 1

ITEM	Alpha	X ²	P-Value	SE	MH D-DIF (MH D-DIF)
A 1*	1.409	4.490	0.034	-0.805	0.366
CR 2***	3.828	58.544	0.000	-3.154	0.421
A 3**	1.499	7.658	0.006	-0.951	0.337
B 4***	1.893	17.548	0.000	-1.500	0.360
A 5	1.221	1.657	0.198	-0.469	0.345
A 6**	0.660	7.393	0.007	0.975	0.347
A 7	0.889	0.507	0.476	0.276	0.353
A 8	1.080	0.222	0.638	-0.181	0.333
CF 9***	0.394	36.552	0.000	2.188	0.363
A 10	1.373	3.722	0.054	-0.744	0.371
A 11	1.171	0.988	0.320	-0.371	0.354
CF 12***	0.481	26.683	0.000	1.718	0.340
A 13	0.915	0.264	0.607	0.209	0.354
A 14	1.267	2.682	0.101	-0.556	0.331
A 15	1.208	1.641	0.200	-0.445	0.333
B 16***	0.527	19.443	0.000	1.507	0.339
CF 17***	0.478	26.349	0.000	1.734	0.336
CR 18***	2.692	43.063	0.000	-2.327	0.360
B 19***	1.601	9.508	0.002	-1.107	0.354
A 20	0.796	2.314	0.128	0.537	0.337

A 21* 1.412 4.995 0.025 -0.811 0.349

Results for Pass Number: 2

WARNING: Insufficient Data Found in Level: 16 - 20

SE					
ITEM	Alpha	X ²	P-Value	MH D-DIF	(MH D-DIF)
A 1*	1.393	4.274	0.039	-0.778	0.364
CR 2***	3.589	52.415	0.000	-3.003	0.420
A 3*	1.459	6.482	0.011	-0.888	0.341
B 4***	1.863	16.191	0.000	-1.463	0.362
A 5	1.158	0.823	0.364	-0.344	0.350
B 6***	0.637	8.530	0.003	1.062	0.353
A 7	0.873	0.668	0.414	0.318	0.356
A 8	1.063	0.120	0.729	-0.143	0.337
CF 9***	0.397	35.051	0.000	2.169	0.365
A 10	1.330	3.171	0.075	-0.670	0.362
A 11	1.187	1.151	0.283	-0.403	0.356
CF 12***	0.490	25.284	0.000	1.675	0.337
A 13	0.886	0.519	0.471	0.285	0.356
A 14	1.244	2.277	0.131	-0.512	0.331
A 15	1.152	0.845	0.358	-0.333	0.339
B 16***	0.520	20.342	0.000	1.537	0.340
CF 17***	0.491	24.751	0.000	1.671	0.333
CR 18***	2.487	35.744	0.000	-2.141	0.362
B 19***	1.537	8.025	0.005	-1.011	0.351
A 20	0.770	3.041	0.081	0.614	0.339
A 21*	1.395	4.706	0.030	-0.782	0.348

Logistic Regression Procedure. Note: Items purged from previous step

Item Number: 1	Estimate	SE Estimate	Z	p-value
Intercept	-1.652878	0.235124	-7.029809	0.000000
Trait	0.346377	0.029350	11.801576	0.000000
Group	0.477279	0.427583	1.116223	0.264595
Trait x Group	-0.032716	0.058700	-0.557348	0.577415

Item Number: 2	Estimate	SE Estimate	Z	p-value
Intercept	-0.638006	0.279083	-2.286078	0.022458
Trait	0.275046	0.029849	9.214496	0.000000
Group	1.368469	0.476517	2.871817	0.004167
Trait x Group	-0.017676	0.059699	-0.296083	0.767228

Item Number: 3	Estimate	SE Estimate	Z	p-value
Intercept	-2.434836	0.215465	-11.300378	0.000000
Trait	0.325339	0.025050	12.987570	0.000000

Group	-0.057076	0.417242	-0.136793	0.891222
Trait x Group	0.047397	0.050100	0.946052	0.344351

Item Number: 4	Estimate	SE Estimate	Z	p-value
Intercept	-3.296371	0.228883	-14.401998	0.000000
Trait	0.312392	0.024888	12.551715	0.000000
Group	-0.584901	0.466576	-1.253604	0.210279
Trait x Group	0.126748	0.049777	2.546328	0.011035

Item Number: 5	Estimate	SE Estimate	Z	p-value
Intercept	-2.990967	0.230880	-12.954636	0.000000
Trait	0.384128	0.026967	14.244256	0.000000
Group	-0.477281	0.447713	-1.066043	0.286661
Trait x Group	0.068537	0.053934	1.270755	0.204111

Item Number: 6	Estimate	SE Estimate	Z	p-value
Intercept	-2.135645	0.216009	-9.886832	0.000000
Trait	0.322107	0.027130	11.872623	0.000000
Group	0.014066	0.417764	0.033671	0.973146
Trait x Group	-0.074609	0.054260	-1.375027	0.169431

Item Number: 7	Estimate	SE Estimate	Z	p-value
Intercept	-3.731405	0.245792	-15.181174	0.000000
Trait	0.378708	0.027737	13.653354	0.000000
Group	0.148673	0.502313	0.295976	0.767310
Trait x Group	-0.047946	0.055475	-0.864287	0.387638

Item Number: 8	Estimate	SE Estimate	Z	p-value
Intercept	-2.025879	0.209461	-9.671888	0.000000
Trait	0.298517	0.025383	11.760519	0.000000
Group	0.344189	0.406327	0.847074	0.397156
Trait x Group	-0.049702	0.050766	-0.979034	0.327800

Item Number: 9	Estimate	SE Estimate	Z	p-value
Intercept	-3.028474	0.244609	-12.380902	0.000000
Trait	0.394680	0.029211	13.511311	0.000000
Group	-0.612037	0.472952	-1.294080	0.195937
Trait x Group	-0.061359	0.058422	-1.050269	0.293849

Item Number: 10	Estimate	SE Estimate	Z	p-value
Intercept	-2.760910	0.260903	-10.582140	0.000000
Trait	0.452417	0.031954	14.158520	0.000000
Group	-0.042157	0.473226	-0.089085	0.929033
Trait x Group	0.033070	0.063907	0.517462	0.604948

Item Number: 11	Estimate	SE Estimate	Z	p-value
Intercept	-1.642769	0.220958	-7.434751	0.000000
Trait	0.316532	0.026407	11.986576	0.000000

Group	-0.683098	0.403382	-1.693427	0.090686
Trait x Group	0.108895	0.052814	2.061839	0.039481

Item Number: 12	Estimate	SE Estimate	Z	p-value
Intercept	-3.008446	0.211304	-14.237512	0.000000
Trait	0.286126	0.022629	12.644134	0.000000
Group	-1.968990	0.430571	-4.572970	0.000005
Trait x Group	0.119514	0.045258	2.640723	0.008402

Item Number: 13	Estimate	SE Estimate	Z	p-value
Intercept	-3.056540	0.252873	-12.087269	0.000000
Trait	0.435906	0.031761	13.724388	0.000000
Group	0.521486	0.485724	1.073625	0.283250
Trait x Group	-0.103757	0.063523	-1.633381	0.102704

Item Number: 14	Estimate	SE Estimate	Z	p-value
Intercept	-2.514142	0.200666	-12.528968	0.000000
Trait	0.239275	0.022255	10.751692	0.000000
Group	-1.257212	0.405867	-3.097599	0.002005
Trait x Group	0.161621	0.044509	3.631184	0.000296

Item Number: 15	Estimate	SE Estimate	Z	p-value
Intercept	-2.593956	0.218811	-11.854778	0.000000
Trait	0.345599	0.025604	13.497869	0.000000
Group	-0.553084	0.422801	-1.308143	0.191125
Trait x Group	0.076480	0.051208	1.493522	0.135616

Item Number: 16	Estimate	SE Estimate	Z	p-value
Intercept	-2.637241	0.202120	-13.047879	0.000000
Trait	0.231187	0.022459	10.293767	0.000000
Group	-1.314516	0.414833	-3.168785	0.001577
Trait x Group	0.061751	0.044918	1.374748	0.169518

Item Number: 17	Estimate	SE Estimate	Z	p-value
Intercept	-2.292465	0.205433	-11.159160	0.000000
Trait	0.258437	0.023228	11.126054	0.000000
Group	-0.275756	0.409650	-0.673151	0.501007
Trait x Group	-0.063274	0.046456	-1.362018	0.173499

Item Number: 18	Estimate	SE Estimate	Z	p-value
Intercept	-2.302057	0.248337	-9.269880	0.000000
Trait	0.362570	0.027196	13.331932	0.000000
Group	0.666170	0.449435	1.482239	0.138592
Trait x Group	0.025536	0.054391	0.469481	0.638829

Item Number: 19	Estimate	SE Estimate	Z	p-value
Intercept	-3.195740	0.253713	-12.595879	0.000000

Trait	0.444093	0.029981	14.812284	0.000000
Group	-0.095322	0.478530	-0.199197	0.842149
Trait x Group	0.055611	0.059963	0.927421	0.353932

Item Number: 20	Estimate	SE Estimate	Z	p-value
Intercept	-2.714549	0.205425	-13.214307	0.000000
Trait	0.267331	0.023098	11.573544	0.000000
Group	-0.797595	0.417516	-1.910335	0.056376
Trait x Group	0.052308	0.046197	1.132289	0.257784

Item Number: 21	Estimate	SE Estimate	Z	p-value
Intercept	-2.830913	0.219759	-12.881900	0.000000
Trait	0.258345	0.024209	10.671417	0.000000
Group	0.692059	0.450976	1.534583	0.125203
Trait x Group	-0.049844	0.048418	-1.029443	0.303521

EK 4
CALCBİAS PROGRAMI ANALİZ ÇIKTILARI

KANADA (İNGİLİZCE-FRANSIZCA FORM)

item	zadiff	zbdiff	diff	area	wdiff	warea
1	9999	9999	-0,10	0,24	0	0
2	9999	9999	3,30	3,69	0	0
3	9999	9999	0,63	0,63	0	0
4	9999	9999	0,45	1,68	0	0
5	9999	9999	0,05	0,07	0	0
6	9999	9999	0,28	2,12	0	0
7	9999	9999	0,16	0,43	0	0
8	9999	9999	0,54	0,60	0	0
9	9999	9999	0,03	0,16	0	0
10	9999	9999	0,08	0,08	0	0
11	9999	9999	0,47	0,89	0	0
12	9999	9999	0,35	0,91	0	0
13	9999	9999	0,38	0,59	0	0
14	9999	9999	0,31	0,40	0	0
15	9999	9999	0,13	0,65	0	0
16	9999	9999	0,55	0,56	0	0
17	9999	9999	1,05	1,67	0	0
18	9999	9999	-0,39	0,50	0	0
19	9999	9999	-0,02	0,51	0	0
20	9999	9999	0,20	2,92	0	0
21	9999	9999	0,16	0,40	0	0

AVUSTRALYA-İNGİLTERE

item	zadiff	zbdiff	diff	area	wdiff	warea
1	9999	9999	0,23	0,54	0	0
2	9999	9999	-4,24	5,64	0	0
3	9999	9999	-0,06	0,37	0	0
4	9999	9999	-0,31	0,32	0	0
5	9999	9999	-0,39	0,42	0	0
6	9999	9999	-0,53	0,58	0	0
7	9999	9999	-0,28	0,38	0	0
8	9999	9999	-0,48	0,50	0	0
9	9999	9999	-0,38	0,42	0	0
10	9999	9999	-0,22	0,27	0	0
11	9999	9999	-0,25	0,25	0	0
12	9999	9999	-0,26	0,67	0	0
13	9999	9999	-0,68	0,86	0	0
14	9999	9999	-0,55	0,69	0	0
15	9999	9999	-0,65	0,66	0	0
16	9999	9999	-0,37	3,36	0	0
17	9999	9999	-0,30	0,30	0	0
18	9999	9999	0,44	0,54	0	0
19	9999	9999	-0,04	0,10	0	0
20	9999	9999	-0,80	0,80	0	0
21	9999	9999	-0,25	0,25	0	0

TURKİYE İNGİLTERE 1. KİTAPÇIK

item	zadiff	zbdiff	diff	area	wdiff	warea
1	9999	9999	-24,00	30,95	0	0
2	9999	9999	34,84	39,89	0	0
3	9999	9999	2,68	3,00	0	0
4	9999	9999	-0,26	1,37	0	0
5	9999	9999	-0,15	3,13	0	0
6	9999	9999	-1,71	1,71	0	0
7	9999	9999	1,48	1,78	0	0
8	9999	9999	0,57	1,31	0	0
9	9999	9999	1,82	1,82	0	0
10	9999	9999	1,62	1,62	0	0
11	9999	9999	0,18	4,27	0	0
12	9999	9999	0,78	0,78	0	0
13	9999	9999	-1,23	1,23	0	0
14	9999	9999	-1,26	1,99	0	0
15	9999	9999	2,30	2,84	0	0
16	9999	9999	1,14	1,29	0	0
17	9999	9999	34,69	39,51	0	0
18	9999	9999	35,92	40,59	0	0
19	9999	9999	35,97	40,98	0	0

TURKIYE İNGİLTERE 5. KİTAPÇIK

item	zadiff	zbdiff	diff	area	wdiff	warea
1	9999	9999	0,93	0,93	0	0
2	9999	9999	2,82	2,82	0	0
3	9999	9999	1,03	1,39	0	0
4	9999	9999	2,02	2,74	0	0
5	9999	9999	0,73	0,88	0	0
6	9999	9999	0,12	0,14	0	0
7	9999	9999	0,54	0,54	0	0
8	9999	9999	0,71	0,72	0	0
9	9999	9999	-0,28	0,41	0	0
10	9999	9999	0,68	0,86	0	0
11	9999	9999	0,18	2,30	0	0
12	9999	9999	0,10	3,57	0	0
13	9999	9999	0,50	0,50	0	0
14	9999	9999	3,05	8,68	0	0
15	9999	9999	0,72	1,72	0	0
16	9999	9999	0,22	3,17	0	0
17	9999	9999	-0,41	0,84	0	0
18	9999	9999	1,46	1,52	0	0
19	9999	9999	0,87	1,02	0	0
20	9999	9999	0,59	1,52	0	0
21	9999	9999	1,14	1,16	0	0

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Nezaket Bilge BAŞUSTA
Doğum Yeri ve Tarihi : Kahramanmaraş 24\11\1982

Eğitim Durumu

Lisans Öğrenimi : Hacettepe Üniversitesi Fizik Öğretmenliği
Yüksek Lisans : Hacettepe Üniversitesi Fizik Öğretmenliği (Tezsiz)
Öğrenimi : Hacettepe Üniversitesi Ölçme-Değerlendirme (Tezli)
Doktora Öğrenimi : Hacettepe Üniversitesi Ölçme-Değerlendirme
Bildiği Yabancı Diller : İngilizce, Almanca

İş Deneyimi

Çalıştığı Kurumlar : Hacettepe Üniversitesi Tıp Eğitimi ve Bilişimi Anabilim
Dalı (2006-)

İletişim

E-Posta Adresi : buzun@hacettepe.edu.tr