



HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

TEOG SINAVININ YABANCI DİL ALT TESTİNE AİT MADDELERİN YANLILIĞININ İNCELENMESİ

Münevver ARSLAN

Yüksek lisans tezi

Ankara, 2020

Liderlik, arařtırma, inovasyon, kaliteli eđitim ve deđiřim ile

Daha ileriye ... En İyiyeye ...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

TEOG SINAVININ YABANCI DİL ALT TESTİNE AİT MADDELERİN
YANLILIĞININ İNCELENMESİ

INVESTIGATION OF ITEM BIAS FOR ENGLISH SUBTEST OF TEOG

Münevver ARSLAN

Yüksek lisans tezi

Ankara, 2020

Öz

Bu çalışmada, TEOG 2015 - 2016 Eğitim - Öğretim Yılı II. Dönem İngilizce alt testine ait maddeler, öğrencilerin cinsiyet ve öğrenim gördüğü okul türüne (özel-devlet) göre değişen madde fonksiyonu açısından incelenmiştir. Değişen madde fonksiyonu tespit edilen maddeler yanlılık gösterip göstermediğinin belirlenmesi amacıyla uzman görüşüne sunulmuştur. Araştırmanın Değişen Madde Fonksiyonlaşması (DMF) analizi aşamasında Mantel-Haenszel (MH), Lojistik Regresyon (LR) ve SIBTEST yöntemleri kullanılmıştır. Analizler için kullanılan veri seti, İngilizce alt testi maddelerine cevap veren 1.133.857 sekinci sınıf öğrencisinden rastgele seçilerek oluşturulan 9.108 öğrencinin 20 maddelik test için verdiği cevapları kapsamaktadır. DMF analizi öncesinde İngilizce alt testinin tek boyutluluğu incelenmiş ve cinsiyet ve okul türü değişkenine göre oluşturulan her alt teste ait betimsel istatistikler, güvenilirlik analizleri ve madde istatistikleri hesaplanmıştır. DMF analizi aşamasında ise önce cinsiyete göre analizler gerçekleştirilmiş olup 5, 9, 11 ve 16 numaralı maddelerin B (orta) düzeyinde DMF gösterdiği tespit edilmiştir. Bu maddelerden 5 numaralı olan sadece SIBTEST, 9 ve 11 numaralı olanlar hem MH hem SIBTEST ve 16 numaralı maddenin ise sadece MH yöntemine göre DMF gösterdiği belirlenmiştir. Bu maddeleri yanlılık özelliği bakımından yorumlayan uzmanlar, sadece 11 numaralı maddenin erkekler lehine avantaj sağlayabileceğini belirtmiştir. Okul türü değişkenine göre yapılan analizlere göre de sadece 4 numaralı maddede MH ve SIBTEST yöntemlerine göre B (orta) düzeyinde DMF tespit edilmiştir. Bu madde hakkında uzmanların yorumu ise maddenin özel okul öğrencileri lehine avantaj sağladığı yönündedir. Analiz yöntemlerinin uyumu incelendiğinde ise hesaplanan spearman rho katsayısına göre MH ve SIBTEST yöntemlerinin birbiriyle uyumunun yüksek, LR yöntemi ile uyumunun ise düşük olduğu sonucuna ulaşılmıştır.

Anahtar sözcükler: TEOG, değişen madde fonksiyonu, madde yanlılığı, Mantel Haenszel yöntemi, Lojistik Regresyon yöntemi, SIBTEST yöntemi

Abstract

In this study, differential functioning items according to gender and school type (private vs public) in English subtest of TEOG in 2015 – 2016 Academic Year Spring Semester were investigated. Expert opinions were sought for the items, that were found as showing differential item functioning, in order to determine whether there was an item bias. Mantel-Haenszel (MH), Logistic Regression (LR) and SIBTEST methods were used in the study of the Differential Item Functioning (DMF) analysis. The data includes the answers in 20 items English test given by 9108 8th grade students randomly selected from total 1133857 students. Before the analysis, the unidimensionality of the English subtest was examined. Then, descriptive statistics, reliability analyzes and item statistics were calculated for each subtest created according to gender and type of school. In the DMF analysis, firstly gender-based analyzes were carried out and it was revealed that items 5, 9, 11 and 16 differed at B (medium) level. Among these items, item 5 showed DIF in only SIBTEST, item 9 and 11 showed DIF in both MH and SIBTEST, and item 16 showed DIF in only MH. Based on expert opinions, it was decided that only item 11 might show bias in favor of male students. According to the analysis made based on school type variable, only item 4 was determined as showing B level DMF in both MH and SIBTEST methods. Experts opinions on this item indicated that this item might show bias in favor of private school students.

Keywords: TEOG, differential item functioning, item bias, Mantel Haenszel method, Lojistik Regresyon method, SIBTEST method

Eşim Okan ve Oğlum Uraz'a...

Teşekkür

Tez çalışma süresince bilgisini, desteğini ve güvenini esirgemeyen tüm yoğunluğuna rağmen değerli vaktini bana ayıran değerli hocam Prof. Dr. Nuri Doğan'a, tez jürimde yer alan ve dönütleriyle tezime katkı sağlayan değerli jüri üyelerim Prof. Dr. Selahattin GELBAL'a, ve Doç. Dr. Deha DOĞAN'a;

Hayatım boyunca yaptığım her işte bana tereddütsüz güvendiklerini tüm kalbimle hissettiğim, ve verdikleri bu güç sayesinde hayata karşı çok daha dik durabilmemi sağlayan, varlıklarının ve sevgilerinin bedelini ödeyemeyeceğim babam Özdemir SAYGILI ve annem Yasemin SAYGILI'ya, yardımlarını hiç eksik etmeyen her ihtiyacım olduğunda yanımda olan abim Cengizhan SAYGILI'ya ve tez yazım sürecinde çocuğumu güvenle emanet ettiğim ve hep yanımda olduklarını bildiğim ailem, Fatma Birgül ARSLAN ve Rüstem ARSLAN'a;

Gerek anne karnında gerekse kucağımda yanımda olan, bazen annesinden mahrum kalan canım oğlum Uraz ARSLAN'a;

Yoğun çalışma dönemimde yaşadığım tüm zorluklarda sabır ve anlayışı ile beni hep destekleyen, hep yanımda olan, fedakarlıklarını göz ardı edemeyeceğim, sevgili eşim Okan Arslan'a

teşekkür ederim.

İçindekiler

Öz.....	ii
Abstract.....	iii
Teşekkür.....	v
Tablolar Dizini.....	ix
Şekiller Dizini.....	xi
Simgeler ve Kısaltmalar Dizini.....	xii
Bölüm 1 Giriş.....	1
Problem Durumu.....	1
Araştırmanın Amacı ve Önemi.....	6
Araştırma Problemi.....	8
Sayıltılar.....	9
Sınırlılıklar.....	9
Tanımlar.....	10
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	11
Araştırmanın Kuramsal Temeli.....	11
İlgili Araştırmalar.....	27
İlgili Araştırmalar Özet.....	38
Bölüm 3 Yöntem.....	41
Araştırmanın Türü.....	41
Evren ve Örneklem.....	41
Veri Toplama Süreci.....	42
Veri Toplama Araçları.....	42
Verilerin Analizi.....	52
Bölüm 4 Bulgular ve Yorumlar.....	57
Alt Problem 1' e İlişkin Bulgular ve Yorumlar.....	57
Alt Problem 2' ye İlişkin Bulgular ve Yorumlar.....	65

Alt Problem 3' e İlişkin Bulgular ve Yorumlar	71
Bölüm 5 Sonuç, Tartışma ve Öneriler	76
Sonuç.....	76
Alt Problem 1' e İlişkin Tartışmalar.....	77
Alt Problem 2' e İlişkin Tartışmalar.....	83
Alt Problem 3' e İlişkin Tartışmalar.....	85
Öneriler	88
Kaynaklar	91
EK-A: TEOG Sınavına Ait İngilizce Alt Testi.....	107
EK-B: Uzman Görüşü Anket Formu	111
EK-C: İngilizce Alt Testine Ait Madde Faktör Yükleri.....	119
EK-Ç: Cinsiyete Göre İngilizce Alt Testi Maddelerinin MH Analizi İçin Yapılan EZDIF Program Çıktısı	120
EK-D: Cinsiyete Göre İngilizce Alt Testi Maddelerinin LR Analizi İçin Yapılan EZDIF Program Çıktısı	121
EK-E: Zumbo (1999)'nun Syntaxı.....	123
EK-F: Cinsiyete Göre İngilizce Alt Testi Maddelerinin SIBTEST Analizi İçin Yapılan SIBTEST Program Çıktısı.....	124
EK-G: Okul Türüne Göre İngilizce Alt Testi Maddelerinin MH Analizi İçin Yapılan EZDIF Program Çıktısı	126
EK-Ğ: Okul Türüne Göre İngilizce Alt Testi Maddelerinin LR Analizi İçin Yapılan EZDIF Program Çıktısı	127
EK-H: Okul Türüne Göre İngilizce Alt Testi Maddelerinin SIBTEST Analizi İçin Yapılan SIBTEST Program Çıktısı	129
EK-I: Etik Komisyon İzin Muafiyeti Formu.....	131
EK-İ: Etik Beyanı	132
EK-J: Yüksek Lisans Tez Çalışması Orijinallik Raporu	133
EK-K: Thesis Originality Report.....	134

EK-L: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı	135
--	-----

Tablolar Dizini

Tablo 1 <i>Potenza ve Dorans (1995)'a Ait İkili ve Çok Kategorili DMF Yöntemlerinin Parametrik Olup Olmama Durumlarına Göre Çapraz Sınıflandırması</i>	13
Tablo 2 <i>Eşleştirilen Odak ve Referans Gruba Ait Veri Düzeni</i>	16
Tablo 3 <i>Eşleştirilen Odak ve Referans Gruba Ait Olasılık Tablosu</i>	17
Tablo 4 <i>$\Delta\alpha_{MH}$ Değerlerini Yorumlama Ölçütleri</i>	18
Tablo 5 <i>Zumbo ve Thomas (1996), Jodoin ve Gierl (2001) ve Bakan-Kalaycıoğlu ve Kelecioğlu (2011) Tarafından Düzenlenen ΔR^2 Değerlerini Yorumlama Ölçütleri</i>	23
Tablo 6 <i>β Değerlerini Yorumlamak için Roussos ve Stout (1996) Tarafından Önerilen Sınıflama Ölçütleri</i>	25
Tablo 7 <i>Örneklem Grubunun Öğrenci Sayıları ve Yüzdeleri</i>	42
Tablo 8 <i>Örneklem ve Alt Gruplara Ait Basıklık ve Çarpıklık Değerleri</i>	46
Tablo 9 <i>İngilizce Alt Testinin Test İstatistikleri</i>	47
Tablo 10 <i>Evren ve Örnekleme Ait İngilizce Alt Testi Madde İstatistikleri</i>	48
Tablo 10 <i>Evren ve Örnekleme Ait İngilizce Alt Testi Madde İstatistikleri</i>	49
Tablo 11 <i>Cinsiyet ve Okul Türü Gruplarına Göre Madde İstatistikleri</i>	50
Tablo 12 <i>Öz Değerler Tablosu</i>	53
Tablo 13 <i>İngilizce Alt Testi Maddelerinin Cinsiyete Göre MH Analiz Sonuçları</i>	57
Tablo 13 <i>İngilizce Alt Testi Maddelerinin Cinsiyete Göre MH Analiz Sonuçları</i>	58
Tablo 14 <i>Cinsiyete Göre MH Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi</i>	58
Tablo 15 <i>İngilizce Alt Testi Maddelerinin Cinsiyete Göre LR Analizi β Estimate Değerleri</i>	59
Tablo 16 <i>İngilizce Alt Testi Maddelerinin Cinsiyete Göre LR Analiz Sonuçları</i>	60
Tablo 17 <i>İngilizce Alt Testi Maddelerinin Cinsiyete Göre SIBTEST Analiz Sonuçları</i>	61
Tablo 18 <i>Cinsiyete Göre SIBTEST Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi</i>	62
Tablo 19 <i>5 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi</i>	63
Tablo 20 <i>9 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi</i>	63
Tablo 21 <i>11 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi</i>	64
Tablo 22 <i>16 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi</i>	64
Tablo 23 <i>İngilizce Alt Testi Maddelerinin Okul Türüne Göre MH Analiz Sonuçları</i>	65

Tablo 23 <i>İngilizce Alt Testi Maddelerinin Okul Türüne Göre MH Analiz Sonuçları</i>	66
Tablo 24 <i>Okul Türüne Göre MH Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi</i>	66
Tablo 25 <i>İngilizce Alt Testi Maddelerinin Okul Türüne Göre LR Analizi β Estimate Değerleri</i>	67
Tablo 26 <i>İngilizce Alt Testi Maddelerinin Okul Türüne Göre LR Analiz Sonuçları</i>	68
Tablo 27 <i>İngilizce Alt Testi Maddelerinin Okul Türüne Göre SIBTEST Analiz Sonuçları</i>	69
Tablo 28 <i>Okul Türüne Göre SIBTEST Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi</i>	70
Tablo 29 <i>4 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi</i>	70
Tablo 30 <i>Cinsiyete Göre MH, LR ve SIBTEST Yöntemlerine Ait Analiz Bulgularının Karşılaştırılması</i>	71
Tablo 31 <i>Cinsiyete Göre MH, LR ve SIBTEST Yöntemlerinin Spearman Rho Katsayı Değerleri</i>	73
Tablo 32 <i>Okul Türüne Göre MH, LR ve SIBTEST Yöntemlerine Ait Analiz Bulgularının Karşılaştırılması</i>	73
Tablo 33 <i>Okul Türüne Göre MH, LR ve SIBTEST Yöntemlerinin Spearman Rho Katsayı Değerleri</i>	75

Şekiller Dizini

Şekil 1. Tek Biçimli DMF (TB DMF) grafiği.	20
Şekil 2. Tek Biçimli Olmayan DMF (TBO DMF) grafiği.	20
Şekil 3. Yabancı dil (İngilizce) alt testinin evrene ait puan dağılımı.	44
Şekil 4. Yabancı dil (İngilizce) alt testinin örnekleme ait puan dağılımı.	44
Şekil 5. Cinsiyet değişkenine göre İngilizce alt testi puan dağılım grafiği.	45
Şekil 6. Okul türü değişkenine göre İngilizce alt testi puan dağılım grafiği.	45
Şekil 7. KMO ve Bartlett Kürsellik Testi sonuçları.	52
Şekil 8. Yamaç birikinti grafiği.	53
Şekil 9. TEOG İngilizce alt testine ait 5 numaralı madde.	77
Şekil 10. TEOG İngilizce alt testine ait 9 numaralı madde.	78
Şekil 11. TEOG İngilizce alt testine ait 11 numaralı madde.	80
Şekil 12. TEOG İngilizce alt testine ait 16 numaralı madde.	82
Şekil 13. TEOG İngilizce alt testine ait 4 numaralı madde.	84

Simgeler ve Kısaltmalar Dizini

AFA: Açımlayıcı Faktör Analizi

DFA: Doğrulayıcı Faktör Analizi

DMF: Değişen Madde Fonksiyonu

ETS: Educational Testing Sevice

KTK: Klasik Test Kuramı

LR: Lojistik Regresyon

MEB: Milli Eğitim Bakanlığı

MH: Mantel –Haenszel

MKE: Madde Karakteristik Eğrisi

MTK: Madde Tepki Kuramı

SIBTEST: Simultaneous Item Bias Test

TB: Tek biçimli

TBO: Tek Biçimli Olmayan

YİBO: Yatılı İlköğretim Bölge Okulu

Bölüm 1

Giriş

Bu bölümde bu çalışmanın yapılmasını gerekli kılan problem durumuna, araştırmanın amacı ve önemine, çalışmanın düzen içinde yapılmasını sağlayan problem cümlelerine, sayıltı ve sınırlılıklara yer verilmiştir.

Problem Durumu

Öğrenme–öğretme sürecinde öğrencilerin kazandığı bilgi ve becerilerin düzeyinin belirlenmesi amacıyla gerçekleştirilen ölçme ve değerlendirme çalışmaları, öğrenme sürecinin verimliliğinin değerlendirilmesi aşamasında çok önemli bir yere sahiptir. Bu nedenle ülkeler eğitim politikalarını hazırlarken ölçme ve değerlendirme çalışmalarına önem vermektedirler. Bu çalışmalar, aynı zamanda eğitim sisteminde yapılan reformlara da geri bildirim sağlamaktadır.

Türkiye gibi gelişmekte olan ve kalabalık nüfusa sahip ülkelerde uygulanan merkezi sınavlar, eğitimin her kademesinde geçmişten günümüze eğitim sistemlerinin değişilmez bir parçasını oluşturmaktadır. İlköğretimden ortaöğretime geçiş aşamasında merkezi sınav olarak ilk 1998 yılında OKÖSYS uygulanmaya başlamıştır. Yapılan değerlendirme çalışmaları ile yıllar geçtikçe ihtiyaçlar doğrultusunda bu merkezi sınavların (LGS, SBS, TEOG, LGS (2018)) sistemselsel olarak değişime uğradığı görülmüştür. Sınav sisteminin içeriği ve adı değişse de değişmeyen tek özelliğinin çoktan seçmeli test maddelerinin kullanılması olduğu gözden kaçmamaktadır. Çünkü çoktan seçmeli test maddeleri ile diğer tekniklere oranla daha fazla davranış yoklanabilmekte ve değerlendirme işlemi daha kısa sürede ve objektif olarak gerçekleştirilebilmektedir (Turgut ve Baykul, 2010). Ayrıca istatistiksel analiz imkânı sağlayan bu maddeler, hem ölçme aracının niteliğinin değerlendirilmesini hem de varsa eksiklerinin giderilmesi konusunda da yetkili mercilere geri bildirim imkânı sağlama konusunda da avantajlıdır.

Test puanlarından faydalanarak bireyler hakkında karar verme özelliği taşıyan merkezi sınavların, ölçme sonuçlarında olması gereken özellikleri barındıracak şekilde titizlikle hazırlanması önemlidir. Bu özellikler, ölçme aracının amaca uygun bir şekilde sadece ölçülmek istenen özelliği ölçmesi ve ölçme aracının aynı koşullarda tekrarlandığında da aynı sonuçları verebilmesidir (Kan, 2007a).

Ölçme aracının tutarlı ve kararlı sonuçlar verebilmesi ölçmenin mümkün olduğunca hatasız bir şekilde yapılmasını gerekli kılmaktadır (Crocker ve Algina, 1986). Bu durum testin güvenilirliği ile ilgilidir. Eğitim ve psikolojide, ölçmede hatanın sıfır olmasının neredeyse imkânsız olması yüzde yüz güvenilir bir test hazırlanması durumunu engellemektedir (Fraenkel ve Wallen, 2007). Bu nedenle yapılan çalışmaların çoğu hata kaynaklarının tespit edilerek önlenmesine yöneliktir.

Ölçme işlemine karışan hatalar; ölçmeciden, ölçme koşullarından, cevaplayıcılardan ya da ölçme aracından kaynaklanabilir. Buradaki amaç, bu kaynakların ölçme sonucunu ne düzeyde etkilediğini tespit edip olabildiğince önlem almak olmalıdır. Eğitim ve psikoloji alanında en çok karşılaşılan kaynağının ve yönünün tespit edilemediği tesadüfi hatalar, maalesef bu konuda önlem almayı zorlaştırmaktadır. Bu aşamada alan yazında bu konuda yapılan çok sayıdaki çalışma sayesinde hata kaynakları tespit edilerek önleyici çalışmalar yapılabilir. Hata kaynakları; grup homojenliği, test uzunluğu, sınav süresi, olumsuz sınav ortamı, cevaplayıcı motivasyonu, maddelerin cevaplayıcılara yönelik açık ve anlaşılır şekilde ifade edilmemesi, şans başarısı gibi çoğaltılabilmektedir (Demircioğlu, 2011). İfade edilen kaynaklara dikkat edilerek özenle gerçekleştirilen bir ölçme söz konusu olduğunda bireylerin becerileri hakkında daha güvenilir çıkarımlar yapılabilir.

Crocker ve Algina (1986), ölçmenin güvenilir olmasının bireyler hakkında çıkarım yapmak için tek başına yeterli olmadığını ve test sonuçlarının geçerlik özelliğini de taşıması gerektiğini ifade etmişlerdir. Bununla birlikte, geçmişten günümüze geçerlik kavramının tanımı hakkında ortak bir anlayışa varılamadığı alan yazında görülebilmektedir (Kelecioğlu ve Göçer-Şahin, 2014). Örneğin, Thorndike ve Hagen (1961) geçerlik kavramını, ölçme aracının sadece ölçmeyi amaçladığı özelliği ölçebilme yeterliliği olarak tanımlamıştır (Akt. Baykul, 2010). Cronbach (1984) ise ölçmecinin, test puanlarından elde ettiği çıkarımları desteklemek için delil toplama süreci olarak ifade etmiştir. Öte yandan Fraenkel ve Wallen (2007) araştırmacıların elde ettiği çıkarımların uygunluğu, anlamlılığı, doğruluğu ve kullanılabilirliği olarak açıklamıştır. Tüm bu tanımlar dikkate alındığında araştırmacıların geçerlik kavramına farklı açılardan yaklaştığı görülmüştür. Bu durum da tek bir geçerlik değil, geçerlik yaklaşımlarının olduğunu ortaya koymaktadır (Kan, 2007a). Bu sebeple ölçmecinin, elindeki ölçme aracı için hangi

geçerlik türüne dikkat edeceğine, test puanlarını kullanım amacına göre sürecin başında karar vermesi beklenir (Baykul, 2010; Crocker ve Algina, 1986). Merkezi sınavların amacının öğrenci başarısının değerlendirilmesi olduğu bilindiğine göre, ölçme aracının ilk etapta kazanılması beklenen tüm davranışları temsil etmesi beklenir (Fraenkel ve Wallen, 2007). Bu ifade kapsam geçerliğini işaret etmektedir. Kapsam geçerliğinin yüksek olduğu bir testin hazırlanabilmesi için bazı koşulların dikkate alınması gerekmektedir. Bunun için test hazırlama sürecinde öncelikle ölçülmek istenen davranışları ve düzeylerini ifade eden belirtke tablolarının hazırlanması önemlidir (Baykul, 2010). Daha sonra belirtke tablolarının yardımıyla maddelerin yazımı işlemine geçilir. Bu aşamada maddeler yazılırken maddenin, herhangi bir şekilde farklı etnik grup ve kültürel geçmiş, dil veya cinsiyet gibi özelliklerin madde performansını etkilememesine özen gösterilmelidir (Crocker ve Algina, 1986).

İlköğretimden ortaöğretime geçişte uygulanan merkezi sınavlarda olmasının beklendiği bir diğer özellik ise, test puanlarının ölçülmek istenen davranış bakımından öğrencinin gelecekteki performansını yordayabilmesidir (Crocker ve Algina, 1986). Yordama, ölçüt geçerliği kapsamında açıklanan bir kavramdır. Bireylerin geleceği hakkında karar verilmesi gereken ölçmelerde test puanlarının ölçüt puanlarını yordayabilme ölçüsü yordama geçerliği olarak ifade edilir (Doğan-Başokçu ve Doğan, 2005).

Ölçme aracının ölçüt ve kapsam geçerliğinin istenen seviyede olduğu tespit edildikten sonra, aracın yapı geçerliğinin de incelenmesi gerekmektedir (Cronbach ve Meehl, 1955, akt. Kelecioğlu ve Göçer-Şahin, 2014). Yapı geçerliği kavramı, Lord ve Novick (1968) tarafından ölçülmek istenen yapının belirlenen ölçme aracı ile ölçülebilir derecesi olarak tarif edilmiştir (Akt. Baykul, 2010). Ölçme aracı, ölçülmek istenen yapının dışında başka bir değişkeni ve/veya özelliği de istemsiz olarak ölçüyorsa bu durum ölçme aracının yapı geçerliğini olumsuz yönde etkilemektedir (Doğan ve Öğretmen, 2005). Örneğin bir başarı testinde, testi alan kişilerinin herhangi bir özelliği (cinsiyet, dil, kültür vb.) maddeyi doğru cevaplama ihtimalini olumlu veya olumsuz yönde etkiliyorsa, bu durum testin yapı geçerliğini tehdit etmektedir. Dolayısıyla, ölçme aracında sistematik hataya yol açmaktadır (Osterlind, 1983). Burada sistematik hata kavramı, ölçmeciyeye, ölçülen özelliğe ve ölçme koşullarına bağlı olarak miktarı değişen ve tüm ölçmelere aynı oranda etki

etmeyen bir hata türü olarak tanımlanmaktadır (Turgut, 1995). Sistematik hatalar belirli bir grup için yanlılık oluşturur ve geçerliği tehdit etmektedir (Camilli 2006; Turgut ve Baykul, 2010). Burada yanlılık, belli bir özellikten ötürü oluşturulan alt grupların özellikleri ile ilgilenmeden testin ölçmek istediği özelliği ölçüyor mu sorusuna cevap aramaktır (Shepard, Camilli ve Williams, 1984). Yanlılığa sebep olabilecek olası kaynakların iyi tespit edilmesi gerekmektedir. Aksi halde, testin tarafsızlığı, dolaylı olarak da geçerliği olumsuz yönde etkilenmektedir (Clauser ve Mazor, 1998; Kristanjansson, Aylesworth, McDowell ve Zumbo, 2005).

Yanlılık çalışmalarının istatistiksel kısmını Değişen Madde Fonksiyonu (DMF) analizleri oluşturmaktadır. Değişen madde fonksiyonu ise bir gruptaki bireylerin, maddeye doğru cevap verme olasılığının benzer yetenek düzeyinde başka bir gruptaki bireylerden daha fazla olması durumunu tanımlamak için kullanılan istatistiksel bir terimdir (Hambleton ve Rogers, 1989; Zumbo, 2007). Bireylerin sahip oldukları bazı özellikler nedeniyle test maddeleri bireylerin beceri düzeylerinin farklılaşmasına kaynak oluşturuyorsa bu özelliklerin çok iyi belirlenip madde yazım sürecinin bu koşullar altında daha özenle gerçekleştirilebilmesi için çeşitli önlemler alınmalıdır (Zenisky, Hambleton ve Robin, 2004). Bu amaçla DMF'ye neden olan olası kaynakların (dil, ırk, kültür, cinsiyet, sosyo-ekonomik düzey vb.) belirlenmesi gerekmektedir. Örneğin alan yazında, en çok karşılaşılan olası kaynak, cinsiyet farklılığıdır (Azen, Bronner ve Gafni, 2010; Doolittle ve Cleary, 1987; Henderson, 2001; Kalaycıoğlu ve Kelecioğlu, 2011; Karakaya, 2012; Mendes-Barnett ve Ercikan, 2006;). Özellikle başarı testlerinde yapılan çalışmalarda gerek ilgi alanı (bkz. Stricker ve Emmerich, 1999) gerekse beceri türünden (bkz. O'Neill ve McPeck, 1993) ötürü cinsiyete göre farklılaşmalara rastlanmaktadır. Örneğin; görsel şekil, diyagram veya mekânsal işlem gerektiren maddelerde erkek öğrencilerin daha başarılı olduğu daha önce yapılan çalışmalarda tespit edilmiştir (Benbow ve Stanley, 1982; Halpern, 1997; Ingels, 1990; Lohman, 1993). Educational Testing Service (1998), erkeklerin özellikle spor dalları ve askeri konularla ilgili maddelerde kadınlara oranla daha başarılı olduğunu tespit etmiş olup madde yazım sürecinde olabildiğince bu alanlardan uzaklaşılması konusunda test geliştiricileri uyarmaktadır. Ayrıca Hamilton (1999), çalışmasında yaptığı görüşmelerde okul dışında edinilen bilgilerin yine erkek öğrenciler tarafından ilgi çekici bulunması nedeniyle bu konuları içeren maddelerde daha başarılı olduklarını tespit etmiştir. Fennema ve Sherman

(1997) ise aynı şekilde bireylerin hayatları boyunca biriktirdikleri tutumların, ders dışı etkinliklerin veya sosyo-kültürel etkinliklerinin öğrenme sürecinin arka planında özellikle çok etkili olduğunu ifade ederek bu durumu desteklemektedir. İnsan ilişkileri içerikli sosyal alan konularında ise kadınların daha başarılı olduğunu ifade eden çalışmalara rastlanmaktadır (O'Neill ve McPeck, 1993; Stricker, 1981). Ek olarak kadınların estetik, insan hakları ve sembol içeren öğeler konusunda da daha başarılı sonuçlar elde ettiği çalışmalar vardır (Doolittle ve Cleary, 1987; Harris ve Carlton, 1993; O'Neil ve McPeck, 1993). Yine çalışmalar gösteriyor ki, açık uçlu madde içeren testlerde kadınlar erkeklere oranla daha başarılı iken çoktan seçmeli maddelerde ise erkekler kadınlara oranla daha başarılı sonuçlar elde etmektedir (Zenisky, Hambleton ve Robin, 2004). Bu çalışmalar, kadınların sözel ve dilsel yetenekleri daha iyi kullandığı fikrini desteklemektedir. Özellikle yabancı dil eğitimi söz konusu olduğunda aynı yaştaki kadın ve erkek öğrencilerinin dil gelişimlerinin aynı seviyede olmadığı düşünülürse bu konuda araştırma yapılmasının gerekliliği su götürmez bir ihtiyaçtır.

Alan yazında çok fazla rastlanmasa da eğitimde eşitlik ilkesi adına araştırılması gereken bir diğer olası kaynak ise okul türüdür. Örneğin, özel okul imkânlarının devlet okullardan daha iyi olduğu bilinen bir gerçektir. Peki, yapılan ulusal sınavlar eğitimde eşitlik ilkesini göz önünde bulunduracak şekilde maddelere sahip midir? Karakaya ve Kutlu (2012), çalışmalarında SBS verilerini kullanarak Türkçe başarısını karşılaştırmış ve maddelerin özel okul öğrencilerine avantaj sağladığını tespit etmiştir. Kelecioğlu, Karabay ve Karabay (2014) ise okul türü kapsamını genişleterek YİBO'ları da dâhil ettikleri çalışmalarında yine benzer sonuca ulaşarak SBS maddelerinin özel okul öğrencilerine avantaj sağladığını tespit etmiştir. Gök, Kelecioğlu ve Doğan (2010), çalışmalarında her ne kadar yöntem karşılaştırması yapmayı amaçlasalar da alt grup oluşturma kriteri olarak kullandıkları okul türü değişkeninde yine OKS maddelerin özel okul öğrencilerinin lehine olduğunu tespit etmiştir. Bu çalışmalar da gösteriyor ki, okul türü değişkeni madde yazımı sürecinde dikkate alınması gereken önemli bir detaydır. Zaten bu sebeplerdir ki ebeveynler de imkânlarını zorlayarak çocuklarını özel okullara göndermeyi tercih etmektedirler. Örneğin sosyo-ekonomik düzeyi yüksek olan bireyler sahip oldukları imkânlar sayesinde maddeye doğru cevap verme olasılıklarını yükseltebiliyorlarsa bu durum sınav açısından istenmeyen durumlar oluşturur. Günümüzde de

insanlarda böyle bir algı olması nedeniyle imkânı olan ebeveynler çocuklarını özel okullara gönderme eğiliminde oluyorlar. Özellikle yabancı dil eğitiminin özel okullarda daha iyi verildiği düşüncesi bu davranışı pekiştiriyor (Hesapçioğlu ve Nohutçu, 1999). Özel okullarda yabancı dil eğitimi sırasında devlet okullarına oranla daha çeşitli ders araç gereçleri kullanmaları, etütler sayesinde ders saatlerinin fazla olması, sınıf mevcutlarının çok olmaması sayesinde öğrencinin sürece daha çok dâhil olması gibi nedenler öğrenme sürecine katkı sağlamaktadır (Nartgün ve Kaya, 2016). Bu koşulların merkezi sınavlarda öğrencilerin başarılarında farklılık oluşturabilir olması, araştırılması gereken bir problem olarak görülmektedir.

Bu sebeplerden ötürü, bireylerin hayatları hakkında karar verme yetkisi taşıyan merkezi sınavların DMF veya yanlılık göstere maddeler içermemesine dikkat edilmeli ve varsa tespit edilip düzeltilmeli ya da testten çıkartılmalıdır (Kan, 2007b). Kısacası, yanlı madde içeren testler, merkezi sınavların geçerliği konusunda önemli bir problem durumu oluşturmaktadır.

Araştırmanın Amacı ve Önemi

Bu araştırmanın amacı; TEOG 2015 – 2016 Eğitim – Öğretim Yılı II. Dönem İngilizce alt testine ait maddelerin, testi alan öğrencilere cinsiyet ve öğrenim gördüğü okul türüne (özel-devlet) göre yanlılık taşıyor mu sorusuna yanıt bulabilmektir. Bu sayede yüzbinlerce öğrencinin geleceği hakkında karar verme niteliği taşıyan bu sınavın geçerlik düzeyi hakkında bilgi sahibi olunacaktır.

Ülkemizde binlerce öğrenci, TEOG gibi başvurduğu geniş ölçekli merkezi sınavlarda aldıkları puanlarla karşılaştırılarak seçilme sürecine dâhil olmaktadır. Bu sınavlara ait maddelerin geçerlik ve güvenilirlik incelemeleri maalesef sınav öncesinde uygulanamamaktadır. Uzmanlar madde yazımı sürecini ancak bilgi ve öngörülerini dâhilinde gerçekleştirebilmektedir. Sınavlara ait bu özelliklerin araştırma süreci yalnızca sınav yapıldıktan sonra gerçekleştirilebilmekte ve eğer sınav, herhangi bir şekilde adaletsizlik yaratmışsa bu durum telafi edilememektedir. Ayrıca ülkemizde bu sınavlara ait maddelerin kamuoyu ile paylaşılması da geçerliği ve güvenilirliği yüksek maddelerin belirlenerek kaliteli bir madde havuzu oluşturulmasını da engelleyerek süreci daha da zor hale getirmektedir. Aslında olması gereken sürekli yeni madde üretmektense maddelerin istatistiksel özelliklerinin incelenerek o madde için yapılan düzeltmeler sayesinde en kalitelisine ulaşmaya fırsat verilmesi

olmalıdır. Alan yazında bulunan bu ve benzeri çalışmalar madde yazımı sürecinde uzmanlara fikir sunabilmeyi amaçlamaktadır. İşte bu noktada geniş ölçekli sınavların maddeleri ile yapılan araştırma çalışmaları uzmanlara kolaylık sağlamaları açısından önemli görülmektedir.

DMF ve yanlılık alanında yapılan çalışmalar, bireylerin bazı özelliklerinden ötürü maddelerin doğru cevaplanma durumlarının değiştiğini ortaya koymaktadır. Bu nedenledir ki, sınava giren bireyler gerek yaşadıkları coğrafya etkisiyle gerek sosyo-ekonomik düzey gerekse cinsiyet farklılığı gibi birçok nedenden ötürü sınavda gerçek yeteneklerini yeterince sergileyebilmekte mi? sorusu alan yazında sıkça karşımıza çıkarmaktadır. Bu nedenden ötürü DMF kavramının testin geçerliği kapsamında önemi, göz ardı edilmemelidir. Ayrıca DMF'nin altında yatan temel nedenler hala büyük ölçüde bilinmemektedir (Roussos ve Stout, 1996b). Alan yazındaki araştırmalar; cinsiyet, dil, kültür ve sosyo-ekonomik düzey gibi farklılıkların ön plana çıktığını göstermektedir. Bu nedenle bu çalışma, bireylerin eğitim gördüğü okulların farklı olmasının maddelerin DMF belirtip belirtmemesi konusunda etkisini incelemesi açısından da önemlidir.

Bu çalışmanın İngilizce testi verilerine dayanması, özel okulların dil eğitiminde sağladığı ayrıcalığın ulusal sınavlarda adaletsizlik yaratıp yaratmama konusunu incelemesi nedeniyle önemli görülmektedir. Ayrıca, ilgili alan yazın incelendiğinde Fen Bilimleri, Matematik ve Türkçe soruları için okul türü veya cinsiyet değişkenine göre yanlılık ya da DMF çalışmaları (Gök, Kelecioğlu ve Doğan, 2010; Karakaya, 2012; Karakaya ve Kutlu, 2012; Kelecioğlu, Karabay ve Karabay, 2014; Yurdugül ve Aşkar, 2004a) yapılmasına rağmen, dil eğitimi ile ilgili olarak yanlılık çalışmasının yapılmadığı gözlemlenmiştir. Özel okullarda verilen yabancı dil eğitimin devlet okullarındaki eğitim ile aynı seviyede olmamasına rağmen öğrencilerin süreç sonunda aynı sınavla değerlendirilmesi durumunda öğrenciler arasında adaletsizlik oluşuyor mu sorusunun yanıtlanması açısından bu çalışma önem arz etmektedir. Ayrıca maddelerin cinsiyete göre de DMF belirtmesi veya yanlılık taşıyıp taşımadığının incelenmesi bu çalışma için önemli görülmektedir.

Çalışmada, test maddelerinin testi alan bireylere yanlılık gösterip göstermediğini tespit etmek amacıyla ilk aşamada Değişen Madde Fonksiyonu (DMF) analizlerinin yapılması amaçlanmaktadır. Alan yazın incelendiğinde DMF analizlerinde farklı farklı birden fazla yöntemin kullanıldığı ve bu yöntemlerin aynı

test içerisinde farklı maddeleri belirlediği görüldüğünden bu çalışmada da üç farklı DMF belirleme yöntemi kullanılacaktır (Acar, 2008; Ateşok ve Deveci, 2008; Benito ve Ara, 2000; Camilli ve Shepard, 1994; Doğan ve Öğretmen, 2006; Kristjansson, Aylesworth, McDowell ve Zumbo, 2005; Ulutaş, 2012). Veri özelliklerine göre uygun görülerek kullanılacak olan yöntemler; Mantel-Haenzsel, Lojistik Regresyon ve SIBTEST yöntemleridir. Ayrıca bu çalışmada, bu üç yöntemin sonuçlarının karşılaştırılması sonrası gücü ve uyumu hakkında alan yazına katkı sağlamak da amaçlanmaktadır. Sonraki adımda ise, istatistiksel olarak DMF içerdiği belirlenen maddelerin, uzman görüşleri yardımıyla yanlılığı ve olası sebeplerinin belirlenmesi hedeflenmektedir.

Araştırma Problemi

TEOG 2016 Nisan (II. Dönem) sınavının İngilizce alt testine ait maddeleri cinsiyet ve okul türüne göre (devlet-özel) herhangi bir gruba karşı yanlılık göstermekte midir?

Alt problemler.

1. TEOG 2016 Nisan (II. Dönem) sınavının İngilizce alt testine ait maddeleri cinsiyete göre yanlılık göstermekte midir?
 - 1.1. İngilizce alt testine ait maddelerin cinsiyet değişkeni için MH yöntemiyle elde edilen sonuçlara göre DMF düzeyleri nedir?
 - 1.2. İngilizce alt testine ait maddelerin cinsiyet değişkeni için LR yöntemiyle elde edilen sonuçlara göre DMF düzeyleri nedir?
 - 1.3. İngilizce alt testine ait maddelerin cinsiyet değişkeni için SIBTEST yöntemiyle elde edilen sonuçlara göre DMF düzeyleri nedir?
 - 1.4. Cinsiyet değişkenine göre DMF içerdiği belirlenen maddeler, uzman görüşüne başvurulduğunda yanlılık göstermekte midir, varsa olası yanlılık sebepleri nelerdir?
2. TEOG 2016 Nisan (II. Dönem) sınavının İngilizce alt testine ait maddeleri okul türüne göre yanlılık göstermekte midir?

- 2.1. İngilizce alt testine ait maddelerin okul türü değişkeni için MH yöntemiyle elde edilen sonuçlara göre DMF düzeyleri nedir?
 - 2.2. İngilizce alt testine ait maddelerin okul türü değişkeni için LR yöntemiyle elde edilen sonuçlara göre DMF düzeyleri nedir?
 - 2.3. İngilizce alt testine ait maddelerin okul türü değişkeni için SIBTEST yöntemiyle elde edilen sonuçlara göre DMF düzeyleri nedir?
 - 2.4. Okul türü değişkenine göre DMF içerdiği belirlenen maddeler, uzman görüşüne başvurulduğunda yanlılık göstermekte midir, varsa olası yanlılık sebepleri nelerdir?
3. İngilizce alt testi maddelerinin MH, LR ve SIBTEST analiz yöntemleriyle elde edilen sonuçları uyumlu mudur?
 - 3.1. İngilizce alt testi maddelerinin cinsiyete göre MH, LR ve SIBTEST analiz yöntemleriyle elde edilen sonuçları uyumlu mudur?
 - 3.2. İngilizce alt testi maddelerinin okul türü değişkenine göre MH, LR ve SIBTEST analiz yöntemleriyle elde edilen sonuçları uyumlu mudur?

Sayıtlılar

Yanlılık belirlemek için görüşleri alınan uzmanların yanlılık kavramına hakim olarak maddelere cevap verdiği varsayılmıştır.

Sınırlılıklar

Bu araştırma, 2015 -2016 Eğitim - Öğretim yılı Nisan ayında (II. Dönem) Türkiye genelinde 1.133.857 öğrencinin girdiği TEOG sınavında İngilizce alt testini alan ve A kitapçığını işaretleyenlerden rastgele seçilen 9.108 öğrenciye ait cevapların yer aldığı örneklem grubu ile sınırlıdır.

Bu çalışma DMF belirlemede kullanılan MH, LR ve SIBTEST yöntemleri ile sınırlıdır.

Tanımlar

LGS: Liselere Giriş Sınavı

SBS: Seviye Belirleme Sınavı

PISA: Programme for International Student Assessment

TEOG: Temel Eğitimden Ortaöğretime Geçiş

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Araştırmanın Kuramsal Temeli

Bu bölümde araştırmanın temelini oluşturan madde yanlılığı kavramının ne olduğu, etkileri, nedenleri ve tespit etmek için yapılması gereken işlemler detaylı bir şekilde açıklanacaktır. Ayrıca bu çalışmanın istatistiksel kısmını oluşturan değişen madde fonksiyonu kavramı da yine aynı titizlikle etkileri, nedenleri ve çeşitli yöntemleri ile ele alınarak avantajları ve dezavantajları kapsamında açıklanacaktır.

Madde yanlılığı ve değişen madde fonksiyonu. Bireylerin yetenekleri hakkında karar vermek amacıyla yapılan ölçme işleminde ölçme aracının bu işlemi hassas bir şekilde yapması beklenmektedir. Bu nedenle tüm koşulların ölçme kuralları çerçevesinde özenle oluşturulması gerekmektedir. Fakat elde edilen ölçme sonuçlarının, ölçme aracının ölçmeyi amaçladığının haricindeki başka özelliklerden etkilenmesinin önüne geçilmesi de maalesef zor bir konudur (Crocker ve Algina, 1986). Bu durumdan ötürü alan yazında, madde yanlılığı kavramına sıkça rastlanmaktadır.

Eğer, eşit yetenek düzeyinde farklı alt gruplardan gelen bireylerin maddeyi doğru cevaplama olasılıkları eşitlik göstermiyorsa o maddenin farklı fonksiyonlaştığından bahsedilebilmektedir (Ackerman ve Evans, 1992; Angoff, 1993; Camilli ve Averill, 1981; Hambleton ve Rogers, 1989; Ironson, 1982; Linn ve Drasgov, 1987; Linn, Levine, Hasting ve Wardrop, 1981; Shepard, Raju, 1988; Zumbo 1999). Bu farklılaşmanın nedeni maddenin belirli bir gruba yanlı davranması olabilir. Bu kapsamda madde yanlılığı kavramı, maddenin aynı yetenek düzeyinde fakat testte ölçülmek istenen özellik haricindeki bazı özelliklere de sahip olan bireylere avantaj sağlarken, diğerlerine dezavantaj sağlaması olarak ifade edilebilir (Osterlind ve Everson, 2009). Bu özellikteki maddelerin, bireylerin karşılaştırılarak bir okula ya da meslek grubuna seçilmesi söz konusu olduğu sınavlarda kullanımı, alınan kararların geçerliğine şüphe düşürmektedir. Bir başka deyişle yanlılık kavramı, bireylerin bulunduğu grubun özelliğinden kaynaklı maddeden elde edeceği puana, sistematik hata karışması olarak da tanımlanabilmektedir (Camilli ve Shepard, 1994; Zumbo, 1999).

Yanlılık çalışmaları, sistematik hataların kaynaklarını belirleyebilmek için yapılan önemli araştırmalardandır. Aslında bir geçerlik araştırmasıdır. Yanlılık çalışmalarında başvurulan yöntemler gereğince yeterince nesnel olmaması ve bu nedenden ötürü sonuçların güvenilirliğine olumsuz yönde etki etmesi, istatistiksel analiz yapma imkânı sağlayan değişen madde fonksiyonu teriminin çalışmalara dâhil edilmesi ile bu çalışmaların daha kolay ve öznellikten uzak bir şekilde yapılmasına imkân sağlamaktadır (Angoff, 1993).

Değişen madde fonksiyonu (DMF); bir gruptaki bireylerin, maddeye doğru cevap verme olasılığının benzer yetenek düzeyinde başka bir gruptaki bireylerden daha fazla olması durumunu tanımlamak için kullanılan istatistiksel bir terimdir (Hambleton ve Rogers, 1989; Zumbo, 2007). Bir başka deyişle; bir ölçme ölçütü ile eşleşen ölçme gruplarından birinin diğer gruba oranla madde performansının daha yüksek olması bu maddenin DMF sergilediği anlamına gelmektedir (Buzick ve Stone, 2011). Burada, o grubun madde performanslarının farklı sonuçlanmasının nedeni ölçme hedefleri ile açıklanamıyorsa bu neden, bireylerin hesaba katılmayan özelliklerinden kaynaklanıyor olabilir (Buzick ve Stone, 2011; Crocker ve Algina, 1986). Bu özellikler; cinsiyet, sosyo-ekonomik farklılıklar, bilgiye ulaşma imkânları (internet, bilgisayar, dersane vb.), bireyin ilgi alanları, yaşanan coğrafya, ırk, dil, yaş vb. ölçme hedeflerinde olmayan bireysel özelliklerdir (Ellis ve Raju, 2003). Bu durum da ölçmeye dâhil olan öğrencilerden DMF yaratabilecek bu olası kaynaklara göre oluşturulan alt grupların karşılaştırılması temelinde yapılan analizlerle araştırılabilmektedir (Ironson and Craig, 1982; Shepard, Camilli and Williams, 1985).

Yanlılık tespiti amacıyla yapılan analizlerde kullanılan teknikler; gerek verilerin parametrik olma veya olmama durumu, gerek kullanılan puanların gözlenen (KTK) veya örtük (MTK) olması durumu gerekse yine puanların iki kategorili veya çok kategorili olma durumu göz önüne alınarak yapılan sınıflandırmalarla alan yazında yerini almıştır. Bu kapsamda alan yazında karşılaşılan farklı birçok sınıflandırma bulunmaktadır. Bu çalışmada Potenza ve Dorans (1995) tarafından yapılan sınıflandırma dikkate alınmıştır. Potenza ve Dorans (1995)'a ait sınıflandırma Tablo 1'de verilmiştir.

Tablo 1

Potenza ve Dorans (1995)'a Ait İkili ve Çok Kategorili DMF Yöntemlerinin Parametrik Olup Olmama Durumlarına Göre Çapraz Sınıflandırması

	Eşleşme Değişkeni	Parametrik Olan	Parametrik Olmayan
İki Kategorili	Gözlenen Puan (KTK)	Log-Doğrusal Modeller Lojistik Regresyon	Mantel Haenszel Standartlaştırma Ki-Kare
	Örtük Değişken (MTK)	Genelleştirilmiş Lojistik Regresyon Lord'un Ki-Karesi	SIBTEST
Çok Kategorili	Gözlenen Puan (KTK)	Çok Kategorili Lojistik Regresyon	Çok Kategorili Mantel - Haenszel Çok Kategorili Standartlaştırma
	Örtük Değişken (MTK)	Genelleştirilmiş Lojistik Regresyon	Çok Kategorili SIBTEST

Bu çalışmada kullanılan verilerin TEOG puanlarına dayalı olması yani çoktan seçmeli testten elde edilmesi, kullanılacak yöntemlerin iki kategorili olmasını gerekli kılmıştır. Alan yazında yapılan incelemeler doğrultusunda gözlenen puanlara (KTK) dayalı yöntemlerin kullanım kolaylığı ve etkili sonuçlar vermesi nedeniyle sıkça kullanılması, bu çalışmada da bu yöntemlerden Mantel – Haenszel ve Lojistik Regresyon yöntemlerinin tercih edilmesinde önemli rol oynamıştır. Bu amaçla KTK daha yakından incelenmiştir.

Klasik Test Kuramı, Psikometri tarihinin temelini oluşturan ve başlangıcından günümüze kadar da kullanılan bir kuram olup, gerçek puanı hesaplamak için gözlenen puan ve hata puanına başvurmaktadır (Crocker ve Algina, 1986). KTK formülü aşağıdaki gibidir (Lord ve Novick, 1968).

$$T = X + E$$

T: Gerçek puan

X: Gözlenen puan

E: Hata puanı

Bu formüle göre, gözlenen puanı elde etmek amacıyla yapılan ölçme işlemine ne kadar hata puanı karışırsa, gözlenen puan gerçek puandan o derece uzaklaşır. Hata puanının sıfır "0" olduğu durumda gözlenen puan gerçek puanı yansıtır. Fakat bu durum uygulamada imkânsız olarak kabul edilir. Bu nedenle yapılan ölçmelerde hata puanını mümkün olduğunca sıfıra yaklaştırmak amaçlanmaktadır.

DMF analizleri, KTK'ye dayalı yöntemler kullanılarak yapılacaksa oluşturulan alt grupları için maddelerin güçlük ve ayırt edicilik özelliklerinin karşılaştırılmasına başvurulabilir (Gomez-Benito ve Navas-Ara, 2000). Bu analiz yöntemleri sadece maddenin özelliğini içermez, aynı zamanda bireyin yeteneğini de barındırır. Bu durum, analiz sonuçlarının grubun özelliğine göre değişkenlik göstereceğine kanıt olarak sunulur ve başka gruplar için genelleme yapılamaz. Her ne kadar bu özellik, KTK'ye dayalı DMF analizlerinin genellenebilirliğini olumsuz olarak etkilese de sağladığı başka avantajlar nedeniyle günümüzde alan yazında kullanımı hala devam etmektedir (Demir ve Köse, 2014; Doğan ve Öğretmen, 2008; Erdem-Keklik, 2014; Kan, Sünbül ve Ömür, 2013; Ross ve Okabe, 2006; Yurdugül ve Aşkar, 2004a).

Mantel-Haenszel, Standartlaşma, Lojistik Regresyon, Olabilirlik Oran Testi ve Ki-Kare yöntemleri KTK' ye dayalı yöntemler olarak alan yazında yer almaktadır.

KTK'ye dayalı yöntemler; analiz sonuçlarının genellenememe özelliği ve oluşturulan alt gruplardaki bireylere bağımlı olma özelliği gibi nedenlerle eleştirilmektedir (Rodney ve Drasgow, 1990; Shepard, Camilli ve Williams, 1985). Bu nedenle yeni yöntemler arayışı içinde örtük değişken dikkate alınarak hazırlanan Madde Tepki Kuramına dayalı yeni teknikler geliştirilmiştir.

Madde Tepki Kuramı, Klasik Test Kuramına alternatif olarak üretilmiş olup bireyin kendi yetenekleri doğrultusunda maddeye vereceği tepkileri olasılık hesaplamaları yaparak ortaya koymayı hedefleyen bir kuramdır (Crocker ve Algina, 1986; Hambleton ve Swaminathan, 1985). MTK'nin bu hedefi doğrultusunda, bireyin gözlenebilen performansı ile ölçülmek istenen yeteneği arasındaki ilişkiden faydalanılır. Son zamanlarda alan yazında sıkça karşılaşılan bu model, KTK'den daha üstün olmasına karşın bazı sayılıtların karşılanması koşulu nedeniyle uygulamada zorlaşmaktadır. Bu sayılıtlar; tek boyutluluk, yerel bağımsızlık ve değişmezlik olarak ifade edilir.

Tek boyutluluk. Bireyin madde performansının ölçülmesi sırasında ölçmenin, sadece tek bir yetenek düzeyinin etkisi altında olmasının beklentisidir. Psikolojik testlerde bu beklentinin karşılanmasının neredeyse imkânsız olması bu sayılıtnın karşılanmasını olanaksız kılmaktadır (Hambleton ve Swaminathan, 1985). Bu durum MTK'nin dezavantajlarındanıdır. Alan yazında, bu durumu ekarte etmek amaçlı geliştirilmiş çok boyutlu modeller de yer almaktadır. Tek boyutluluk sayılıtı faktör analizi yapılarak incelenebilir (Crocker ve Algina, 1986; Sünbül ve Erkuş, 2013).

Yerel bağımsızlık. Yerel bağımsızlık, testte yer alan bir maddeye verilen cevabın başka bir maddeye verilen cevapla ilişkisinin olmaması durumudur (Hambleton ve Swaminathan, 1985). Bu durumu test etmek amacıyla testte yer alan maddelerin ayrı ayrı ikili korelasyonlarına bakılabilir (Crocker ve Algina, 1986; Gelbal, 1994). Testte ölçülmek istenen yetenek düzeyi eşit tutulmak koşuluyla maddelerin her birine verilen cevapların istatistiksel olarak birbirinden bağımsız olduğunun tespiti yerel bağımsızlık sayılıtısının karşılanmasını sağlar. Böylece testin tek boyutluluk sayılıtısını karşılaması durumu, yerel bağımsızlık sayılıtısının da karşılandığını dolaylı olarak göstermektedir (Doğan ve Tezbaşaran, 2003). Bu mantıktan hareketle tek boyutluluk sayılıtısının test edilmesi için yapılan faktör analizi, yerel bağımsızlık sayılıtı hakkında da bilgi verebilmektedir (Hambleton ve Swaminathan, 1985).

Değişmezlik. KTK'ye göre yapılan analizler ile bireyin yeteneği belirlenirken kullanılan ölçme aracı eğer madde güçlük değeri yüksek olan maddeler içeriyorsa bireyin aldığı puan yüksek, madde güçlük değeri düşük olan maddeler içeriyorsa da bireyin aldığı puan düşük olarak belirleniyor. Yani bireyin yeteneği, ölçme aracının betimsel özelliğine göre değerlendirilmektedir. Hâlbuki olması gereken bireyin yetenek düzeyinin ölçme aracına göre değişmemesidir (Hambleton ve Jones, 1993). KTK'nin bu kıstasta eleştirilmesi, MTK kuramının değişmezlik sayılıtısını ön planı çıkarmıştır. Çünkü MTK'nin değişmezlik sayılıtısına göre madde parametreleri, ölçmenin uygulandığı gruba göre değişiklik göstermemelidir (Doğan ve Kılıç, 2017). Bu özellik de parametreleri belirlenen maddelerin, başka yetenek grupları için birçok kez kullanılmasına imkân sağlamaktadır (Çelen, 2008; Kelecioğlu, 2001). Madde parametrelerinin değişmezlik sayılıtısını test etmek için, ölçme aracının uygulandığı

alt grupların madde parametrelerinin karşılaştırması yapılmaktadır (Hambleton, Swaminathan ve Rogers, 1991).

SIBTEST, Genelleştirilmiş LR ve Lord'un Ki-kare Testi, gibi yöntemler MTK'ye dayalı yöntemler olarak alan yazında yer almaktadır. Bu çalışma kapsamında da SIBTEST yönteminin kullanılması tercih edilmiştir.

Değişen madde fonksiyonunu belirlemede kullanılan yöntemler. Bu aşamada sadece bu çalışmada kullanılması tercih edilen yöntemler detaylı olarak incelenmiştir. Bu yöntemler; KTK'ye dayalı yöntemlerden parametrik olmayan Mantel – Haenszel yöntemi, yine KTK'ye dayalı yöntemlerden parametrik olan Lojistik Regresyon ve MTK'ye dayalı parametrik olmayan SIBTEST yöntemidir.

Mantel – Haenszel yöntemi. Sağlık bilimi ile uğraşan Mantel ve Haenszel (1959) çalışmalarında, tabakalara ayrılmış grupları daha kolay karşılaştırmak amacıyla, temeli ki-kare tekniğine dayanan bir yöntem geliştirmişlerdir. Holland ve Thayer (1986) ise bu yöntem üzerinde çalışmalar yaparak yöntemin, aynı yetenek düzeyindeki farklı gruplara göre maddenin farklılaşmasının tespitinde kullanılmasını önermiştir. Bu kapsamda Mantel-Haenszel (MH) adıyla ifade edilen yöntem, ilk DMF belirleme tekniği olarak kullanılmıştır (Holland ve Wainer, 1993).

MH yöntemi, aynı yetenek düzeyinde farklı gruplara ait bireylerin, grup üyeliği ile test puanı arasındaki ilişkinin anlamlılığını test etmektedir (Ackerman ve Evans, 1992; Millsap ve Everson, 1993). Bu amaçla, test puanları dikkate alınarak oluşturulan odak grubun performansı ile referans grubun performansını karşılaştırılmaktadır. Karşılaştırmalar, Tablo 2'de gösterildiği gibi her bir madde için 2x2'lik matrisler oluşturularak yapılmaktadır (Holland ve Thayer, 1986).

Tablo 2

Eşleştirilen Odak ve Referans Gruba Ait Veri Düzeni

		Madde Puanı		
		Doğru (1)	Yanlış (0)	Toplam
Grup	Referans	A_i	B_i	N_{Ri}
	Odak	C_i	D_i	N_{Oi}
	Toplam	N_{Di}	N_{Yi}	T_i

Maddeye verilen doğru ve yanlış cevapların toplamı referans grubu için N_{Ri} değeri ile ifade edilirken, odak grubu için ise N_{Oi} değeri ile ifade edilmektedir. Tablo

2’de verilen N_{Di} ve N_{Yi} değerleri ise maddenin doğru ve yanlış cevaplanma toplamını ifade etmektedir. İstatistiksel hipotezleri tam olarak belirtebilmek için Tablo 2’den yola çıkarak maddelerin gruplara göre doğru cevaplanma ve yanlış cevaplanma olasılıkları hesaplanmaktadır.

Tablo 3

Eşleştirilen Odak ve Referans Gruba Ait Olasılık Tablosu

		Madde Puanı		Toplam
		Doğru (1) cevaplanma oranı	Yanlış (0) cevaplanma oranı	
Grup	Referans	p_{Ri}	q_{Ri}	1
	Odak	p_{Oi}	q_{Oi}	1

Tablo 3’te yer alan p_{Ri} ifadesi Tablo 2’de verilen A_i ’nin N_{Ri} ’ ye oranı alınarak hesaplanmaktadır. Aynı şekilde p_{Oi} değeri de hesaplanırken yine Tablo 2’de verilen C_i ’nin N_{Oi} ’ye oranı alınmaktadır. Bu işlemler sayesinde odak ve referans grubunun maddeyi doğru cevaplama olasılıkları hesaplanmış olmaktadır. Eğer eşleştirilmiş olan bu iki grubun maddeyi doğru cevaplama olasılıkları eşit ($p_{Ri} = p_{Oi}$) ise maddenin DMF özelliği taşımadığı ifade edilmektedir (Holland ve Thayer, 1986; Zwick, Donohue ve Grima, 1993). Eğer eşitlik yoksa yani madde DMF özelliği taşıyorsa MH yöntemi, odd-oranı parametresini kullanarak etki büyüklüğünün yorumlanması için de imkân sağlamaktadır (Gierls, Rogers ve Klinger, 1999). i yetenek düzeyinde maddeye doğru ve yanlış cevap veren odak ve referans grubunda yer alan bireylerin sayısı yardımıyla Odd-oranın kestiricisi olarak α değeri hesaplanmaktadır (Phillips ve Holland, 1987). Tablo 2’de yer alan değerler kullanılarak elde edilen α değerinin hesaplama formülü aşağıdaki gibidir.

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

α olasılık oranıdır; yani test puanına göre eşleştirilmiş referans grubu ile odak grubuna ait bireylerin maddeyi doğru cevaplama olasılığının oranıdır. α değeri, referans grubu için 1 ila ∞ arasında değer alırken odak grubu için ise 0 ila 1 arasında değer alır (Clauser ve Mazor, 1998). Bu durumun bulguyu yorumlamayı zorlaştırması sebebiyle Holland ve Thayer (1988) lojistik bir dönüşüm önerisinde bulunmuşlardır. Dönüşüm formülü aşağıdaki gibidir.

$$\Delta\alpha_{MH} = -2,35 \ln \alpha_{MH}$$

α_{MH} katsayısının logaritmik bir dönüşüm yapılarak elde edilen $\Delta\alpha_{MH}$ değeri, hem maddenin hangi gruba avantaj sağladığını hem de etki büyüklüğünü yorumlamasını kolaylaştırmaktadır (Dorans ve Holland, 1992; Zwick, Donoghue ve Grima, 1993). $\Delta\alpha_{MH}$ değerinin pozitif değeri maddenin, odak gruba avantaj sağladığını ifade ederken negatif değeri ise referans gruba avantaj sağlandığını göstermektedir. Eğer hesaplanan bu değer yaklaşık olarak sıfıra eşitliği tespit edilirse de incelenen bu maddenin gruplara göre farklılaşmadığı ifade edilebilmektedir (Holland ve Thayer, 1986).

$\Delta\alpha_{MH}$ değeri, eğer maddenin farklılaştığını ifade ediyorsa testin geçerliğinin daha iyi ifade edilebilmesi için DMF'nin etki büyüklüğünün hesaplanmasına ve yorumlanmasına ihtiyaç duyulmaktadır. Bu aşamada etki büyüklüğünün yorumlanması amacıyla Zwick ve Ercikan (1989) ve Zieky (1993) tarafından Educational Testing Service (ETS) için üç ögeli bir sınıflama oluşturulmuştur. Bu sınıflamaya ait ölçütler Tablo 4'te verilmiştir.

Tablo 4

$\Delta\alpha_{MH}$ Değerlerini Yorumlama Ölçütleri

Düzyey	Değer Aralıkları	DMF Miktarı
A	$ \Delta\alpha_{MH} < 1$	Yok veya ihmal edilebilir
B	$1 \leq \Delta\alpha_{MH} < 1,5$	Orta düzey
C	$ \Delta\alpha_{MH} \geq 1,5$	Yüksek düzey

Tablo 4'te görüldüğü üzere maddenin DMF miktarı, düzeylerle (A, B ve C) ifade edilmiştir. A düzeyindeki maddeler ihmal edilebilirken B ve C düzeyindeki maddeler, testin geçerliğine olumsuz yönde etki edeceği için bu maddeler ya düzeltilmeli ya da testten çıkartılmalıdır (Kan, 2007b).

MH yöntemi, örneklem sayısının az olduğu çalışmalarda istatistiksel analiz yapma imkânı sağlaması açısından avantajlıdır. Ayrıca maddeler için tek tek elle hesap yapmak yerine bazı yazılım veya programların kullanılmasına da olanak sağlamaktadır (Clauser ve Mazor, 1998). Bu programlara SAS, Stata, Systat, Jmetrik, DIFAS ve EZDIF örnek olarak verilebilmektedir (Ayala, 2009; Koyuncu, Aksu ve Kelecioğlu, 2018). MH yönteminin, $\Delta\alpha_{MH}$ değerinin eldesinin sağlanması

nedeniyle DMF'nin yorumlanması açısından kolaylık sağlıyor olması da yine bu yöntemin avantajlarındandır (Clauser ve Mazor, 1998).

MTK'ye dayalı istatistiksel analizler, gruptan bağımsız olarak madde parametreleri ile analiz yapılabilme özeliğinden ötürü daha güvenilir sonuçlar ortaya koymaktadır. MH tekniği de MTK'ye ait bir yöntem olmamasına rağmen bu yöntemlerle paralel bulgulara ulaşmaktadır. İşte bu nedenlerden ötürü MH tekniğine alan yazında çok fazla rastlanmaktadır (Hambleton ve Rogers, 1989; Henderson, 2001; Millsap ve Everson, 1993; Yıldırım ve Berberoğlu, 2009). Eğitim alanında referans sayılan Educational Testing Service (ETS) tarafından da çok tercih edilen bir yöntem olmuştur. Fakat tek biçimli olmayan DMF'yi belirleme konusunda yetersiz olması MH yönteminin en büyük sınırlılığıdır (Bertrand ve Boiteau, 2003; Clauser ve Mazor, 1998; Swaminathan ve Rogers 1990). Ayrıca test puanlarına dayalı analiz gerçekleştirmesi sebebi ile 20 maddeden az testlerde I. tip hatanın görülmesi olasıdır (Uttaro ve Millsap, 1994; Zwick, 1990).

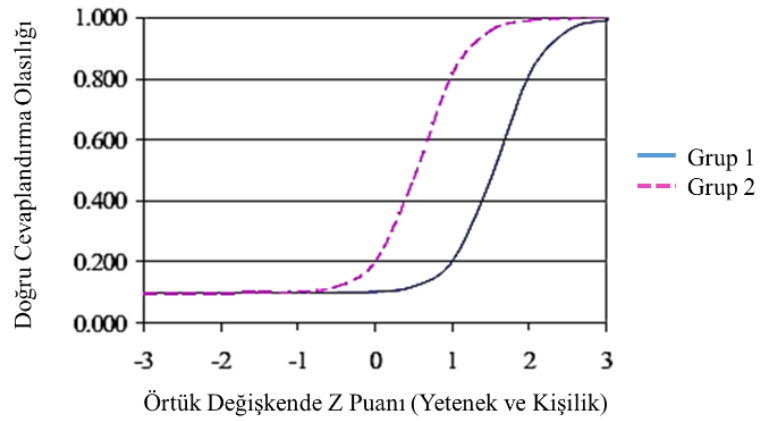
Alan yazında MTK'ye dayalı birçok DMF belirleme yöntemi önerilmesine rağmen araştırmalarda hala en çok tercih edilen yöntem, KTK'ye ait yöntemlerden MH yöntemidir. Buna rağmen, Mellenberg (1982)'in tek biçimli ve tek biçimli olmayan DMF'yi tanımlamasından sonra MH yönteminin tek biçimli olmayan DMF'yi belirleme konusunda yetersiz olduğu görülmüştür (Bertrand ve Boiteau, 2003; Clauser ve Mazor, 1998; Swaminathan ve Rogers 1990).

Tek biçimli (TB) DMF, yetenek seviyesi ile grup üyeliği arasında bir etkileşim olmadığı zaman mevcuttur. Yani, maddenin doğru cevaplanma olasılığının, bir grup için tüm yetenek seviyelerinde diğer gruba göre daha büyük olması, hep aynı gruba avantaj sağlanmasıdır (Finch ve French, 2007; Swaminathan ve Rogers, 1990). Tek biçimli olmayan (TBO) DMF ise yetenek seviyesi ile grup üyeliği arasında bir etkileşim olduğunda ortaya çıkar. Yani, maddeyi doğru cevaplama olasılığındaki farkın tüm yetenek düzeylerinde iki grup için aynı olmaması, avantajın bir gruptan diğer gruba göre değişkenlik göstermesidir (Finch ve French, 2007; Swaminathan ve Rogers, 1990;). TB ve TBO DMF, MTK'de yer alan madde karakteristik eğrileri ile gösterilmektedir (Hambleton ve Swaminathan, 1985).

Madde karakteristik eğrisi (MKE), bireyin madde için sergilediği performansa göre başarılı olma olasılığının ve maddeyi doğru cevaplama yeteneğinin birlikte

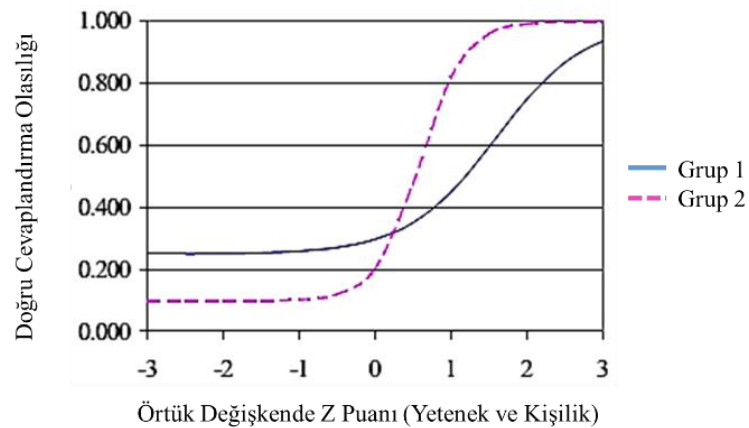
verildiği bir grafik çizimidir (Crocker ve Algina, 1986; Hambleton ve Swaminathan, 1985). Belirlenen alt grupların aynı madde için gösterdiği performans, madde karakteristik eğrisinde farklılaşma gösteriyorsa o madde için DMF gösteriyor yorumu yapılabilmektedir (Camili ve Shepard, 1994; Zumbo, 1999).

Madde, TB DMF özelliği gösteriyorsa bu maddeye ait MKE Şekil 1.'de verilen grafikte yer aldığı gibi tüm yetenek düzeylerinde odak grubun (grup 1) maddeyi doğru cevaplama olasılığı referans gruptan (grup 2) tutarlı olarak düşük olmaktadır (Hambleton ve diğerleri, 1993).



Şekil 1. Tek Biçimli DMF (TB DMF) grafiği.

Eğer madde, TBO DMF özelliği gösteriyorsa da grupların maddeyi doğru yanıtlama olasılığındaki farklılık, tüm yetenek düzeylerinde Şekil 2.'de gösterildiği gibi tutarsızlık belirtir (Hambleton ve diğerleri, 1993).



Şekil 2. Tek Biçimli Olmayan DMF (TBO DMF) grafiği.

TBO DMF'yi belirlemede yetersiz olan MH yöntemine alternatif olarak Swamithan ve Rogers (1990) tarafından geliştirilen yöntem, Lojistik Regresyon yöntemidir.

Lojistik Regresyon yöntemi. Lojistik Regresyon yöntemi, benzer yetenek düzeyindeki bireylerin maddeyi doğru yanıtlama olasılıklarını, test puanını ölçüt olarak belirlemeyi amaçlayan istatistiksel bir yöntemdir. Swaminathan ve Rogers (1990)'ın bir maddenin doğru cevaplanma olasılığını tahmin etmek için kullandığı lojistik regresyon modeli formül 1'deki gibidir.

$$P(u = 1|\theta) = \frac{e^{(\beta_0 + \beta_1 \theta)}}{[1 + e^{(\beta_0 + \beta_1 \theta)}]} \quad \text{formül 1}$$

Modelde yer alan parametreler; u = maddeye verilen cevap, θ = Bireyin gözlenen yeteneği, β_0 =kesişim parametresi ve β_1 = eğim parametresi olarak ifade edilmektedir. Lojistik regresyon modeli, DMF analizi amacıyla belirlenen iki alt grup için ayrı ayrı hesaplanmaktadır.

$$P(u_{ij} = 1|\theta_{ij}) = \frac{e^{(\beta_{0i} + \beta_{1j} \theta_{1j})}}{[1 + e^{(\beta_{0i} + \beta_{1j} \theta_{1j})}]} , i = 1, \dots, n_j, j = 1, 2.$$

Burada $u_{ij} = j$ grubundaki kişinin i maddesine verdiği cevap, $\beta_{0j} = j$ grubundaki kesişme parametresi, $\beta_{1j} = j$ grubu için eğim parametresi ve $\theta_{ij} = j$ grubundaki bireyin i maddesi için gözlenen yeteneği olarak ifade edilmiştir.

Aynı yetenek düzeyine sahip alt gruplardaki bireylerin aynı maddede gösterdiği başarı durumunun aynı olmaması o madde için DMF varlığını işaret etmektedir (Hambleton & Swaminathan, 1985). Ya da eşleştirilmiş iki alt grup için lojistik regresyon eğrileri aynı ise o maddede DMF özelliği gözlemlenmemektedir (Swaminathan ve Rogers, 1990). Yani bu durum alt gruplara ait kesişim parametrelerinin ve eğim parametrelerinin eşitliğinde ($\beta_{01} = \beta_{02}$ ve $\beta_{11} = \beta_{12}$) gözlenmektedir.

Logistik regresyon eğrileri eğer paralellik gösteriyorsa yani çakışma yoksa Tek Biçimli DMF'ye rastlanmaktadır. Bir başka deyişle, alt gruplara ait eğim parametrelerinin eşitliği fakat kesişim parametrelerinin eşitsizliği ($\beta_{11} = \beta_{12}$ ve $\beta_{01} \neq \beta_{02}$) durumudur. Eğer, logistik regresyon eğrileri paralellik göstermiyorsa tek biçimli olmayan DMF'na rastlanmaktadır. Bu durumda da alt gruplara ait kesişim

parametrelerin eşitliği fakat eğim parametrelerinin eşitsizliği ($\beta_{01} = \beta_{02}$ ve $\beta_{11} \neq \beta_{12}$) gözlenmektedir.

Yukarıda verilen formül 1'e eşit ama farklı ifade edilen alternatif bir formül ise şu şekilde olur:

$$P(u = 1) = \frac{e^z}{1+e^z} \quad \text{formül 2}$$

Bu formülde u maddeye verilen cevabı ifade ederken z parametresi de şuna eşittir:

$$z = \tau_0 + \tau_1\theta + \tau_2g + \tau_3(\theta g)$$

z parametresinde θ bireyin gözlenen yeteneğini, g değişkeni ise gruba aitliği tanımlar. Eğer g kişisi 1 nolu gruba aitse formülde g yerine 1 (θ_1), 2 numaralı gruba aitse de 0 (θ_0) yazılmaktadır. Burada (θg) olarak gösterilen değişken, bireyin ait olduğu grup ve gözlenen yeteneği olan iki bağımsız değişkenin çarpımı yani bileşkesidir. z parametresinde τ_2 o madde için gruplar arasındaki performans farkını ($\tau_2 = \beta_{01} - \beta_{02}$) ortaya koyarken τ_3 ise grup ile yetenek arasındaki etkileşimi ($\tau_3 = \beta_{11} - \beta_{12}$) göstermektedir. Bu durumda $\tau_2 \neq 0$ ve $\tau_3 = 0$ ise madde TB DMF, $\tau_3 \neq 0$ ve $\tau_2 = 0$ veya $\tau_2 \neq 0$ ise madde TBO DMF göstermektedir.

Doğrusal regresyon eşitliği (z), maddeyi doğru cevaplama olasılığının doğru cevaplayamama olasılığına bölünmesinin doğal logaritmasıdır. Yani z değeri odd oranının doğal logaritmasını belirtmektedir. Bu eşitlik şöyledir:

$$z = \ln \left[\frac{pi}{(1-pi)} \right] = \tau_0 + \tau_1\theta + \tau_2g + \tau_3(\theta g)$$

TB ve TBO DMF serbestlik derecesi 2 olan χ^2 testi ile test edilebilmektedir (Swaminathan ve Rogers, 1990 ve Zumbo, 1999).

Lojistik regresyon, bağımlı değişkenin kategorik (sürekli olmayan) olduğu koşullarda kullanılan bir yöntemdir (Pedrajita ve Talisayon, 2009; Zumbo, 1999). Bu yöntemde, sürekli bağımsız değişkenler yardımıyla, sürekli olmayan (kategorik veya ikili) bağımlı değişkeni yordamayı amaçlayan bir istatistiksel model kurulmaktadır (Clauser ve Mazor, 1998). İki kategorili (1-0) bağımsız değişken için oluşturulan lojistik regresyon modelinde toplam puan (X), ait olduğu grup (g) ve ait olduğu grup ile toplam puanın (X ile g) etkileşimi incelenmektedir (Gierl, Jodoin ve Ackerman, 2000). Bu yöntemle bireyin maddeyi doğru veya yanlış cevaplamış

olmasının asıl sebebinin bu değişkenlerden hangisinden kaynaklandığı belirlenmektedir (Penfield ve Camilli, 2007). Bu amaçla değişkenler modele sırayla dâhil edilmektedir (Gierl, Jodoin ve Ackerman, 2000).

Model 1, toplam puanlar dikkate alınarak oluşturulmaktadır.

$$z = \tau_0 + \tau_1\theta \quad \text{Model 1}$$

Model 1'e grup değişkeni eklenerek Model 2 kurulmaktadır.

$$z = \tau_0 + \tau_1\theta + \tau_2g \quad \text{Model 2}$$

Model 3'te ise önceki değişkenlere etkileşim değişkeni eklenmektedir.

$$z = \tau_0 + \tau_1\theta + \tau_2g + \tau_3(\theta g) \quad \text{Model 3}$$

Bu aşamaların hepsinde serbestlik derecesi 2 olan χ^2 testi uygulanırken Zumbo (1999), Swaminathan ve Rogers'ın (1990), Model 3 ile Model 1'de elde edilen Nagelkerke R^2 (Nagelkerke, 1991) farklarının χ^2 testine karşılık geldiğini ifade etmiştir. TB DMF'nin varlığı Model 2 ve Model 1'den elde edilen R^2 farkı (ΔR^2), TBO DMF'nin varlığı ise Model 3 ve Model 2'den elde edilen R^2 farkının tespiti ile yoklanmaktadır (Zumbo, 1999). Bu analizler SPSS paket programı ile hesaplanabilmektedir. İki model arasındaki R^2 farkı (ΔR^2), DMF miktarının yorumlanmasında kullanılan etki büyüklüğünü belirtmektedir (Swaminathan ve Rogers, 1990). ΔR^2 'lerinin yorumlanmasında alan yazında yer alan sınıflama ölçütleri Tablo 5'te verilmiştir.

Tablo 5

Zumbo ve Thomas (1996), Jodoin ve Gierl (2001) ve Bakan-Kalaycıoğlu ve Kelecioğlu (2011) Tarafından Düzenlenen ΔR^2 Değerlerini Yorumlama Ölçütleri

DMF Düzeyi	Açıklama	Zumbo ve Thomas (1996)'a göre değer aralıkları	Jodoin ve Gierl (2001)'e göre değer aralıkları	Bakan-Kalaycıoğlu ve Kelecioğlu (2011)'e göre değer aralıkları
A	Yok veya ihmal edilebilir	$\Delta R^2 < 0,13$	$\Delta R^2 < 0,035$	$\Delta R^2 \leq 0,010$
B	Orta düzey	$0,13 \leq \Delta R^2 < 0,26$	$0,035 \leq \Delta R^2 < 0,070$	$0,010 \leq \Delta R^2 < 0,020$
C	Yüksek düzey	$0,26 \leq \Delta R^2$	$0,070 \leq \Delta R^2$	$0,020 \leq \Delta R^2$

Jodoin ve Gierl (2001), farklı büyüklüklerdeki örneklerle yaptıkları simülasyon çalışmalarında LR için kullanılan etki büyüklüğü sınıflandırmasının büyük örneklerde I. tip hatayı arttırdığını ifade etmişler ve daha hassas olan kendi sınıflandırma kriterlerini oluşturmuşlardır. Benzer sebepten ötürü Bakan-Kalaycıoğlu ve Kelecioğlu (2011) da kendi sınıflama sistemlerini oluşturmuşlardır.

Lojistik Regresyon yöntemi; TBO DMF'nu tespit edebilmesi, küçük örneklem gruplarıyla çalışılabilmesi ve yazılım programlarının kullanılabilirliği gibi sebeplerden ötürü alan yazında en çok önerilen yöntemlerden biridir (Camilli ve Shepard, 1994; Clauser ve Mazor, 1998; Swaminathan ve Rogers, 1990).

SIBTEST yöntemi. Temelleri Dorans ve Kulick (1986)'in Standartlaştırma yöntemine dayanan SIBTEST (Simultaneous Item Bias Test) yöntemi, Shealy ve Stout (1993) tarafından geliştirilmiştir. SIBTEST yöntemi, bireylerin karşılaştırılmasında örtük puanlar yerine toplam puanlar aracılığıyla gerçek puanların kestirimini kullanan bir DMF analiz yöntemidir (Abbott, 2007). Bu özelliği nedeniyle de I. tip hatayı kontrol edebilmesi diğer yöntemlere göre avantajlı bir yapı oluşturur (Nandakumar, 1993). TB DMF belirlemede örtük puanlar kullanan MH yöntemiyle benzer sonuçlar üretmektedir (Clauser ve Mazor, 1998). Ayrıca parametrik olmayan bir yöntem olan SIBTEST, istatistiksel hipotez ile test edilmektedir. (Gierl, Khaliq ve Boughton, 1999).

$$H_0: B(T) = P_R(T) - P_O(T) = 0$$

$$H_1: B(T) = P_R(T) - P_O(T) \neq 0$$

Belirtilen hipotezlerde $P_R(T)$ T gerçek puanı ile referans gruptaki bireylerin maddeyi doğru cevaplama olasılığı iken $P_O(T)$ ise odak gruptaki bireylerin maddeyi doğru cevaplama olasılığını belirtmektedir. Bu iki ifade arasındaki fark da $B(T)$ ifadesini açıklamaktadır. Eşitlik (H_0) söz konusu olduğunda odak ve referans gruplarda gerçek puanları bakımından eşleştirilen bireylerin, maddeyi doğru cevaplama olasılığı, her iki grupta da aynı olmaktadır. Bu durum maddenin DMF özelliği göstermediğinin ifadesi olarak tanımlanmaktadır. Eşitliğin sıfırdan farklı olması (H_1) ise DMF'nin varlığını ifade etmektedir. DMF, odak ve referans gruplarındaki bireylerin toplamına ait gerçek puan dağılımından elde edilen $B(T)$ 'in beklenen değeri olan β ile tanımlanır (Abbott, 2007). Ayrıca DMF miktarı da etki

büyüklüğü kullanılarak belirlenebilmektedir. DMF etki büyüklüğünü yorumlamak için Roussos ve Stout (1996) tarafından bir sınıflama ölçütü önerilmiştir.

Tablo 6

β Değerlerini Yorumlamak için Roussos ve Stout (1996) Tarafından Önerilen Sınıflama Ölçütleri

Düzye	Değer Aralıkları	DMF Miktarı
A	$\beta < 0,059$	Yok veya ihmal edilebilir
B	$0,059 \leq \beta < 0,088$	Orta düzey
C	$\beta \geq 0,088$	Yüksek düzey

Hipotez testinde β değerinin pozitif olması maddenin referans grubun lehine farklılaştığını ifade ederken negatif olması ise odak grubun lehine olduğunu gösterir (Abbott, 2007).

Olası değişen madde fonksiyonu kaynakları. Yanlılık çalışmalarında DMF kaynaklarının belirlenmesi, test hazırlama sürecinde oluşabilecek adaletsizliğin önüne geçilmesi açısından önemli bir adımdır (Clauser ve Mazor, 1998). DMF belirten kaynakların tespiti, madde yazım sürecini de daha kaliteli hale getirmektedir. Olası kaynaklar hakkında ne kadar çok bilgi sahibi olunursa madde yazımı sürecinde bu kaynaklar göz önünde bulundurularak yazılan maddelerin geçerlik değeri de oldukça yüksek olabilecektir. Bu sebepten ötürü alan yazında bulunan DMF kaynaklarının belirlenmesine yönelik çalışmalar çok kıymetlidir.

Bu çalışma, İngilizce dersi kapsamında hazırlanan maddeleri incelemektedir; fakat alan yazın incelendiğinde İngilizce dersi kapsamında hazırlanan maddeler ile ilgili DMF belirleme çalışmalarına pek rastlanmamaktadır. Bu nedenle diğer dersler kapsamında yapılan çalışmalar incelenerek çalışma için öngörü oluşturulması hedeflenmiştir.

Alan yazındaki DMF veya yanlılık çalışmaları incelendiğinde; cinsiyet, dil, kültür ve sosyo-ekonomik düzey gibi farklılıklar ön plana çıkmaktadır. Aynı yetenek düzeyindeki iki bireyin bu özelliklerinden ötürü maddeyi doğru cevaplama olasılığı değişiklik gösteriyorsa o madde tekrar incelenmeli ve gerekiyorsa düzenlemeler yapılmalıdır. Testin düzeni (madde sıralaması, kitapçık türü vb.) de bireylerin farklı ortalama madde puanlarına sahip olmasına neden olabilir. Bu alanlarda çok fazla

çalışma olmasa da Yakar ve Yavuz (2014)'un SBS'ye ait farklı kitapçıklardan elde ettiği veriler ışığında %35 düzeyinde farklılaşma tespit etmesi, bu özelliklerin de DMF belirtilebileceğini göstermektedir.

Sonuç olarak, yurt içinde ve yurt dışında yapılan çalışmalar ışığında DMF'ye nelerin neden olduğu ve onu en aza indirmek için ne yapılabileceği konusunda daha fazla araştırma yapılmasına ihtiyaç duyulmaktadır.

Çalışmanın bu bölümüne kadar, aynı yetenek düzeyindeki bireylerin maddeyi doğru cevaplama olasılığındaki farklılığın madde yanlılığından kaynaklanabileceği ifade edilmiştir. Fakat bu farklılık bireyin ölçülmek istenen özellik kapsamındaki gerçek bilgi ve becerisinden de kaynaklanabilir. Bu durum da madde etkisi kavramına dikkat çekmektedir.

Madde etkisi. Maddenin geçerliğini etkileyen DMF kavramı incelenirken eşleştirilen alt grupların benzer özellikte olması kritik özelliktir. Aksi takdirde maddenin farklılaşmasının nedeni gruplar arasındaki gerçek farklılıklar olabilir (Clauser ve Mazor, 1998). Bu durum *madde etkisi* (item impact) olarak ifade edilmektedir.

Madde etkisi, maddeyi doğru cevaplama olasılığının incelendiği gruplar arasında oluşan farklılaşmanın, karşılaştırılan grupların yetenek düzeyindeki gerçek farklılığıdır (Ackerman, 1992b; Zumbo, 1999). Bu durum maddenin geçerliğini etkilememektedir. Çünkü farklılaşma maddenin özelliğinden değil, bireyin maddeyi cevaplama aşamasında göstermiş olduğu performansından kaynaklanmaktadır. Ya da bireyin önceden edindiği bilgi ve beceri maddeyi doğru cevaplama olasılığını etkilemiş olabilir. Bu durumda maddenin yanlılık gösterdiğini ifade etmek doğru olmaz.

Madde etkisini veya maddenin yanlılığını ortaya koymak için istatistiksel bir yöntem bulunmamaktadır. Bu aşamada uzman bilgisine başvurulabilir. Uzmanlar; maddeye verilen cevaplardaki farklılaşmanın, maddenin ölçmeyi hedeflediği özellikten mi yoksa bireyin sahip olduğu beceriden mi kaynaklı olduğunu inceleyerek sürece katkı sağlayabilir.

İlgili Araştırmalar

Bu bölümde araştırmaya yön veren yanlılık konulu çalışmalar incelenmiştir. Alan yazında yanlılık kapsamında yapılan birçok çalışma bulunmaktadır. Bu çalışmaların hepsine burada yer vermek yerine araştırmamanın amacı ile benzer olan çalışmalar, araştırma ile ilişkilendirilerek bu bölümde özetlenmiştir.

Türkiye’de yabancı dil alanıyla ilgili yapılan yanlılık çalışmaları. Türkiye’de yabancı dil alanıyla ilgili geniş ölçekli sınavlarla yapılan yalnızca bir tane yanlılık çalışmasına rastlanmıştır.

ÜDS 2005 Mart Sosyal Bilimler ve İngilizce testine ait maddeleri yanlılık açısından inceleyen Ateşok-Deveci ve Çıkrıkçı-Demirtaşlı (2008), analizleri cinsiyet ve alan özelliklerinin (Eğitim Bilimleri (EB)- Beşeri ve İdari Bilimler (BİB)) karşılaştırmasını yaparak gerçekleştirmişlerdir. DMF tespit edilen maddelerden doğum oranları, ekonomi, para ve sanat konulu maddeler kadınlar lehine, askerlik, döviz kuru, tarih ve ticaret konulu maddeler ise erkekler lehine davranmıştır. Alan özelliklerine göre yapılan karşılaştırmaya göre ise; siyaset, ekonomi, bankacılık ve bilgisayar konulu maddeler, Beşeri ve İdari Bilimler alanındaki bireylere; tarih, askeri operasyon ve arkeoloji konulu maddeler ise, Eğitim Bilimleri alanındaki bireylere avantaj sağlamıştır.

Yurtdışında yabancı dil alanıyla ilgili yapılan yanlılık çalışmaları. Alan yazında yer alan yurtdışı çalışmalarından yabancı dil (İngilizce) maddelerinin kullanıldığı bir kaç araştırmaya rastlanmıştır.

Kore Üniversite Giriş Sınavı (2003) İngilizce alt testi ile cinsiyet kapsamında madde yanlılığı analizi yapan Pae ve Park (2006), olasılık oran ölçümü (Likelihood-Ratio) ve Doğrulayıcı Faktör Analizi yöntemlerini kullanmışlardır. Okuduğunu anlama yeteneğini ölçen 33 maddeden 22’sinde DMF tespit edilmiştir. Bu maddelerden 13 tanesi erkekler lehine DMF göstermiş olup yanlılık nedenleri deniz şarkıları, e-ticaret, iş danışmanlığı, Doğu-Batı ilişkisi, Korece gazeteler gibi konular olarak ifade edilmiştir. DMF gösteren diğer 9 madde ise kadınlar lehine çalışmış olup yanlılık nedenleri ise çocukluk deneyimi, seyahat deneyimi, özür dilemenin önemi, insani ilişkiler sanat gibi konularla ilişkilendirilmiştir.

Ross ve Okabe (2006), İngilizceyi yabancı dil olarak öğrenen 5 Japon Enstitüsünden öğrencilerle okuduğunu anlama becerisine dayalı bir dil testinin

cinsiyete göre yanlılık taşıyıp taşımadığına dair bir çalışma yapmışlardır. Çalışmada 4 farklı konudan (kelime bilgisi-görsel-futbol-taşımacılık) oluşan 20 maddelik bir ölçme aracı hazırlanarak kullanılmıştır. Öncelikle ölçme aracı uzmanlara gösterilmiştir. Uzmanlar; deneyimli - deneyimsiz ve cinsiyete (kadın-erkek) göre ayrılmış öğretmen gruplarından oluşturulmuştur. Ölçme aracındaki maddeler öğretmenlere gösterildiğinde deneyimli ve deneyimsiz erkek öğretmenler ölçme aracının genel olarak ön yargılı bir şekilde, konular nedeniyle erkeklerin lehine DMF gösterebileceğini ifade etmiştir. Daha sonra MH, LR ve SIBTEST yöntemleri kullanılarak DMF analizleri yapılmıştır. 20 maddelik ölçme aracından sadece futbolla ilgili olan maddelerden bir tanesinde DMF görülmüştür. Bu sonuçla uzmanlara tekrar başvurulduğunda ise deneyimli ve deneyimsiz erkek öğretmenler edinilen bulgunun görüşlerinden farklı çıkmasının nedenini bu gruptaki kadın cevaplayıcıların genel olarak daha başarılı olması ile ifade etmişlerdir.

DMF yöntemlerinin karşılaştırılmasıyla ilgili araştırmalar. Yanlılık araştırmasının istatistiksel kısmını oluşturan DMF analizleri bölümünde tercih edilen yöntemlerin avantaj ve dezavantajlarının gözlenmesi için alan yazında bulunan çalışmalar incelenmiştir. Yöntemler KTK'ye ve MTK'ye göre değişmektedir. Bu araştırmada kullanılması düşünülen MH, LR ve SIBTEST yöntemleri ön planda tutularak alan yazın incelemesi yapılmıştır. Bunlardan bazılarına bu bölümde yer verilmiştir.

Altıncı sınıfta (2009), yedinci sınıfta (2010) ve sekizinci sınıfta (2011) SBS'ye giren öğrenciler için sınavın alt testlerine (Türkçe, Matematik, Fen ve Teknoloji, Sosyal Bilgiler ve Yabancı dil) ait maddelerin, cinsiyete göre DMF göstermesi Kan, Sünbül ve Ömür (2013) tarafından incelenmiştir. İncelemelerinde amaç, KTK'ye dayalı yöntemlerle MTK'ye dayalı yöntemlerin tutarlılığını karşılaştırmaktır. Bu kapsamda DMF belirleme yöntemlerinden; Dönüştürülmüş madde güçlüğü, MH, LR, Lord'un x^2 ve Raju'nun alan ölçüsü yöntemleri kullanılmıştır. Cinsiyet değişkenine göre DMF analizinde yabancı dil alt testinde KTK'ye ait DMF belirleme yöntemlerinden sadece MH yönteminde bir madde B düzeyinde DMF göstermiş olup diğerlerinde bulunmamıştır. MTK'ye ait yöntemlerde ise Lord'un x^2 tekniğinde 13 madde, Raju'nun alan ölçülerinde ise 11 madde DMF göstermiştir. Çalışma sonucunda KTK'ye ait yöntemlerin kendi aralarında tutarlı ve az sayıda DMF belirlediği, MTK'ye ait yöntemlerin ise yine kendi aralarında tutarlı fakat KTK'ye

oranla daha fazla sayıda DMF belirlediği saptanmıştır. MTK'ye ait yöntemlerin fazla DMF belirlemesinin nedeni, örneklem büyüklüğü ile açıklanmıştır.

DMF belirlemede kullanılan yöntemleri karşılaştırmak amacıyla Atalay, Gök, Kelecioğlu ve Arsan (2012), bir simülasyon çalışması yapmıştır. Bu amaçla gözlenen puan yöntemleri olarak ifade edilen MH ve LR yöntemleri ve örtük puan yöntemleri olarak ifade edilen MTK-OO ve SIBTEST yöntemlerini karşılaştırmışlardır. Test uygulamalarının olası koşulları olan; örneklem büyüklüğü, yetenek dağılımı ve DMF gösteren madde oranı dikkate alınarak oluşturulan alt gruplar ile çalışılmıştır. Tüm koşullar için test uzunluğu 60 madde ile sabit tutulmuştur. Örneklem büyüklüğünün az ya da çok olduğu durumda DMF belirleme yöntemlerinin nasıl işlediğini görmek için 400'er ve 1500'er kişiden oluşan örneklem oluşturulmuştur. Büyük örneklem grubunda yöntemlerin I. tip hataya daha fazla düşerek tüm maddelerin ihmal edilebilir (A) düzeyde DMF gösterdiği, küçük örnekte ise I. tip hataya düşme durumunun daha az ve belirlenen DMF'li maddelerin önemli düzeyde olduğu sonucuna ulaşılmıştır. Bir diğer simülasyon koşulu olan yetenek dağılımı kapsamında, yetenek dağılımı eşit olan ve olmayan alt gruplar oluşturulmuştur. Son olarak da DMF'li madde oranı %5 ve %10 olan alt gruplar oluşturulmuştur. Çalışmanın sonucuna göre örtük puan (MTK-OO ve SIBTEST) yöntemlerinin gözlenen puan (MH ve LR) yöntemlerine oranla daha duyarlı ve etkili bir şekilde DMF'li madde belirlediği tespit edilmiştir. Örtük puan yöntemleri kendi aralarında kıyaslandığında ise MTK-OO yöntemi daha duyarlı sonuçlar vermiştir. Ayrıca gözlenen puan yöntemleri de kendi aralarında kıyaslandığında MH yönteminin LR yöntemine oranla daha duyarlı olduğu sonucuna varılmıştır. Çalışmada, belirlenen yöntemlerin türetilen farklı koşullarda kıyaslanması nedeniyle, hangi yöntemin hangi koşullarda daha duyarlı ve etkili sonuçlar verebileceği hakkında öneriler oluşturulmuştur. Küçük örneklerde MH, LR ve MTK-OO teknikleri; büyük örneklerde ise I. tip hata kontrol edilebildiği için SIBTEST tekniği önerilmiştir.

Uyar ve Kaya-Uyanık (2016), yaptıkları DMF çalışmasında, alan yazında yaygın olarak tercih edilen KTK'ye ait MH ve LR yöntemlerini kullanmışlardır. Bu amaçla çalışmalarında PISA 2012 sınavında yer alan matematik ve fen okuryazarlığı alt testine ait bilişsel becerileri ölçen maddelerin verilerini kullanmışlardır. Odak ve referans grupları belirlerken OECD ortalamasının üstünde

puan alan Hong Kong ve Finlandiya, ortalamanın altında puan alan Türkiye ve Rusya'ya ait veriler kullanılmıştır. Matematik alt testinin analizlerine göre 28 maddeden 24'ü, fen okuryazarlığı alt testinde ise 34 maddeden 30'unda DMF tespit edilmiştir. Bulgulara göre, OECD ortalamasının üstünde puan alan Hong Kong ile ortalamanın altında puan alan Türkiye arasında fen testinde yer alan maddelerin diğer ülkelere oranla daha az farklılaşmıştır. Bu da demek oluyor ki, maddelerin farklılaşma nedeni sadece ülkelerin başarı durumu kaynaklı demek doğru değildir. DMF'li maddelerin belirlenmesinin nedeni yüksek ihtimalle farklı dil ve kültür özellikleri ya da çeviri problemleri olabilir. Çalışmadan çıkarılan bir diğer sonuç ise kullanılan DMF belirleme yöntemlerinin I. tip hata ve istatistiksel gücü hakkında bilgi vermesidir. Çalışmada kullanılan MH ve LR teknikleri yakın sonuçlar elde etmesine karşın MH tekniği daha fazla DMF'li madde tespit etmiştir. Bu durum, MH tekniğinde I.tip hataya düşme olasılığının daha fazla olmasından kaynaklanabileceği şekilde açıklanmıştır. Ayrıca DMF tespit edilen maddelerin etki büyüklükleri incelendiğinde tekniklere göre değişiklik gösterdiği görülmüştür. Bu durum da, etki büyüklüğü belirlemede kullanılan sınıflandırma farkı ile açıklanmıştır.

Demir ve Köse (2014), MH, SIBTEST ve LR yöntemlerini kullanarak DMF analizi çalışması yapmışlardır. Çalışma, PISA 2009 Matematik okuryazarlığı alt testinde yer alan maddeleri incelemek amacıyla cinsiyet ve kültürel özelliklere göre odak ve referans gruplar oluşturularak gerçekleştirilmiştir. Cinsiyet değişkenine ait analizler değerlendirildiğinde MH tekniği ile iki (1 ve 11), LR tekniği ile üç (1, 11 ve 17) ve SIBTEST tekniği ile dört (1, 5, 6 ve 11) maddede DMF tespit edilmiştir. Bu maddelerin çoğu erkekler lehine DMF göstermiştir. Kullanılan üç yöntemin belirlediği DMF'li madde sayısında ve DMF düzeyinde farklılık görülse de maddelerin avantaj sağladığı gruplar benzeşmektedir. Üç yöntemin de farklı sonuçlar vermesinin nedeni DMF belirlemede kullanılan sınıflandırmaların farklılığından kaynaklandığı ifade edilmiştir.

Doğan ve Öğretmen (2008), 2003 OKÖSYS Fen Bilgisi alt testine katılan 600.000 öğrenciden 3.345'i yansız olarak seçerek oluşturdukları örneklem ile DMF belirlemede kullanılan MH, Ki-Kare ve LR tekniklerini cinsiyet grupları üzerinde karşılaştırmışlardır. DMF belirlenen madde sayısının en çok MH tekniği ile belirlenmesine ek olarak bu üç yöntem DMF açısından farklı bulgular ortaya koymuştur. Fakat yine de yöntemlerden elde edilen Ki-Kare değerleri arasındaki

ilişkinin manidar olduğu tespit edilmiştir. Bu tespitler, DMF belirlemede bu üç yöntemin birbiri yerine kullanılabileceği yorumuna dayanak sağlamıştır.

Atalay-Kabasakal, Arsan, Gök ve Kelecioğlu (2014) çalışmalarında DMF belirlenmesinde MTK olabilirlik oranı, SIBTEST ve MH yöntemlerinin performanslarının I. tip hata ve güç kapsamında karşılaştırmasını yapmışlardır. Bu kapsamda Türkiye'deki gerçek veri uygulamalarına benzer simülasyon ile veri üretmişlerdir. Simülasyonda örneklem büyüklüğü oranı, yetenek dağılımı farklılıkları, test uzunluğu, DMF içeren madde oranı ve model türü değişken olarak seçilmiştir. Sonuçta tek biçimli DMF'nin belirlenmesinde gücünün en yüksek olduğu yöntem MH çıkarken I. tip hatanın en yüksek olduğu yöntemin ise SIBTEST olduğu bulunmuştur.

Narayanan ve Swaminathan (1996), DMF belirleme yöntemlerinden; MH, LR ve SIBTEST yöntemlerinin I. tip hata oranlarını ve tek biçimli olmayan DMF tespit etme kapasitelerini karşılaştırmak amacıyla bir çalışma yapmışlardır. Analizler; simülasyonla oluşturulan verilerle gerçekleştirilmiştir. Odak ve referans grup belirlemede ise örneklem büyüklüğü, gruplar arasındaki yetenek dağılımı, belirlenen DMF'li madde oranı, DMF etki büyüklüğü ve madde tipi özellikleri dikkate alınmıştır. Analizlerin çoğunda tek biçimli olmayan DMF tespitinde LR ve SIBTEST yöntemleri daha başarılı sonuçlar verirken I. tip hata oranlarında ise MH yöntemi diğer iki yöntemle oranla daha düşük değerler ortaya koymuştur.

Tek biçimli DMF'nin belirlenmesinde ön plana çıkan MH ve LR yöntemlerinin I. tip hata oranları ve istatistiksel güç değerlerinin karşılaştırılması amacıyla Erdem-Keklik (2014) tarafından bir çalışma yapılmıştır. Çalışmada odak ve referans grubun yetenek dağılımlarının, örneklem büyüklüklerinin ve alt grupların örneklem büyüklüklerinin oranlarının değiştirilerek hazırlandığı toplam 18 simülasyon koşulu kullanılmıştır. Analizler sonucunda benzer yetenek düzeyine sahip odak ve referans gruplarının MH ve LR yöntemine göre I. tip hata oranlarının birbirine yaklaşık ve nominal α düzeyine yakın, yetenek düzeyleri farklı gruplarda ise bu yöntemlere göre I. tip hata oranları arası farkın yüksek çıktığı görülmüştür. Özellikle farklı yetenek dağılımına sahip büyük örneklem gruplarında ise I. tip hata oranının daha yüksek olduğu gözlenmiştir. MH ve LR yöntemlerinin istatistiksel güçleri karşılaştırıldığında ise yetenek dağılımlarının benzer olduğu durumda yöntemlerin DMF belirleme gücünün yeterli olmadığı ifade edilmiştir. Yetenek dağılımlarının farklı olduğu

durumlarda ise her iki yönteminde DMF belirleme gücünün çok yüksek olduğu ifade edilmiştir. Örneklem büyüklüğünün artması ise bu yöntemlerin DMF belirleme gücünü olumsuz yönde etkilemiştir. Alan yazında bu bulguya paralel sonuçlara da aksi yönde sonuçlara da ulaşıldığı görülmüştür. Bu durumun madde sayısının az olmasından (madde sayısı:20) kaynaklanabileceği ifade edilmiştir.

Koyuncu, Aksu ve Kelecioğlu (2018), DMF belirleme yöntemlerinden KTK'ye ait ve MTK'ye ait olan yöntemlerin etkililiğini karşılaştırmak amacıyla bir çalışma yapmışlardır. Bu amaçla; MH, LR ve Olabilirlik Oranı yöntemleri PISA 2012 uygulamasının öğrenci anketinden elde edilen veriler üzerinde test edilmiştir. Odak ve referans gruplar öğrencilerin Matematik başarıları dikkate alınarak oluşturulmuştur. DMF belirlemek için kullanılacak yazılım programları birden fazla olacak şekilde tercih edilmiş ve onlar arasında da karşılaştırma yapılmıştır. MH analizleri için JMETRIK ve DIFAS yazılımları kullanılmıştır. Sonuçlar çok farklılaşmasa da aynı yöntem farklı programlarda farklı maddelerde DMF belirlemiştir. LR analizleri için Zumbo SPPS syntax'ı kullanılmış ve diğer yöntemlerden farklı olarak nerdeyse tüm maddelerde DMF tespit edilmiştir. Olabilirlik Oran yöntemi için ise IRTLRF ve IRTPRO programları kullanılmış ve MH analizlerine benzer sonuçlar vermişlerdir. Bu durum da farklı kuramsal temele dayanan farklı yöntemlerin farklı yazılım programlarında benzer tespitler yaptığı yorumunu doğrulamaktadır. Çalışmada bu sonucun alan yazındaki çalışmalarla benzeştiği de vurgulanmıştır.

Alan yazında yer alan bu çalışmalar gösteriyor ki, sözü geçen bu üç yöntem (MH, LR ve SIBTEST) birbirlerine yakın sonuçlar vermektedir. Aralarındaki küçük farklılıkların nedeni yöntemlerin kendilerine has etki büyüklükleri ve I. tip hatadan etkilenme dereceleri olarak ifade edilmektedir. I. tip hatadan etkilenme bakımından en güçlü SIBTEST en zayıf olan ise MH yöntemi olarak değerlendirilebilir. MH yönteminin daha fazla etkilenmesi daha fazla DMF belirlemesi ile ilişkilendirilmektedir. Bu çalışmalara göre örneklem büyüklüğü ve madde sayısı da I. tip hatayı etkilemektedir. Analizlerin elle yapılma külfetinin önüne geçen analiz programları da incelenmiş ve benzer sonuçlar verdiği tespit edilmiştir.

Cinsiyet değişkenine ilişkin yapılan yanlılık çalışmaları. Alan yazındaki yanlılık çalışmaları incelendiğinde gerek yabancı gerekse Türkiye özelindeki çalışmalarda alt grupların çoğunlukla cinsiyet değişkenine göre hazırlandığı

görülmektedir. Bu araştırmada da cinsiyet karşılaştırılması yapılmıştır. Alan yazında cinsiyet karşılaştırılmasının yapıldığı birkaç çalışmaya bu bölümde yer verilmiştir.

Köse (2015), yapmış olduğu DMF belirleme çalışmasında PISA 2009 öğrenci anketinde bulunan okul öğrenmelerine ilişkin algı ve öğretmene yönelik algı alt testine ait maddeleri cinsiyet ve farklı dil ve kültür değişkeni kapsamında incelemiştir. Aynı dil ve benzer kültüre sahip İngiltere ve İrlanda, aynı dil ve farklı kültüre sahip ABD ve İngiltere, farklı dil ve farklı kültüre sahip ABD ve Türkiye karşılaştırması kullanılarak ve cinsiyet değişkeni için de olmak üzere alt gruplar oluşturulmuştur. PISA maddeleri çoklu puanlandığı için aşamalı tepki modeli altında olabilirlik oran testi ile DMF belirleme araştırması yapılmıştır. Sonuçta, İrlanda-İngiltere örnekleminde iki ve ABD-Türkiye örnekleminde ise dört maddede DMF belirlenmiştir. Bu durum kültürel ve dilsel farklılıklar arttıkça maddelerdeki farklılaşmanın da arttığı yorumunu yaptırmaktadır. Cinsiyet değişkenine göre de iki madde DMF özelliği göstermiştir.

Öğretmen ve Doğan (2004), MTK çerçevesinde işaretli ve işaretsiz alan indeksleri kullanarak 2002 OKÖSYS Matematik alt testinin cinsiyete göre yanlılık içerip içermediğini araştırmışlardır. Bu kapsamda sınava katılan öğrencilerden Ankara ilindeki farklı bölgelerden 3.344 öğrenci tesadüfi olarak seçilmiştir. Yapılan analiz sonuçlarına göre işaretli ve işaretsiz alan indekslerinin büyüklükleri arasında anlamlı bir fark bulunmazken testin kadınların lehine yanlı davrandığı tespit edilmiştir.

Yurdugül ve Aşkar (2004b), OÖK-ÖSYS ait maddelerle, öğrencilerin cinsiyet özelliğine göre DMF belirleme araştırması yapmışlardır. Araştırmalarında MH yöntemi kullanarak analizlerini gerçekleştirmişlerdir. Araştırmanın bulgularına göre, Matematik alan testinde bir madde erkekler lehine DMF göstermiş ve bu madde için uzman bilgisine başvurulmuştur. Uzmanların çoğu bu maddede yer alan “basketbol” ifadesinin erkeklerin ilgilerini daha çok çekmesi nedeniyle ile daha başarılı olduklarını ifade etmişlerdir. Fen bilimleri alan testinde ise iki madde de yine erkeklerin lehine DMF göstermiştir. Bu maddelerde de “futbol topu” ve “araba” ifadelerinin geçmesi erkeklerin daha başarılı olmasıyla ilişkilendirilmiştir.

ÖSS 2005’e ait maddeler, Bakan-Kalaycıoğlu ve Kelecioğlu (2011) tarafından yanlılık açısından incelenmiştir. Çalışmanın amacı, maddelerin cinsiyet özelliğine

göre belli bir gruba avantaj sağlayıp sağlamadığını araştırmaktır. Bu amaçla öncelikle MH ve LR yöntemleri kullanılarak DMF analizi yapılmış ve DMF belirlenen maddeler sonrasında uzmanlara sunularak maddelerin yanlılığı hakkında karar verilmiştir. Analizde elde edilen bulgulara göre; Türkçe testinde hiç DMF'ye rastlanmazken Sosyal Bilimler testinde 7, Matematik testinde 3 ve Fen Bilimleri testinde ise 3 maddede DMF gözlenmiştir. Analiz sonuçları detaylı incelendiğinde MH ve LR yöntemleri genel olarak benzer fakat aynı şekilde DMF belirlememiştir. LR yöntemi, MH yöntemine oranla daha az sayıda ihmal edilmeyecek düzeyde DMF belirlemiştir. DMF belirlenen maddeler uzman görüşüne sunulduğunda, uzmanlar sadece tek bir maddenin yanlılık gösterdiği diğer maddelerin cinsiyete göre DMF belirlenmesinin nedenini madde etkisi olarak açıklamışlardır.

SBS 2009'da altıncı, yedinci ve sekizinci sınıfların aldığı Matematik ve Fen ve Teknoloji dersi alt testlerinin maddeleri cinsiyet özelliğine göre bir grubun lehine çalışıyor mu diye araştırmak amacıyla Karakaya (2012), bir yanlılık çalışması yapmıştır. Bu nedenle öncelikle maddelerin DMF analizi yapılmış sonra sonuçlar, uzman görüşüne sunulmuştur. DMF belirlemek için MH yöntemi kullanılmıştır. Bu analiz yöntemi ile Fen ve Teknoloji alt testinde altıncı sınıfta B düzeyinde 2 adet, sekizinci sınıfta B düzeyinde 3 adet DMF belirlenmiş olup yedinci sınıfta önemsenecek düzeyde DMF belirlenmemiştir. Genel olarak, fizik ve kimya konu alanına ait maddelerin erkeklerin lehine, biyoloji konu alanına ait maddelerin ise kadınların lehine DMF belirlendiği ifade edilmiştir. Matematik alt testinde ise altıncı sınıfta 1 adet B düzeyinde, yedinci sınıfta 1 adet B ve son olarak sekizinci sınıfta da 2 adet B düzeyinde DMF belirlenmiştir. Genel olarak, doğa ve geometri alanındaki maddeler erkekler lehine, algoritmik hesaplama gerektiren maddelerin ise kadınların lehine DMF belirlendiği ifade edilmiştir. Algoritmik hesaplama gerektiren maddelerde kadınlar lehine DMF belirlenmesinin alan yazınla çeliştiği vurgulanmıştır. DMF belirlenen Matematik alt testi ve Fen ve Teknoloji alt testine ait maddeler uzman görüşüne sunulduğunda, bu maddelerin yanlılık taşımadığı, bu sonucun da madde etkisinden kaynaklandığı ifade edilmiştir.

Alan yazında yer alan bu çalışmaların ortak yanı maddelerin bireylerin cinsiyet özelliklerine göre DMF belirlendiği yönündedir. Bu nedenle test hazırlama sürecinde bu durumun göz önünde bulundurulmasına dikkat çekilmektedir. Cinsiyete göre DMF belirlenmesinin nedeni; genellikle kadın ve erkek bireylerin

farklı alanlara ilgi duyması ya da toplumun cinsiyetlere yüklediği yükler veya cinsiyet özelliklerine göre kendilerini geliştirdiği alanlar olarak ifade edilmektedir. Bu çalışmalardan bazılarında da uzmanların düşüncelerine göre yanlılık oluşturmayacak farklılaşmalar tespit edilmiş ve bu durumun madde etkisinden kaynaklandığı ifade edilmiştir.

Okul türü değişkenine ilişkin yapılan yanlılık çalışmaları. Okul türü değişkeninin dikkate alınarak yapılan yanlılık çalışmaları da bu araştırma açısından önemli yer tutmaktadır. Alan yazın incelemesinde çok fazla örneğine rastlanmasa da son zamanlarda araştırmacılar tarafından incelenen bir konu olma özelliği taşımaktadır. Burada okul türünün oluşturduğu fark, bireylerin katıldığı öğretimin farkıdır. Özel okul, devlet okulu, yatılı bölge okulları, liseler açısından düşünüldüğünde Anadolu liseleri ve genel liseler gibi farklı okullar çalışmalarda kullanılmıştır. Bu bölümde de bu çalışmalara yer verilmiştir.

Özdemir (2003), 7 özel okul ve 8 devlet okulundan kadın ve erkek öğrencilerin Türkçe okuduğunu anlama yeteneğini ölçen çoktan seçmeli bir teste verdiği cevaplar ile MH ve LR yöntemini kullanarak DMF analizi yapmıştır. Çalışmada özel okul öğrencilerin lehine 1 tane ve erkekler lehine 1 tane olmak üzere B düzeyinde DMF belirlenmiştir. B düzeyinin çok yüksek olmaması nedeniyle hazırlanan test, yansız olarak kabul edilmiştir.

Alan yazında, DMF belirleme amaçlı yapılan çalışmalarda çokça kullanılan yöntemlerden olan MH ve LR yöntemlerini karşılaştırmak amacıyla, Gök, Kelecioğlu ve Doğan (2010) bir çalışma yapmışlardır. Çalışmalarında veri seti olarak 2005 OKS Matematik ve Fen Bilgisi alt testini cevaplayan öğrencilerin cinsiyet ve okul türü değişkenine göre gruplandırarak oluşturulan alt gruplara ait veriler kullanılmıştır. Matematik alt testi için yapılan analizlerde genel olarak LR yöntemi MH yöntemine oranla daha az DMF belirlemiştir. Cinsiyet değişkenine göre yapılan analizlerde iki yöntem arasındaki uyum incelendiğinde yöntemlerin DMF belirlenen maddelerin etki büyüklüğü bakımından uyum gösterirken DMF'li madde sayısı bakımından düşük uyum göstermektedir. Okul türü değişkenine ait bulgular incelendiğinde ise yöntemlerin gösterdiği uyum tam tersi sonuç vermiştir. MH ve LR analizlerinden elde edilen maddelere ait χ^2 değerlerine, Spearman'ın sıra farkları korelasyonu ile bakıldığında da önemli bir ilişki olmadığı sonucuna ulaşılmıştır. Fen Bilgisi alt testine

ait analizlerde de benzer şekilde LR yöntemiyle daha az sayıda DMF belirlenirken yöntemlerin gösterdiği uyum çok düşüktür.

SBS 2009 altıncı, yedinci ve sekizinci sınıfların aldığı Türkçe alt testi maddelerini cinsiyet ve okul türü özelliğine göre bir gruba avantaj sağlıyor mu sorusuna yanıt bulmak amacıyla Karakaya ve Kutlu (2012), bir yanlılık çalışması yapmıştır. Bu amaçla öncelikle MH ve LR yöntemlerini kullanarak DMF analizlerini gerçekleştirmişlerdir. Daha sonra sonuçları uzman görüşüne sunmuşlardır. Cinsiyet değişkenine göre yapılan çalışmalarda altıncı sınıfta her iki yöntemle de DMF belirlenmemiştir. Yedinci sınıfta MH yöntemi ile 2 adet, LR yöntemi ile 1 adet ve sekizinci sınıfta ise MH yöntemine göre 5 adet LR yöntemine göre de 3 adet maddede DMF belirlenmiştir. DMF belirlenen bu maddeler uzman görüşüne sunulduğunda, sadece sekizinci sınıf düzeyinde 1 maddede yanlılık olduğu kanaatine varılmıştır. Diğer maddelerin DMF belirtmesinin nedeni madde etkisi olarak açıklanmıştır. Okul türü değişkeni temel alındığında ise alt gruplar devlet okulu ve özel olarak oluşturulmuştur. Analiz sonucunda altıncı sınıfta MH yöntemi ile 7 adet ve LR yöntemi ile de 5 adet maddede, yedinci sınıfta MH yöntemi ile 2 adet ve LR yöntemi ile 1 adet maddede ve son olarak sekizinci sınıfta ise sadece MH yöntemi ile 2 adet maddede DMF belirlenmiştir. DMF belirlenen maddeler uzmanlar tarafından değerlendirildiğinde ise durumun madde etkisinden kaynaklandığı şeklinde yorumlar alınmıştır. Son olarak araştırmada kullanılan MH ve LR analiz yöntemlerinin DMF belirleme güçleri kıyaslandığında yöntemlerin bulgularının benzeşmediği sonucuna varılmıştır.

Şenferah (2015), çalışmasında 2010 SBS Matematik alt testine ait maddelerin cinsiyet ve okul türü değişkenine göre DMF analizini yapmıştır. Bu kapsamda MH ve LR yöntemlerini kullanarak tespit ettiği DMF'li maddeler için, bu maddelerin herhangi bir gruba yanlı davranıp davranmadığını araştırmak amacıyla Delphi tekniği yardımıyla uzman bilgisine başvurmuştur. Çalışmanın sonucuna göre; madde 8'in erkekler lehine yanlılık gösterdiği kabul edilmiştir. Yanlılık nedeni ise soru kökünde geçen "fabrika, iş kazası ve risk" ifadelerinin erkeklere daha yakın olduğu uzmanlar tarafından belirtilmiştir. Madde 3'ün ise özel okul öğrencilerinin lehine yanlılık gösterdiği kabul edilmiştir. Yanlılık nedeni ise soru kökünde geçen "geometri tahtası" ifadesinin materyal olarak özel okul öğrencileri tarafından tanınabileceği şeklinde belirtilmiştir.

Yıldırım (2015), çalışmasında 2012 SBS Matematik alt testine ait maddelerin cinsiyet ve okul türü değişkenine göre DMF analizini yapmıştır. Bu kapsamda MH ve LR yöntemlerini kullanarak tespit ettiği DMF'li maddeler için, bu maddelerin herhangi bir gruba yanlı davranıp davranmadığını araştırmak amacıyla Delphi tekniği yardımıyla uzman bilgisine başvurmuştur. Çalışmanın sonucuna göre; cinsiyete göre, bir maddenin kadınlar lehine bir maddenin ise erkekler lehine yanlılık taşıdığı kabul edilmiştir. Kadınlar lehine yanlılık taşıyan maddenin, ayrıntı ve soyut düşünme özelliği taşıması nedeniyle kadınların yeteneklerine daha yakın olduğu ifade edilmiştir. Erkekler lehine yanlılık gösteren maddenin ise, futbol maçlarındaki puan hesaplamasına benzer hesaplama içermesi nedeniyle erkeklerin ilgilerine daha yakın olduğu ifade edilmiştir. Okul türü değişkenine göre ise 5 maddenin özel okul öğrencilerinin lehine yanlılık gösterdiği ifade edilmiştir. Maddelerden ikisinin, özel okullarda kullanılan eğitim-öğretim materyalleri (tangram, simetri aynası vb.) nedeni ile yanlılık oluşturması, bir diğerinin maddenin oyun tarzında yazılması nedeniyle cep telefonu ve bilgisayar kullanan özel okul öğrencilerinin daha yetkili olması, bir diğerinin özel okullarda günlük hayat problemlerinin proje tabanlı öğretim tekniklerinin kullanılmasıyla, yanlılık gösteren son maddenin ise yanlılık gösterme nedeni, devlet okullarına oranla özel okullarda daha fazla soru çözülmesi ile ilişkilendirilmiştir.

Kelecioğlu, Karabay ve Karabay (2014), çalışmalarında 2009 SBS maddelerini cinsiyet ve okul türü değişkenine göre yanlılık açısından incelemişlerdir. Bu amaçla hem istatistiksel hem de uzman görüşüne dayalı araştırmalar yapmışlardır. Çalışmanın istatistiksel aşaması MH, LR ve SIBTEST yöntemlerinin kullanılması ile DMF belirleme süreci olarak gerçekleştirilmiştir. Bu aşamada okul türü değişkenine göre yapılan analizlerde en fazla Türkçe ve Matematik testinde DMF'ye rastlanmıştır. Test maddelerinin okul türü değişkenine göre hazırlanan alt gruplardan (devlet okulu-YİBO-özel okul) en fazla özel okul öğrencileri lehine DMF gösteren madde tespit edilmiştir. Cinsiyet değişkeni kapsamında yapılan analizlerde ise DMF belirlenen beş maddeden üçünün kadınların, ikisinin ise erkeklerin lehine farklılaştığı tespit edilmiştir. Çalışmanın ikinci aşamasında ise bu maddeler uzman görüşüne sunulmuş ve matematik testine ait üç maddenin yanlılık gösterdiği diğerlerinin ise madde etkisinden kaynaklanabileceği yönünde yorumlar alınmıştır.

ÖSS 2005'te Türkçe, Sosyal Bilimler, Matematik ve Fen Bilimleri maddelerine verilen cevapları, MH ve LR yöntemi kullanarak cinsiyet ve okul türü (Anadolu lisesi-genel lise-özel lise) değişkenleri kapsamında analiz eden Bakan-Kalaycıoğlu (2008), birçok maddede DMF tespit etmiştir. DMF tespit edilen maddeler uzman bilgisine sunulmuştur. Birçok maddede DMF tespit edilmesine karşın; cinsiyet değişkenine göre yalnızca Fen Bilimleri testindeki bir madde ve okul türü değişkenine göre de yine yalnızca Fen Bilimleri testindeki bir maddenin yanlılık gösterdiği kabul edilmiştir. Cinsiyet değişkenine göre yanlılık gösteren madde, erkeklerin lehine çalışmış olup Fizik dersi kapsamında hız konusuna aittir. Maddede geçen "otomobil" ifadesinin erkeklerin daha çok ilgisini çekmesi nedeniyle yanlılık oluşturduğu ifade edilmiştir. Okul türü değişkenine göre yanlılık gösteren madde ise Genel lise – Anadolu lisesi karşılaştırmasında Anadolu lisesi lehine, Genel lise – Özel lise karşılaştırmasında ise Özel lise lehine yanlılık oluşturan maddeler tespit edilmiştir. Uzmanlar bu yanlılığın nedenini, madde kökünde geçen "binici" ifadesinin İngilizce "rider" ifadesinden dilimize çevrildiğini belirterek maddenin, İngilizce eğitimi fazla olan bireylerin lehine olduğunu ifade etmiştir.

Bu çalışmaların hepsinde yanlılığın tespit edildiği ve maddelerin hep özel okul öğrencilerine yanlılık gösterdiği ifade edilmiştir. Uzmanların ifadesine göre bu durumun nedeni özel okulların fiziki imkânlarından ya da kullanılan materyallerden kaynaklanmaktadır. Bazı durumlarda da özel okullardaki öğretim tekniklerinin de etkili olduğu ifade edilmiştir. Bunlara ek olarak da maddelerin farklılaşmasına özel okulda öğrenim gören bireylerin sosyo-ekonomik düzeylerinin daha yüksek olması da neden olarak gösterilmiştir. Bu durum özellikle yatılı bölge okullarında öğrenim gören bireylerin başarı seviyesinin daha aşağıda kalmasının nedeni olarak da ifade edilmiştir. Ancak bu çalışmalarda yabancı dil alanındaki maddeler için yanlılık çalışması yapılmadığı görülmektedir.

İlgili Araştırmalar Özeti

Araştırmada, Türkiye'de uygulanan geniş ölçekli sınav olan TEOG Yabancı dil (İngilizce) alt testine ait maddelerin kullanılması nedeniyle öncelikle alan yazında yer alan yabancı dil verilerinin kullanıldığı yanlılık çalışmaları incelenmiştir. Bu incelemelerde TEOG Yabancı dil (İngilizce) alt testine ait maddelerle yapılan yanlılık çalışmalarına rastlanmamıştır. Ayrıca Türkiye'de uygulanan farklı isimlerdeki

ortaöğretime geçiş sınavlarında da yer alan yabancı dil testi ile ilgili bir kaç çalışma hariç yanlılık çalışmasına pek rastlanmamıştır. Ortaöğretime geçiş sınavları olmasa da alan yazında yer alan yerli ve yabancı çalışmalar incelendiğinde İngilizce maddelerinin doğasından ötürü herhangi bir gruba karşı kesin yanlılık oluştur denilebilecek bir kanıt ulaşılmamıştır.

Alan yazında yer alan çalışmaların birçoğunda maddelerin bireylerin cinsiyet özelliklerine göre DMF belirtip belirtmediği incelenmiştir. Bu nedenle test hazırlama sürecinde bu durumun göz önünde bulundurulmasına dikkat çekilmektedir. Maddelerin cinsiyete göre DMF belirtmesinin nedeni; genellikle kadın ve erkek bireylerin farklı alanlara ilgi duyması ya da toplumun cinsiyetlere yüklediği yükler veya cinsiyet özelliklerine göre kendilerini geliştirdiği alanlar olarak ifade edilmektedir. Bazı çalışmalarda ise kadın ve erkek bireylerin bilişsel ve sosyal becerilerinin benzeşmediği de vurgulanmıştır. Örneğin; şekil, tablo ve grafik içeren veya askeri, spor ve siyaset konulu maddelerin doğru yanıtlanma olasılığının erkekler için daha yüksek olduğunu belirten çalışmalar vardır. Benzer şekilde sözel becerilerin ön planda olduğu maddelerde ise kadınların daha avantajlı olduğunu belirten çalışmalar vardır. Bu durumda yapılması gereken bunun gibi olası kaynakların belirlenip madde yazım sürecinde bu kaynaklara dikkat edilmesidir.

Gerek yabancı gerekse Türkiye özelindeki çalışmaların hepsinde yanlılığın tespit edildiği ve maddelerin hep özel okul öğrencilerinin lehine çalıştığı ifade edilmiştir. Uzmanların ifadesine göre bu durumun, özel okulların fiziki imkânlarından ya da kullanılan materyallerden kaynaklandığı belirtilmiştir. Bazı çalışmalarda da özel okullardaki öğretim tekniklerinin bu durumda etkili olduğu ifade edilmiştir. Bunlara ek olarak maddelerin farklılaşmasına özel okulda öğrenim gören bireylerin sosyo-ekonomik düzeylerinin daha yüksek olması da neden olarak gösterilmiştir. Bu durum özellikle yatılı bölge okullarında öğrenim gören bireylerin başarı seviyesinin daha aşağıda kalmasının nedeni olarak da ifade edilmiştir. Ancak bu çalışmalarda yabancı dil alanındaki maddeler için yanlılık çalışması yapılmadığı görülmektedir.

Alan yazında yer alan çalışmalar gösteriyor ki, DMF belirlemede en fazla tercih edilen MH, LR ve SIBTEST yöntemleri birbirlerine yakın sonuçlar vermektedir. Aralarındaki küçük farklılıkların nedeni yöntemlerin kendilerine has etki büyüklükleri ve I. tip hatadan etkilenme dereceleri olarak ifade edilmektedir. I. tip hatadan etkilenme bakımından en güçlü SIBTEST en zayıf olana ise MH yöntemi olarak

rastlanmaktadır. MH yönteminin daha fazla etkilenmesi daha fazla DMF belirlemesi ile ilişkilendirilmektedir. Bu çalışmalara göre örneklem büyüklüğü ve madde sayısı da I. tip hatayı etkilemektedir. Analizlerin elle yapılma külfetinin önüne geçen analiz programları da incelenmiş ve benzer sonuçlar verdiği tespit edilmiştir.

Bazı çalışmalarda da istatistiksel olarak DMF belirttiği tespit edilen maddelerin, yapılan uzman görüşüleri yanlı olmadığı sonuçlarına ulaşılmıştır. Bu çalışmalarda farklılaşmanın madde etkisinden kaynaklandığı ifade edilmiştir.

Bölüm 3

Yöntem

Bu bölümde; araştırmanın türüne, evren ve örneklem grubunun özelliklerine, verilerin elde edilmesine, veri toplama araçlarının özelliklerine ve verilerin analizine yer verilmiştir. Ayrıca analizlerde kullanılan programlar, KTK'ye bağlı olarak hesaplanan test istatistikleri, madde istatistikleri ve faktör analizine de yer verilmiş olup testin DMF analizi için uygun özellikte olup olmadığı test edilmiş ve açıklanmıştır.

Araştırmanın Türü

TEOG 2015 - 2016 Eğitim - Öğretim Yılı II. Dönem İngilizce alt testine ait maddelerin, öğrencilerin cinsiyet ve öğrenim gördüğü okul türüne (özel-devlet) göre DMF özelliğinin incelenerek ayrıca uzman kanısına da başvurularak maddelerin belli bir gruba yanlılık gösterip göstermediği belirlenip var olan bir durum ortaya konulmaya çalışıldığı için bu çalışmanın türü, betimsel araştırma kapsamında kabul edilebilir. Betimsel araştırmalar; herhangi bir konuda halen var olan bir durumu, var olduğu haliyle betimlemeyi amaçlayan araştırma yaklaşımıdır (Karasar, 1999).

Evren ve Örneklem

Türkiye genelinde Temel Eğitimden Orta Öğretime Geçiş (TEOG) sistemi kapsamında 2015 – 2016 Eğitim – Öğretim Yılı II. Dönemde yapılan sınavda 1.134.623 öğrenci yabancı dil alt testine katılmış ve bunlardan 1.133.857 tanesi İngilizce alt testini cevaplamıştır. Bu öğrencilerin %52'si erkek, %48'si kadındır.

Millî Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğünden gerekli izinlerle, yukarıda belirtilen evrene ait verilerden rastgele seçilen bir örneklem talep edilmiştir. Örneklemde İngilizce alt testini yanıtlayan öğrencilerin %54'ü erkek %46'sı kadın ve %56'sı devlet okulu %44'ü özel okul öğrencisidir. Yani örneklem toplamı 9.108 öğrencinin 20 maddelik İngilizce alt testine ait maddelere verdiği cevaplardan ve bu öğrencilerin cinsiyet ve öğrenim gördüğü okul türü (özel-devlet) bilgilerinden oluşmaktadır. Ayrıca bu öğrencilerin hepsi A kitapçığını işaretlemiştir. Bu veri setinin değişkenlerine ait öğrenci sayıları ve yüzdeleri Tablo 7'de verilmiştir.

Tablo 7

Örneklem Grubunun Öğrenci Sayıları ve Yüzdeleri

	Devlet Okulu		Özel Okul		Toplam	
	Öğrenci Sayısı	Yüzde	Öğrenci Sayısı	Yüzde	Öğrenci Sayısı	Yüzde
Kadın	2.415	%47	1.829	%46	4.244	%46
Erkek	2.725	%53	2.139	%54	4.864	%54
Toplam	5.140	%56	3.968	%44	9.108	%100

Alan yazındaki çalışmalar incelendiğinde DMF analizlerinin, çalışılan grupların örneklem büyüklüğünden etkilendiği belirtilmiştir. Örneklem büyüklüğünün yüksek olması analiz yöntemlerinin gücünü olumlu yönde etkilemektedir (Ackerman ve Evans, 1992; Buzick ve Stone, 2011; Clauser ve Mazor, 1988; Finch ve French, 2007; Narayanan ve Swaminathan, 1996). MH, LR ve SIBTEST gibi analiz yöntemleri için her bir alt grup örnekleminin en az 200 kişi taşıma yeterliliğinin ifade edilmesine (Rogers ve Swaminathan, 1993; Zumbo, 1999) ek olarak örneklem büyüklüğünün artması DMF belirleme gücünü artırmaktadır (Narayanan ve Swaminathan, 1994; Rogers ve Swaminathan, 1993). İngilizce alt testi ile yapılan çalışmada ise Tablo 7’de görüldüğü üzere kadın (4.244), erkek (4.864), özel okul (3.968) ve devlet okulu (5.140) alt grupları DMF analizi için örneklem büyüklüğü bakımından istenilen değerlerdedir. Ayrıca alt grupların örneklem büyüklükleri birbirine eşit değil, evren değerlerine paralel olması beklenmektedir. Bu durumun I. tip hatayı önemli oranda azalttığı alan yazında birçok çalışmada ifade edilmektedir (Atalay Kabasakal, Arsan, Gök ve Kelecioğlu, 2014; Gotzmann ve Boughton, 2004; Kristjansson, Aylesworth, McDowell ve Zumbo, 2005).

Veri Toplama Süreci

Bu çalışmada kullanılan TEOG İngilizce alt testine ait veriler, gerekli izinlerle Millî Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğünden temin edilmiştir.

Veri Toplama Araçları

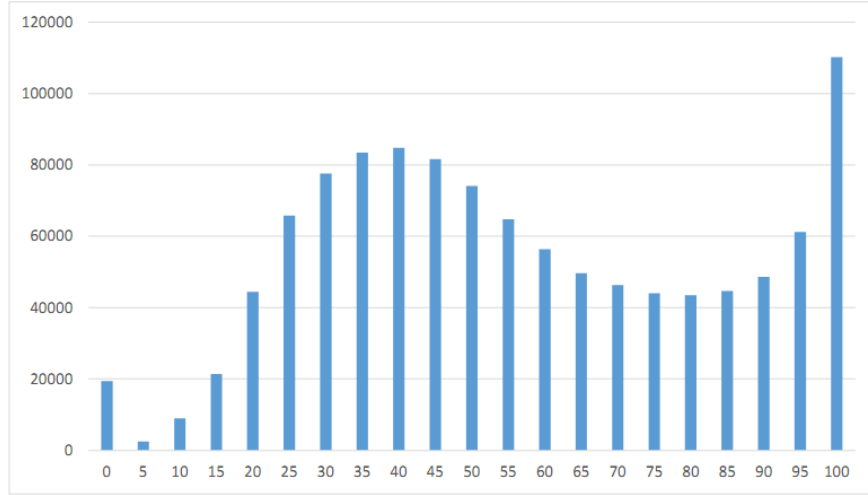
Bu araştırma kapsamında iki adet veri toplama aracı kullanılmıştır. Madde istatistikleri için MEB tarafından hazırlanan ve 2015 - 2016 Eğitim - Öğretim yılı II.

döneminde yine MEB tarafından uygulanan TEOG sınavına ait İngilizce alt testine ait maddeler (EK-A) kullanılmıştır. Uzman görüşü almak için ise hazırlanan Uzman Görüşü Anket Formu (EK-B) kullanılmıştır.

2015 - 2016 Eğitim - Öğretim Yılı II. Dönem TEOG İngilizce alt testi. TEOG İngilizce alt testi, sekizinci sınıfta okuyan öğrencilerin 6 temel dersten girmiş olduğu ortak sınavlardan bir tanesidir. Bu sınavlar, her eğitim – öğretim yılında I. dönem 1 adet ve II. dönem 1 adet olmak üzere toplamda 2 defa uygulanmaktadır. Bu çalışmada II. dönem uygulanan sınav verileri kullanılmıştır. Testte 20 madde yer almaktadır. Her bir madde 4 adet cevap şıkkına sahiptir ve düzeltme formülü uygulanmamıştır. Testin kapsamı, sekizinci sınıf müfredatının ilk sekiz üniteye ait kazanımlarını içermektedir. Bu sınav sisteminde, geçerli bir mazeret sebebiyle sınava giremeyen öğrenciler önceden yeri ve zamanı belirlenen mazeret sınavına girme hakkına sahip olmaktadır (MEB, 2016b). Mazeret sınavı için ortak sınavla benzer özellikte yeni maddeler hazırlanması nedeniyle çalışmada kullanılan veri setinde sadece ortak sınava giren öğrencilerin verileri yer almaktadır. Sınavda her alt test için A, B, C ve D olmak üzere 4 çeşit kitapçık hazırlanmış olup her kitapçıkta aynı maddeler sadece yerleri değişik olacak şekilde konumlanmıştır (MEB, 2016b). Alan yazın incelenmesi ile test maddelerinin sıralamasının, yanıtlayıcıları güdülemesi nedeniyle geçerliği ve dolayısıyla güvenirliliği etkileyen bir durum olduğu belirlenmiştir (Doğan, 2007; Tippets ve Benson, 1989). Bu nedenle bu çalışmada sadece A kitapçığını işaretleyen öğrencilere ait veriler kullanılmıştır.

Test istatistikleri. İstatistiksel analizlerin yapılabilmesi için öncelikle veri setinin analize uygun hale getirilmesi gerekmektedir. Bu amaçla veri setinde boş bırakılan ve birden fazla cevap işaretlenen maddeler yanlış kabul edilerek düzenlenmiştir. Sonrasında veri seti cevap anahtarı yardımıyla doğru yanıtlanan maddeye “1” yanlış yanıtlanan ve boş bırakılan maddeye “0” gelecek şekilde tekrar düzenlenmiştir. TAP (Test Analysis Program (versiyon 16.11.13))’ı kullanılarak betimsel istatistikler ve madde analizleri hesaplanmıştır.

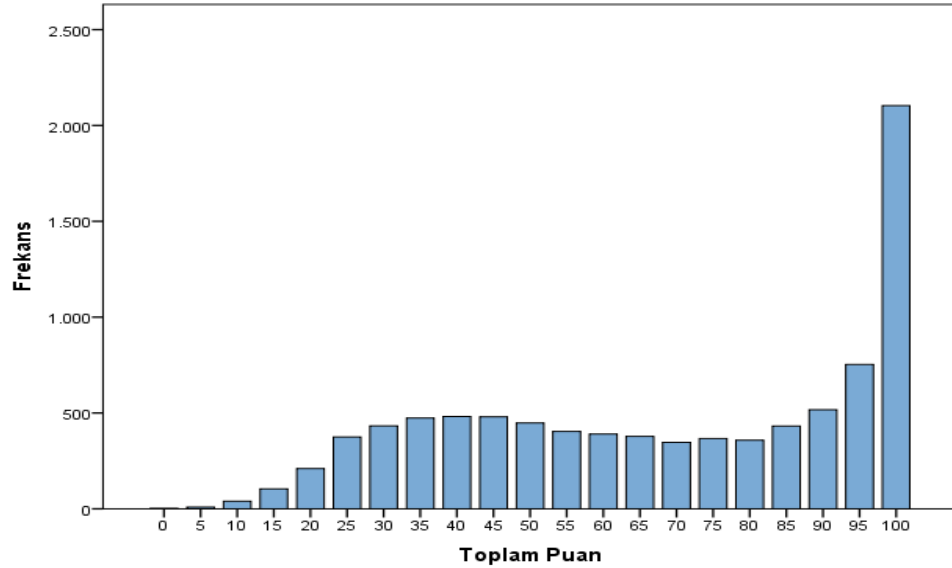
TEOG sınavına giren öğrenciler her dersten ayrı ayrı olmak üzere 100 puanlık sisteme göre değerlendirilmektedir. MEB’in yayımladığı bilgiler doğrultusunda İngilizce alt testini alan öğrencilerin 100 puanlık sisteme göre puan ortalamalarının 57,60 olduğu bilinmektedir. Evren grubundaki öğrencilerin puan dağılımı grafiği Şekil 3.’te verilmiştir (MEB, 2016a).



Şekil 3. Yabancı dil (İngilizce) alt testinin evrene ait puan dağılımı.

Evren değerlerine ait grafik incelendiğinde ortalama ve yüksek puanlarda aşırı yığılmanın olması ancak düşük puanlarda birikmenin az olması nedeniyle dağılım, sola çarpık ve basık olarak yorumlanmıştır (Pallant, 2007).

Bu çalışmada kullanılan örneklem grubuna ait puan ortalaması ise 68,50 olarak hesaplanmıştır. Bu verilere ait puan dağılımı grafiği ise Şekil 4.'te verilmiştir.

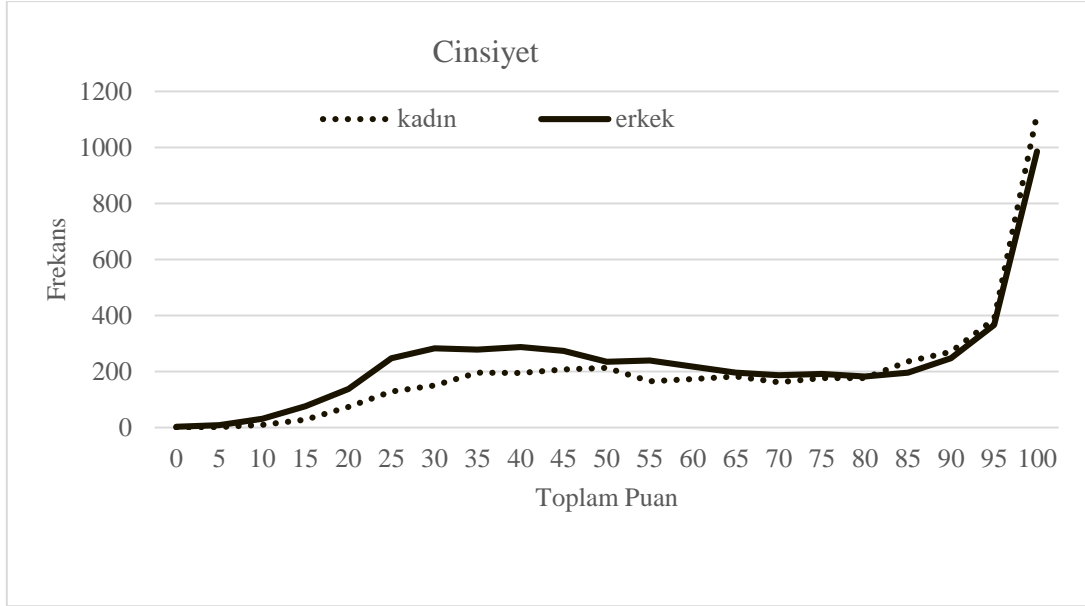


Şekil 4. Yabancı dil (İngilizce) alt testinin örnekleme ait puan dağılımı.

DMF analizleri için seçilen örnekleme ait puan dağılımının evrenin puan dağılımına oldukça benzer olduğu görülmektedir. Örnekleme ait histogram grafiğinde de evrene ait histogram grafiğine benzer şekilde ortalama ve yüksek

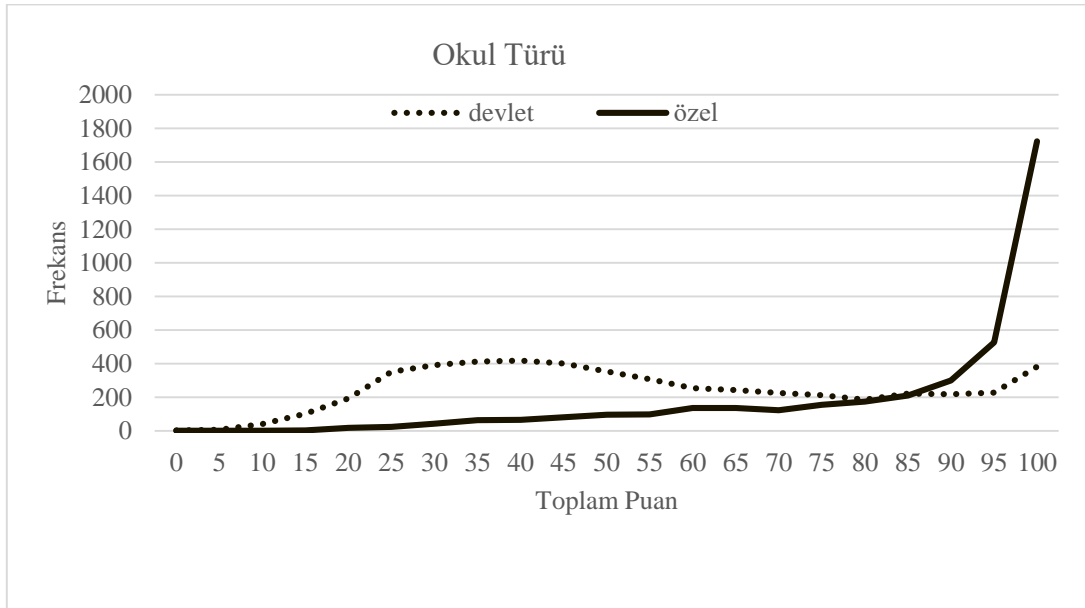
puanlarda aşırı yığılmanın olduğu ancak düşük puanlarda birikmenin az olduğu görülmüş olup dağılım, sola çarpık ve basık olarak yorumlanmıştır (Pallant, 2007).

Örneklem verilerinden elde edilen puan dağılımları bu çalışmada kullanılan değişkenlere göre ayrı ayrı da incelenmiştir. Cinsiyete ait puan dağılımları ile elde edilen grafik Şekil 5.'te verilmiştir.



Şekil 5. Cinsiyet değişkenine göre İngilizce alt testi puan dağılım grafiği.

Yorumlamada karşılaştırma yapabilmek amacıyla okul türü değişkenine göre elde edilen puan dağılım grafiği ise Şekil 6.'da verilmiştir.



Şekil 6. Okul türü değişkenine göre İngilizce alt testi puan dağılım grafiği.

Alt gruplara ait grafiklerin daha kolay yorumlanması adına bu grafiklere ait çarpıklık ve basıklık katsayı değerleri de Tablo 8’de verilmiştir.

Tablo 8

Örneklem ve Alt Gruplara Ait Basıklık ve Çarpıklık Değerleri

	Örneklem	Cinsiyet		Okul Türü	
		Kadın	Erkek	Devlet Okulu	Özel Okul
Basıklık	-1,290	-1,13	-1,36	-1,05	1,01
Çarpıklık	-0,318	-0,50	-0,16	0,30	-1,40

Hesaplanan istatistiksel değerlere göre cinsiyet değişkenine göre oluşturulan alt gruplardan, kadınların puan ortalaması (72,35), erkeklerin puan ortalamasından (65,05) az bir farkla yüksek çıkmıştır. Evren verileri incelendiğinde de benzer şekilde, kadınlar (61,97) erkeklere (51,71) göre daha yüksek puan almıştır (MEB, 2016a). Puan dağılım grafiğine bakıldığında ise puanların da benzer olduğu yorumu yapılabilmektedir. Tablo 8’de verilen değerler de incelendiğinde cinsiyet alt gruplarına ait grafiklerin evren ve örneklem geneline ait grafiklere benzer şekilde sola çarpık ve basık olduğu yorumu yapılabilmektedir.

Okul türü değişkenine göre düzenlenen alt gruplar incelendiğinde ise özel okul öğrencilerinin puan ortalamasının (85,4) devlet okulu öğrencilerinin puan ortalamasından (55,55) yüksek çıktığı görülmüştür. Okul türü gruplarına ait grafikler incelendiğinde de puanların farklılaştığı görülmektedir. Özel okul öğrencilerinin yüksek puanlarda toplandığı ve sivri bir dağılım gösterdiği görülürken özel okul verilerine ait grafik için sola çarpık ve basık yorumu yapılabilmektedir. Devlet okulu öğrencilerinin ise orta değerlerde yığıldığı ve basık bir dağılım gösterdiği görülürken grafiğin ise sağ çarpık ve dik olduğu yorumu yapılabilmektedir.

Örnekleme ait çarpıklık (-0,318) ve basıklık (-1,290) değerlerinin -2,00 ila +2,00 değerleri arasında olması normal dağılımı işaret etse de histogram grafiği incelendiğinde yüksek puanlarda yığılma olduğu görülmüş ve dağılımın sola çarpık ve basık olduğu yorumu yapılmıştır. (Büyüköztürk, Çokluk ve Köklü, 2011; Pallant, 2007).

TEOG İngilizce alt testine ait test istatistikleri evren ve çalışmada kullanılan örneklem grubu karşılaştırması yapılacak şekilde Tablo 9’da gösterilmiştir.

Tablo 9

İngilizce Alt Testinin Test İstatistikleri

		Ortalama Güçlük	Ayırt edicilik	Güvenirlik Katsayısı
	Evren	0,580	0,699	0,875
	Örneklem	0,685	0,657	0,909
Cinsiyet	Kadın	0,720	0,620	0,908
	Erkek	0,650	0,690	0,907
Okul türü	Devlet Okulu	0,550	0,620	0,861
	Özel Okul	0,850	0,410	0,894

Bu tabloda yer alan evrene ait değerler MEB (2016c) tarafından yayımlanmıştır.

Testin ortalama güçlüğü, test maddelerinin güçlük indekslerinin madde sayısına oranı alınarak hesaplanmaktadır (Baykul, 2010). Çıkan değer 0,00'a yaklaştıkça testin zorlaştığı, 1,00'e yaklaştıkça da testin kolaylaştığı yorumu yapılabilmektedir. Buna göre İngilizce alt testi, evren verisine göre 0,58 değerle orta güçlükte bir testtir. Bu çalışmada kullanılan örneklem grubunun ortalama güçlük düzeyi de 0,685 olarak hesaplanmış ve 0,50 değerine yakın olması nedeniyle orta güçlükte bir değer olarak yorumlanmıştır (Baykul, 2010). Çalışmada cinsiyet ve okul türü değişkenine göre alt grupların oluşturulması nedeniyle hazırlanan bu alt grupların da test istatistikleri ayrı ayrı hesaplanmıştır. Cinsiyet değişkeni açısından incelendiğinde kadınlar alt grubuna ait testin ortalama güçlük değerinin (0,72) erkeklere (0,65) göre biraz yüksek çıktığı görülmüştür. Cinsiyet gruplarına göre testin orta güçlükte olduğu ifade edilse de testin kadınlar grubu için biraz daha kolay olduğu yorumu yapılabilmektedir. Okul türü değişkeni açısından incelendiğinde de özel okul alt grubuna ait testin ortalama güçlük değerinin (0,85) devlet okulu (0,55) verilerine göre belirgin bir farkla yüksek çıktığı görülmüştür. Testin devlet okulu öğrencilerine göre orta güçlükte iken, özel okul öğrencileri için kolay olduğu yorumu yapılabilmektedir (Baykul, 2010).

Testin ayırt ediciliği, teste ölçülen özellik bakımından bilen ile bilmeyeni ayırt edebilme derecesidir (Baykul, 2010). Ayırt edicilik değeri 0,699 olan İngilizce alt testi evren değerinin 0,40 değerinden büyük olması sebebiyle ayırt edici olduğu yorumu yapılabilmektedir. Bu çalışmada kullanılan örneklem grubunun ortalama ayırt edicilik indeksi 0,657 olarak hesaplanmış ve evren değerlerine yakın bulunarak ayırt edici olarak yorumlanmıştır (Baykul, 2010). Cinsiyet ve okul türü değişkenine göre

hazırlanan alt gruplar incelendiğinde ise testinin ortalama ayırıcılık gücü kadın (0,62) ve erkeklerde (0,69) yaklaşık değer alırken, devlet (0,62) ve özel okul (0,41) öğrencilerinde belirgin bir farklılaşma görülmektedir.

Testin güvenilirliği, iç tutarlılık katsayısı hesaplama yöntemi ile tayin edilmiştir. İç tutarlılık, testteki maddelerin testin bütünü ile tutarlılığı olarak tanımlanmaktadır (Baykul, 2010). Güvenirlik katsayısı 0,00 ile 1,00 arasında bir değerdir. Güvenirlik katsayısının 0,00'a yaklaştıkça hata puanları varyansına yaklaştığı, 1,00'e yaklaştığında da gerçek puanlar varyansına yaklaştığı yani güvenilirliğin yükseldiği ifade edilmektedir. Buna göre İngilizce alt testinin evren verisine göre 0,875 değerle güvenilirliğinin yüksek olduğu yorumu yapılabilmektedir. Bu çalışmada kullanılan örneklem grubunun da güvenirlik katsayısı, iç tutarlılık yöntemi ile tayin edilmiş olup değeri .909 olarak kestirilmiştir. Bu değere göre, testin güvenilirliğinin yüksek olduğu yorumu yapılabilmektedir. Testin alt gruplara göre güvenirlik katsayıları ayrı ayrı hesaplandığında da Tablo 9'da da görüldüğü üzere kadın (0,908), erkek (0,907), özel okul (0,861), ve devlet okulu (0,894) alt gruplarının güvenirliklerinin yüksek olduğu söylenebilmektedir.

Madde istatistikleri. TEOG İngilizce alt testine ait madde istatistikleri evren ve çalışmada kullanılan örneklem grubu karşılaştırması yapılacak şekilde Tablo 10'da gösterilmiştir.

Tablo 10

Evren ve Örneklem Ait İngilizce Alt Testi Madde İstatistikleri

Madde No	Evren		Örneklem	
	ρ	r_{pb}	ρ	r_{pb}
1	0,46	0,51	0,57	0,57
2	0,49	0,56	0,60	0,66
3	0,45	0,46	0,56	0,58
4	0,69	0,54	0,79	0,57
5	0,50	0,58	0,64	0,65
6	0,39	0,48	0,52	0,60
7	0,57	0,56	0,69	0,63
8	0,71	0,55	0,80	0,57
9	0,78	0,36	0,84	0,37
10	0,78	0,53	0,86	0,54

Tablo 10

Evren ve Örneklem Ait İngilizce Alt Testi Madde İstatistikleri

Madde No	Evren		Örneklem	
	p	r_{pb}	p	r_{pb}
11	0,69	0,57	0,79	0,60
12	0,55	0,62	0,67	0,68
13	0,54	0,57	0,67	0,64
14	0,69	0,53	0,78	0,57
15	0,41	0,51	0,53	0,60
16	0,74	0,55	0,83	0,57
17	0,61	0,56	0,72	0,62
18	0,52	0,64	0,65	0,71
19	0,41	0,61	0,56	0,71
20	0,53	0,58	0,64	0,63

Not: Bu tabloda yer alan evrene ait değerler MEB (2016c) tarafından yayımlanmıştır. p: Madde Güçlük İndeksi, r_{pb} : Madde Ayırıcılık İndeksi

Madde güçlük indeksi (p), maddeyi doğru cevaplayanların maddeyi cevaplayanların tamamına oranı olarak hesaplanabilmektedir (Özçelik, 2010). Madde güçlük indeksi, 0,00 ila 1,00 arasında değer almaktadır. İndeks değerinin 0,00'a yaklaşması maddenin doğru cevaplanma olasılığını azaltırken 1,00'e yaklaşması ise maddenin doğru cevaplanma olasılığını artırmaktadır (Turgut ve Baykul, 2010). İngilizce alt testine ait maddelerin evren değerlerine göre güçlük indekslerine bakıldığında 0,41 ila 0,78 değerleri arasında ve çoğunlukla ortalama (0,50) değerlere rastlanılmaktadır. Özçelik (2010), İngilizce alt testine ait maddelerde olduğu gibi gerçek öğrenme düzeyiyle uyumlu olması nedeniyle madde güçlük indeksinin 0,50 civarında olması gerektiğini ifade etmektedir. Bu çalışmada kullanılan örneklem grubunun madde güçlük indeksleri hesaplandığında da 0,52 ila 0,84 arasında değerler aldığı görülmüştür. Hesaplanan indeks değerlerinin 0,50 civarında olması test maddelerinin orta güçlükte olduğunu göstermektedir (Gelbal, 2013).

Madde ayırıcılık gücü (r_{pb}), test edilen davranışa sahip olan cevaplayıcıları olmayanlardan ayırt etme derecesi olarak ifade edilmektedir (Özçelik, 2010). Madde ayırıcılık indeksi -1,00 ila +1,00 arasında değerler alabilmektedir. İndeks değeri 0,30 değerinden büyükse ayırıcılık gücü yüksek görüldüğünden bu maddelere testte yer

verilebilmektedir (Turgut ve Baykul, 2010). İngilizce alt testine ait maddelerin evren verilerine göre madde ayırıcılık indeks değerlerinin 0,36 ila 0,62 değerleri arasında ve tamamının ayırt edici olduğu görülmektedir. Bu çalışmada kullanılan örneklem grubunun madde ayırıcılık değerleri nokta çift serili ayırıcılık indeksleri hesaplanarak belirlenmiştir. Hesaplanan değerlerinin 0,37 ila 0,71 arasında olması nedeniyle bu maddeler, ayırtedici olarak yorumlanmıştır (Turgut ve Baykul, 2010).

DMF analizleri için cinsiyet ve okul türü değişkenlerine göre oluşturulan alt gruplar için ayrı ayrı madde analizleri yapılmış ve Tablo 11’de gösterilmiştir.

Tablo 11

Cinsiyet ve Okul Türü Gruplarına Göre Madde İstatistikleri

Madde No	Cinsiyet				Okul Türü			
	Kadın		Erkek		Devlet Okulu		Özel Okul	
	p	r _{pb}	p	r _{pb}	p	r _{pb}	p	r _{pb}
1	0,61	0,59	0,53	0,54	0,44	0,49	0,72	0,57
2	0,65	0,68	0,56	0,63	0,45	0,54	0,80	0,70
3	0,60	0,59	0,54	0,57	0,43	0,45	0,74	0,66
4	0,83	0,54	0,76	0,59	0,67	0,53	0,94	0,45
5	0,66	0,66	0,63	0,65	0,48	0,56	0,85	0,62
6	0,55	0,62	0,50	0,58	0,37	0,45	0,72	0,64
7	0,74	0,62	0,64	0,62	0,55	0,53	0,87	0,62
8	0,85	0,55	0,75	0,57	0,70	0,52	0,92	0,53
9	0,88	0,28	0,79	0,42	0,77	0,35	0,92	0,26
10	0,88	0,52	0,64	0,55	0,77	0,52	0,97	0,42
11	0,80	0,63	0,78	0,59	0,67	0,57	0,94	0,49
12	0,72	0,69	0,64	0,66	0,52	0,59	0,87	0,66
13	0,70	0,62	0,64	0,64	0,52	0,56	0,86	0,58
14	0,83	0,57	0,75	0,57	0,67	0,51	0,93	0,51
15	0,55	0,63	0,51	0,58	0,38	0,46	0,72	0,65
16	0,88	0,50	0,78	0,60	0,73	0,54	0,95	0,46
17	0,77	0,59	0,68	0,64	0,59	0,55	0,89	0,59
18	0,69	0,73	0,61	0,69	0,49	0,63	0,85	0,71
19	0,60	0,73	0,53	0,69	0,39	0,58	0,79	0,74
20	0,70	0,63	0,60	0,63	0,50	0,54	0,84	0,61

Not: p: Madde Güçlük İndeksi, r_{pb}: Madde Ayırıcılık İndeksi

Cinsiyete göre oluşturulan alt grupların madde güçlük indeksleri ayrı ayrı analiz edildiğinde kadınlara göre maddelerin güçlük değerinin en düşük 0,55 en yüksek 0,88 değeri aldığı erkeklere göre ise en düşük 0,50 ve en yüksek 0,79 değeri aldığı görülmektedir. Maddelerin tamamının kadınlara göre bir miktar daha kolay geldiği söylenebilir. Okul türü değişkenine göre oluşturulan alt grupların madde güçlük indeksleri analiz edildiğinde ise özel okul öğrencilerine göre maddelerin güçlük değerinin en düşük 0,72 en yüksek 0,97 değeri aldığı erkeklere göre ise en düşük 0,37 ve en yüksek 0,77 değeri aldığı görülmektedir. Maddelerin tamamının özel okul öğrencilerine daha kolay geldiği söylenebilir.

Cinsiyete göre oluşturulan alt grupların madde ayırıcılık indeksleri hesaplandığında ise kadınlarda 0,28 ila 0,73 ve erkeklerde 0,42 ila 0,69 arasında değerler aldığı görülmüştür. Maddelerin ayırıcılık katsayıları her iki grupta da bilen ile bilmeyenin ayırt edebilecek düzeyde güçlü olarak görülmesine rağmen 9. maddenin, kadınlar alt grubundaki 0,28 değeri istenen değerin (0,30) çok az altında kalmıştır. Okul türü değişkenine göre oluşturulan alt grupların madde ayırıcılık indeksleri hesaplandığında ise özel okul öğrencileri için 0,26 ila 0,74 ve devlet okulu öğrencileri için ise 0,35 ila 0,63 arasında değerler kaydedilmiştir. Maddelerin ayırıcılık katsayıları her iki grupta da bilen ile bilmeyenin ayırt edebilecek düzeyde güçlü olarak görülmesine rağmen 9. maddenin özel okul öğrencileri için 0,26 değeri istenen değerin (0,30) çok az altında kalmıştır.

Bu bölümde incelenen istatistiksel değerler İngilizce alt testi üzerinde yapılacak DMF çalışmaları için herhangi bir olumsuz durumun olmadığını göstermektedir.

Uzman Görüşü Anket Formu. 2015-2016 Eğitim-Öğretim Yılı II. Dönem TEOG İngilizce Alt Testine ait maddeler ile yapılan analizlerle sadece “B düzeyinde” DMF gösterdiği belirlenen maddeler, hazırlanan bir form (Ek-B) yardımıyla uzman bilgisine sunulmuştur. Bu formda özel okul öğrencileri lehine DMF gösteren 4 numaralı madde, erkekler lehine DMF gösteren 5 ve 11 numaralı madde ve kadınlar lehine DMF gösteren 9 ve 16 numaralı maddelere yer verilmiştir. Uzmanlara, bu maddelerin DMF göstermesinin nedeninin madde yanlılığından mı yoksa madde etkisinden mi kaynaklandığı sorulmuş olup ayrıca, düşüncelerinin nedenlerini de ifade etmeleri istenmiştir.

Verilerin Analizi

TEOG İngilizce alt testine ait maddelerin yapılan betitimsel istatistikler ile DMF analizleri yapmaya uygun olduğu görülmüştür. Bu bölümde bir de faktör analizi yapılmıştır. Sonraki aşamada MH, LR ve SIBTEST yöntemleri kullanılarak DMF analizleri yapılmıştır. DMF tespit edilen maddeler uzman görüşüne sunulmuştur. Son olarak da analizde kullanılan DMF belirleme yöntemlerinin birbirleriyle uyumu incelenmiştir.

Faktör analizi. DMF analizleri öncesinde, İngilizce alt testi maddelerinin aynı yapıyı ölçüp ölçmediğini tespit etmek amacıyla faktör analizi yapılmıştır. Bu amaçla öncelikli olarak faktör yapısını belirlemek için IBM SPSS Statistic (versiyon 22) programı kullanarak Açıklayıcı Faktör Analizi (AFA) ve sonrasında da belirlenen faktör yapısının doğruluğunu test etmek için LISREL 8.80 programı kullanılarak Doğrulayıcı Faktör Analizi (DFA) yapılmıştır.

İlk olarak testin AFA için uygunluğunu test etmek amacıyla Kaiser-Meyer-Olkin ve Bartlett Küresellik Testi değerlerine bakılmıştır. Şekil 7.'de görüldüğü üzere, KMO değerinin yüksekliği ve Bartlett Küresellik Testinin anlamlı çıkması veri setinin faktör analizi için uygun olduğunu göstermiştir (Büyüköztürk, 2002; Çokluk, Şekercioğlu ve Büyüköztürk, 2010).

KMO ve Bartlett Küresellik Testi

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	0,969	
Bartlett's Test of Sphericity	Approx. Chi-Square	57383,934
	df	190
	Sig.	0,000

Şekil 7. KMO ve Bartlett Küresellik Testi sonuçları.

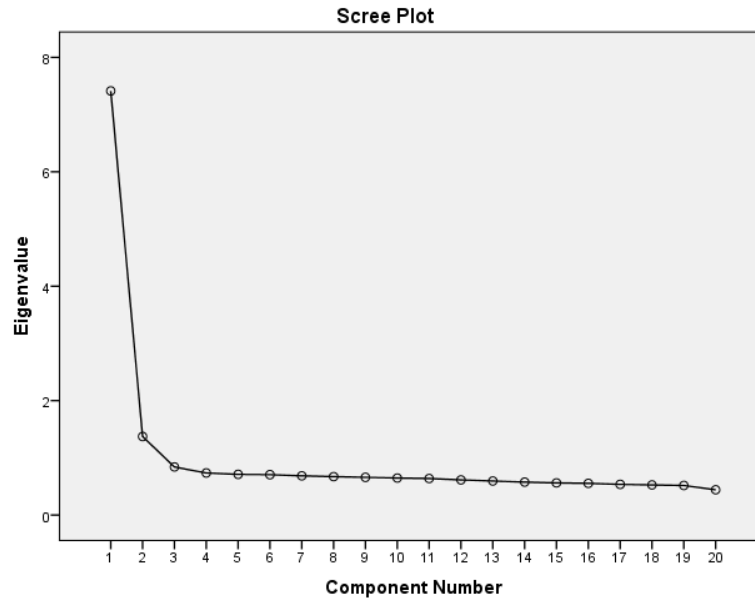
Testin faktör yapısını belirlemek için Öz Değer (Eigen Value) ve Yamaç Birikinti Grafiğine (Scree Plot) bakılmıştır. Tablo 12'de görüldüğü üzere ilk faktör değerinin diğer faktörlerden belirgin olarak büyük olduğu görülmüştür. Bu faktör tek başına toplam varyansın %37'sini açıklamaktadır. Bu oran, oldukça yüksek olarak değerlendirilebilir (Pallant, 2007).

Tablo 12

Öz Değerler Tablosu

Faktör	Öz Değerler		
	Toplam	Varyans Yüzdesi	Toplam Varyans Yüzdesi %
1	7,413	37,067	37,067
2	1,373	6,864	43,931
3	0,840	4,202	48,132
4	0,735	3,677	51,810
5	0,711	3,557	55,366
6	0,705	3,526	58,892

Yamaç Birikinti Grafiği incelendiğinde de grafiğin ikinci noktadan itibaren düzleşmeye başladığı görülmüştür.



Şekil 8. Yamaç birikinti grafiği.

Öz Değerler Tablosu'ndaki birinci faktörün yüksek değeri ve ayrıca Yamaç Birikinti Grafiği'ndeki eğimin ikinci nokta itibariyle düzleşmesi nedeniyle testin tek boyutlu bir faktör yapısına sahip olduğu çıkarımı yapılmıştır. İngilizce alt testine ait madde faktör yükleri EK-C'de verilmiştir.

AFA'da elde edilen tek faktörlü yapıyı test etmek amacıyla DFA yapılmıştır. DFA'da model uyumunu incelemek için birden fazla model uyumu istatistiğinin dikkate alınması tavsiye edilmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2002). Tek faktörlü yapı için Yaklaşık Hataların Ortalama Karekökü (RMSEA) değeri 0,059

olarak hesaplanmıştır ve bu değer yüksek model uyumunu göstermektedir (Hu ve Bentler, 1999; Kelloway, 1998; Kliene, 2005). Ayrıca Normed Fit Index (NFI), Comparative Fit Index (CFI) ve Goodness of Fit Index (GFI) değerleri 0,98 olarak hesaplanmıştır. CFI, NFI ve GFI değerleri için 0,90 üzeri değerler model uyumu açısından yüksek değerler olarak kabul edilmektedir (Hu ve Bentler, 1999; Kelloway, 1998; Kliene, 2005). Elde edilen tüm bu yüksek model uyum değerleri veri seti için tek fatörlü yapının uygun olduğunu göstermektedir.

DMF analizleri. DMF belirleme amaçlı geliştirilen birçok analiz yönteminin olmasına rağmen bu yöntemlerden herhangi birinin kesin doğru sonucu verir yorumu yapılabilmesi için, yapılan çalışmaların henüz yeterli olmaması nedeniyle araştırmalarda birden fazla yöntemin kullanılması haklı görülmektedir (Camilli ve Shepard, 1994; Osterlind ve Everson, 2009; Kristjansson, Aylesworth, McDowell ve Zumbo, 2005). Bu sebeple bu çalışmada DMF incelemesi için MH, LR ve SIBTEST yöntemleri tercih edilmiştir. Bu yöntemlerin tercih edilmesinin bir başka sebebi ise, geniş ölçekli testlerde I.tip hatanın kontrolü açısından güvenilir sonuçlar vermeleridir (Atalay, Gök, Kelecioğlu ve Arsan, 2012; Erdem-Keklik, 2014; Uyar ve Kaya-Uyanık, 2016). Ayrıca bu yöntemler hesaplama ve yorumlama kolaylığı da sağlamaktadır (Bertrand ve Boiteau, 2003; Kelecioğlu, Karabay ve Karabay, 2014).

İngilizce alt testine ait verilerinin betimsel istatistikleri ve faktör analizleri incelendikten sonra testin DMF analizleri yapılmıştır. Test maddelerin DMF özelliği gösterip göstermediği KTK'ye dayalı yöntemlerden parametrik olmayan Mantel – Haenszel yöntemi, yine KTK'ye dayalı yöntemlerden parametrik olan Lojistik Regresyon ve MTK'ye dayalı parametrik olmayan SIBTEST yöntemi kullanılarak test edilmiştir. Bu üç yöntemden en az ikisine göre orta ve yüksek düzeyde DMF gösteren maddeler DMF'li madde olarak kabul edilmiştir.

MH yöntemine göre DMF belirlemek için EZDIF (Waller, 2005) programı yardımıyla maddelerin $\Delta \alpha_{MH}$ değerleri hesaplanmıştır. DMF belirlenen maddelerin etki büyüklüklerinin belirlenmesi amacıyla da alan yazında yaygın kullanımı olan Zieky (1993)' e ait sınıflama kullanılmıştır.

LR yöntemine göre DMF belirleme süreci iki aşamada gerçekleştirilmiştir. Öncelikle oluşturulan alt gruplar EZDIF programı yardımıyla analiz edilerek maddelerin toplam puan, grup ve grup*toplam puan etkileşimine ait β estimate

değerleri ve bu değerlere ait anlamlılık düzeyleri hesaplanmıştır. Grup ve toplam puan*grup etkileşimine ait β estimate değerlerinin anlamlılığı 0,01 ölçüsünde karşılaştırılarak maddelerin DMF göstermesi ve DMF biçimi (TB ve TBO) hakkında karar verilmiştir. DMF gösteren maddelerin etki büyüklüğünün belirlenmesi için serbestlik derecesi 2 olan χ^2 değerinin hesaplanması gerekmektedir (Zumbo ve Thomas, 1990). Bu aşamada Zumbo (1999)'nun syntaxı (Ek-E) kullanılarak IBM SPSS Statistic (versiyon 22) programında LR analizi yapılmıştır. Bu analiz 3 model kurularak gerçekleştirilmektedir. 1. modele toplam puan, 2. modele grup ve 3. modele grup*toplam puan değişkenlerinin etkileşimi dâhil edilmektedir. Analiz sonucunda 2 serbestlik dereceli χ^2 değerine karşılık gelen Nagelkerge (R^2) değerleri hesaplanmaktadır (Zumbo, 1999). Bu ifade 3. model ile 1. modele ait R^2 değerlerinin farkı (ΔR^2) alınarak hesaplanır ve DMF'nin etki büyüklüğünü belirtir (Swaminathan ve Rogers, 1990). DMF belirlenen maddelerin etki büyüklüğünü yorumlamak amacıyla çalışılan veri örnekleminin büyük olması nedeniyle Jodoin ve Gierl (2001)'in oluşturduğu sınıflama kullanılmıştır. Jodoin ve Gierl (2001) oluşturdukları sınıflamayı, simülasyon çalışmalarından elde ettikleri sonuçlar doğrultusunda özellikle büyük örneklem gruplarında I. tip hatayı azalttığı şeklinde savunmuşlardır.

SIBTEST yöntemine göre DMF belirlemek için SIBTEST (Stout, 2005) programı kullanılarak maddelerin β değerleri hesaplanmıştır. Maddelerin DMF düzeylerinin belirlenmesi amacıyla Roussos ve Stout (1996) 'nın yapmış olduğu sınıflama kullanılmıştır.

Analizlerin son aşamasında $\Delta \alpha_{MH}$, ΔR^2 ve β değerlerinin işaretlerine bakılarak maddenin odak grubun mu yoksa referans grubun mu lehine olduğuna karar verilmiştir.

DMF gösteren maddelerin yanlılık durum tespiti. DMF gösteren maddelerin yanlılıktan ötürü mü, yoksa madde etkisinden ötürü mü farklılaştığına karar vermek amacıyla hazırlanan form (EK-B) 8 devlet okulu öğretmeni, 8 özel okul öğretmeni ve 4 ölçme ve değerlendirme uzmanının görüşüne sunulmuştur. Uzmanlardan gelen dönütler ve alan yazındaki çalışmalar yardımıyla da maddeler hakkında ortak sonuca ulaşılmıştır.

MH, LR ve SIBTEST analizleri ile edilen sonuçların uyumu. İngilizce alt testine ait verilerinin cinsiyet ve okul türü değişkenine göre MH, LR ve SIBTEST analiz sonuçlarının uyumu incelenmiştir. Bu amaçla IBM SPSS Statistic (versiyon 22) programı kullanılarak yapılmıştır. Her üç yöneme göre 20 madde için DMF yoksa "0", A düzeyinde DMF varsa "1", B düzeyinde DMF varsa "2" ve C düzeyinde DMF varsa "3" yazılarak kategorik bir veri oluşturulmuştur. Daha sonra her üç yöneme göre DMF tespit edilen maddelerin etki büyüklüklerinin uyumu için 0,01 anlamlılık düzeyinde spearman-brown korelasyon hesaplaması yapılmıştır. Hesaplanan değer $0,10 > r_s > 0,29$ aralığında ise uyum düşük, $0,30 > r_s > 0,49$ aralığında ise uyum orta ve $0,50 > r_s > 1,0$ aralığında ise uyum yüksek olarak belirtilmiştir (Pallant, 2007).

Bölüm 4

Bulgular ve Yorumlar

Bu bölümde araştırmanın problemine ilişkin analizlerin bulgularına ve bu kapsamda yapılan yorumlara yer verilmiştir.

Alt Problem 1' e İlişkin Bulgular ve Yorumlar

Alt Problem 1. TEOG 2016 Nisan (II. Dönem) sınavının İngilizce alt testine ait maddeleri cinsiyete göre yanlılık göstermekte midir?

İngilizce alt testi maddelerinin cinsiyet değişkenine göre DMF değerleri MH, LR ve SIBTEST yöntemleri kullanılarak ayrı ayrı hesaplanması nedeniyle her yöntem için alt başlıklar altında analiz bulguları ayrı ayrı verilmiştir.

Cinsiyete göre MH yöntemi analiz bulguları. İngilizce alt testine ait maddelerin cinsiyet değişkeni temelinde DMF değerlerinin MH yöntemine göre hesaplanması için EZDIF programından faydalanılmıştır. Analiz sonuçları Tablo 13'te verilmiştir. Cinsiyete göre İngilizce alt testi maddelerinin MH analizi için yapılan EZDIF program çıktısı EK-Ç'de verilmiştir.

Tablo 13

İngilizce Alt Testi Maddelerinin Cinsiyete Göre MH Analiz Sonuçları

Madde No	χ^2	p	α	ΔMH	SE ΔMH	DMF Düzeyi
1	3,111	0,078	0,911	0,220	0,123	
2	1,35	0,201	0,927	0,179	0,136	
3	1,231	0,267	1,064	- 0,145	0,128	
4	1,972	0,160	0,911	0,218	0,152	
5	38,219	0,000*	1,439	0,856	0,139	A
6	7,194	0,007*	1,162	- 0,353	0,130	A
7	10,293	0,001*	0,827	0,447	0,138	A
8	27,558	0,000*	0,708	0,812	0,154	A
9	52,411	0,000*	0,629	1,090	0,151	B
10	1,577	0,209	1,105	- 0,234	0,181	
11	67,070	0,000*	1,756	-1,323	0,162	B
12	0,032	0,859	0,987	0,030	0,145	

Tablo 13

İngilizce Alt Testi Maddelerinin Cinsiyete Göre MH Analiz Sonuçları

Madde No	χ^2	p	α	ΔMH	SE ΔMH	DMF Düzeyi
13	3,991	0,046	1,125	- 0,276	0,137	
14	7,837	0,005*	0,835	0,423	0,150	A
15	25,434	0,000*	1,332	- 0,673	0,133	A
16	42,220	0,000*	0,626	1,100	0,169	B
17	8,789	0,003*	0,836	0,422	0,141	A
18	0,154	0,695	1,028	- 0,065	0,151	
19	6,595	0,010	1,179	- 0,388	0,149	
20	10,306	0,001*	0,832	0,433	0,134	A

Not: * $p < 0,01$, $|\Delta \alpha MH| < 1$ ise A Düzeyi, $1 \leq |\Delta \alpha MH| < 1,5$ ise B Düzeyi, $|\Delta \alpha MH| \geq 1,5$ ise C Düzeyi

MH analiz bulguları öncelikle 0,01 anlamlılık düzeyinde incelenerek DMF gösteren maddeler belirlenmiştir. Daha sonra Tablo 13'te verilen ΔMH katsayı değerleri Zieky (1993)'nin sınıflama ölçütü kapsamında incelenmiş olup maddelerin DMF düzeyi tespit edilmiştir. Buna göre üç madde (9, 11 ve 16) B düzeyinde DMF göstermiştir. Sekiz madde (5, 6, 7, 8, 14, 15, 17 ve 20) ise DMF düzeyi A olarak belirlenmiş olup ihmal edilebilir seviyededir. B düzeyinde DMF gösterdiği belirlenen maddelerin avantaj sağladığı grup bilgisi Tablo 14'te verilmiştir.

Tablo 14

Cinsiyete Göre MH Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi

Madde No	ΔMH	DMF Düzeyi	Avantaj Sağladığı Grup
9	1,090	B	Odak grup (Kadın)
11	-1,323	B	Referans grup (Erkek)
16	1,100	B	Odak grup (Kadın)

Not: $\Delta MH < 0,00$ ise referans grup lehine, $\Delta MH > 0,00$ ise odak grup lehine

Maddelerin hangi gruba avantaj sağladığını belirlemek için ΔMH katsayısının işareti incelenmiştir. ΔMH katsayısının pozitifliği maddenin odak grubun lehine, negatifliği ise referans grubun lehine DMF gösterdiğini ifade etmektedir. Bu kapsamda madde 9 ve 16 kadınlar lehine, 11 ise erkekler lehine DMF göstermiştir.

Cinsiyete göre LR yöntemi analiz bulguları. İngilizce alt testine ait maddelerin cinsiyet değişkenine göre DMF özelliğinin incelenmesi için LR yöntemi ile analiz yapılmıştır. Analiz iki aşamada gerçekleştirilmiştir. Öncelikle EZDIF programının çıktısı sayesinde maddelerin DMF gösterip göstermediği ve gösteriyorsa DMF biçimi hakkında yorum yapılabilmesi için gerekli β estimate değerlerinin anlamlılık bilgisine ulaşılmıştır. Cinsiyete göre İngilizce alt testi maddelerinin LR analizi için yapılan EZDIF program çıktısı EK-D'de verilmiştir. Program çıktısından elde edilen grup ve grup*toplam puan değişkenlerine ait β değerleri ile β değerine ait anlamlılık değerleri Tablo 15'te verilmiştir.

Tablo 15

İngilizce Alt Testi Maddelerinin Cinsiyete Göre LR Analizi β Estimate Değerleri

Madde No	Grup		Grup*Toplam Puan		DMF Biçimi
	β	p	β	p	
1	0,599091	0,000121*	-0,049153	0,000007*	TBO
2	0,848831	0,000002*	-0,071070	0,000000*	TBO
3	0,432104	0,006525*	-0,025227	0,023351	TB
4	-0,396465	0,022636	0,035012	0,043174	
5	0,608886	0,000506*	-0,017314	0,205052	TB
6	0,945163	0,000000*	-0,055055	0,000003*	TBO
7	0,095957	0,560601	-0,024487	0,078468	
8	0,148715	0,404366	-0,052456	0,003887*	TBO
9	-1,034291	0,000000*	0,056554	0,000012*	TBO
10	-0,118726	0,600688	0,034350	0,213357	
11	0,526613	0,007228*	0,009013	0,656523	TB
12	0,620810	0,000877*	-0,054345	0,000566*	TBO
13	0,013629	0,934382	0,011201	0,404718	
14	0,351410	0,000000*	-0,053801	0,001742*	TBO
15	1,220927	0,000000*	-0,065693	0,000000*	TBO
16	-0,916000	0,000007*	0,055677	0,013976	TB
17	-0,337199	0,043106	0,016653	0,258803	
18	0,991561	0,000002*	-0,081380	0,000002*	TBO
19	1,064947	0,000000*	-0,066116	0,000011*	TBO
20	0,127626	0,438375	-0,024583	0,058177	

Not: * $p < 0,01$

LR analizinde, grup ve grup*toplam puan etkileşimine ait β estimate değerlerinin 0,01 düzeyinde anlamlılığı incelenmiştir. Sadece grup değişkeninin β estimate değerinin anlamlı olduğu durumlarda TB DMF ve grup*toplam puan etkileşiminin β estimate değerinin anlamlı olduğu durumlarda ise TBO DMF yorumu yapılabilmektedir (Camilli ve Shepard, 1994; Penfield ve Camilli, 2007; Rogers ve Swaminathan, 1990). Bu durumda İngilizce alt testine ait maddelerin on üç tanesinin TBO ve üç tanesinin ise TB DMF içermektedir. Fakat bu analiz, DMF'nin etki büyüklüğü hakkında bilgi sağlamak için yeterli değildir. Bu aşamada SPSS programında Zumbo (1999)'nun syntaxı (EK-E) kullanılarak elde edilen 2 serbestlik dereceli χ^2 değerine karşılık gelen Nagelkerge (R^2) değerleri yardımıyla ΔR^2 hesaplanmıştır. ΔR^2 değeri 3. modele ait R^2 ile 1. modele ait R^2 değerlerinin farkı (ΔR^2) olarak hesaplanır ve DMF'nin etki büyüklüğünü belirtir (Swaminathan ve Rogers, 1990). Cinsiyet değişkenine göre İngilizce alt testi maddelerine ait χ^2 değerlerinin anlamlılığı ve ΔR^2 hesaplamaları Tablo 16'da verilmiştir.

Tablo 16

İngilizce Alt Testi Maddelerinin Cinsiyete Göre LR Analiz Sonuçları

Madde No	R^2 değerleri						DMF Düzeyi
	χ^2	p	Model-1	Model-2	Model-3	$\Delta R^2 (R_3^2 - R_1^2)$	
1	21,826	0,000*	0,397	0,397	0,400	0,003	A
2	30,395	0,000*	0,524	0,524	0,527	0,003	A
3	8,4580	0,015	0,418	0,419	0,419	0,001	
5	47,152	0,000*	0,528	0,533	0,533	0,005	A
6	33,217	0,000*	0,440	0,441	0,443	0,003	A
8	33,791	0,000*	0,474	0,477	0,478	0,004	A
9	50,910	0,000*	0,224	0,232	0,235	0,011	A
11	79,316	0,000*	0,536	0,544	0,544	0,008	A
12	11,953	0,003*	0,573	0,574	0,575	0,002	A
14	16,094	0,000*	0,471	0,472	0,473	0,002	A
15	62,313	0,000*	0,448	0,451	0,454	0,006	A
16	45,669	0,000*	0,519	0,524	0,525	0,006	A
18	25,661	0,000*	0,618	0,618	0,620	0,002	A
19	28,863	0,000*	0,602	0,603	0,604	0,002	A

Not: * $p < 0,01$ $\Delta R^2 < 0,035$ ise A Düzeyi, $0,035 \leq \Delta R^2 < 0,070$ ise B Düzeyi, $\Delta R^2 \geq 0,070$ ise C Düzeyi

DMF gösterdiği belirlenen maddelerin etki büyüklüğünün belirlenmesi amacıyla hesaplanan ΔR^2 değerleri, Jodoin ve Gierl (2001)' in oluşturdukları sınıflama ile karşılaştırılmış ve DMF tespit edilen tüm maddelerin (13 madde) ihmal edilebilir düzeyde (A düzeyi) olduğu bulgusuna ulaşılmıştır.

Cinsiyete göre SIBTEST yöntemi analiz bulguları. İngilizce alt testine ait maddelerin cinsiyet değişkenine göre DMF gösterip göstermediğinin belirlenmesi amacıyla SIBTEST analizi yapılmıştır. Cinsiyete göre İngilizce alt testi maddelerinin SIBTEST analizi için yapılan SIBTEST program çıktısı EK-F'de verilmiştir.

Tablo 17

İngilizce Alt Testi Maddelerinin Cinsiyete Göre SIBTEST Analiz Sonuçları

Madde No	β	SE β	p	DMF Düzeyi
1	0,013	0,011	0,257	
2	0,018	0,010	0,095	
3	- 0,010	0,011	0,374	
4	0,015	0,010	0,121	
5	- 0,065	0,010	0,000*	B
6	- 0,028	0,011	0,011	
7	0,035	0,010	0,001*	A
8	0,048	0,010	0,000*	A
9	0,065	0,009	0,000*	B
10	- 0,019	0,008	0,025	
11	- 0,082	0,009	0,000*	B
12	0,000	0,010	0,966	
13	- 0,019	0,011	0,076	
14	0,027	0,010	0,006*	A
15	- 0,050	0,011	0,000*	A
16	0,055	0,009	0,000*	A
17	0,032	0,010	0,002*	A
18	- 0,006	0,010	0,539	
19	- 0,026	0,010	0,007*	A
20	0,034	0,011	0,002*	A

Not: * $p < 0,01$ $\beta < 0,059$ ise A Düzeyi, $0,059 \leq \beta < 0,088$ ise B Düzeyi, $\beta \geq 0,088$ ise C Düzeyi

Tablo 17’de verilen β katsayısı değerleri Roussos ve Stout (1996)’nın yapmış olduğu sınıflama kapsamında incelenmiş olup maddelerin DMF düzeyleri belirlenmiştir. Buna göre üç madde (5, 9 ve 11) B düzeyinde DMF göstermiştir. Sekiz maddenin (7, 8, 14, 15, 16, 17, 19 ve 20) ise DMF düzeyi A olarak belirlenmiş olup ihmal edilebilir seviyededir. B düzeyinde DMF gösterdiği belirlenen maddelerin avantaj sağladığı grup bilgisi Tablo 18’de verilmiştir.

Tablo 18

Cinsiyete Göre SIBTEST Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi

Madde No	β	DMF Düzeyi	Avantaj Sağladığı Grup
5	- 0,065	B	Odak grup (Erkek)
9	0,065	B	Referans grup (Kadın)
11	- 0,082	B	Odak grup (Erkek)

Not: $\beta < 0,00$ ise odak grup lehine, $\beta > 0,00$ ise referans grup lehine

DMF düzeyi B ve üstü olan maddelerin hangi gruba avantaj sağladığını belirleyebilmek için β katsayısının işareti incelenmiştir. β katsayısının pozitif olması maddenin referans grubun lehine, negatif olması ise odak grubun lehine DMF gösterdiğini ifade etmektedir (Abbott, 2007). Elde edilen değerler doğrultusunda 5. ve 11. maddeler erkeklerin lehine, 9. maddenin ise kadınların lehine avantaj sağladığı tespit edilmiştir.

Cinsiyete göre DMF’li maddeler için yanlılığa ilişkin uzman görüşleri.

Cinsiyet değişkenine göre yapılan analizlerle B düzeyinde DMF tespit edilen maddelerden 5 numaralı maddenin, erkeklere avantaj sağladığı tespit edilmiştir. Bu tespit sadece SIBTEST yöntemi ile elde edilmiştir. Hangi maddelerin Uzman Görüşü Anket Formunda (EK-B) yer alması gerektiğine karar verirken alan yazından elde edilen bilgiler doğrultusunda, analiz yapılan üç yöntemden ikisi ile en az B düzeyinde DMF tespit edilen maddelerin eklenmesi kararı alınmıştır. Ancak 5 numaralı maddenin, MH yöntemine göre de $0,856 (1,0 \leq |\Delta\alpha_{MH}| < 1,5$ ise B Düzeyi) gibi yüksek bir değerle DMF düzeyinin A düzeyinde kalması ile maddenin Uzman Görüşü Anket Formuna eklenmesine karar verilmiştir. Maddenin DMF göstermesinin nedeninin uzmanlara göre madde yanlılığı mı yoksa madde etkisi mi olduğuna dair belirttikleri görüşleri, niceliksel olarak Tablo 19’da verilmiştir.

Tablo 19

5 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi

Uzman Görüşü	Madde Yanlılığı	Madde Etkisi	Toplam
Devlet Okulu Öğretmeni	3	5	8
Özel Okul Öğretmeni	1	7	8
Ölçme-Değerlendirme Uzmanı	1	3	4
Toplam	5 (% 25)	15 (% 75)	20

Tablo 19'dan da anlaşıldığı üzere, 5 numaralı madde için uzmanların % 25'i madde yanlılığı yorumu yapmıştır. Bu yüzde de 5 numaralı maddenin bir grubun lehine yanlılık belirttiği yorumunu yapmak için yeterli görülmemiştir. Uzmanların bu madde hakkındaki yorumlarını dayandırdıkları olası yanlılık sebeplerine sonuçlar ve tartışma bölümünde yer verilmiştir.

Cinsiyet değişkenine göre yapılan analizlerle DMF tespit edilen bir diğer madde ise 9 numaralı maddedir. Yapılan analizlerle, üç yöntemden ikisine (MH ve SIBTEST) göre B düzeyinde DMF gösteren bu maddenin kadınlara avantaj sağladığı tespit edilmiştir. Bu maddenin madde yanlılığı mı yoksa madde etkisinden ötürü mü DMF gösterdiğinin tespiti için uzmanların görüş bildirdiği Uzman Görüşü Anket Formu incelenmiştir. 9 numaralı maddeye ait uzman görüşleri niceliksel olarak Tablo 20'de verilmiştir.

Tablo 20

9 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi

Uzman Görüşü	Madde Yanlılığı	Madde Etkisi	Toplam
Devlet Okulu Öğretmeni	5	3	8
Özel Okul Öğretmeni	4	4	8
Ölçme-Değerlendirme Uzmanı	4	-	4
Toplam	13 (% 65)	7 (% 35)	20

Tablo 20'den de anlaşıldığı üzere, 9 numaralı madde için uzmanların % 65'i madde yanlılığı yorumu yapmıştır. Bu yüzde de 9 numaralı maddenin bir grubun lehine yanlılık belirttiği yorumunu yapmak için yeterli görülmüştür. Uzmanların bu madde hakkındaki yorumlarını dayandırdıkları olası yanlılık sebeplerine sonuçlar ve tartışma bölümünde yer verilmiştir.

Cinsiyet deęişkenine göre yapılan analizlerle farklılaşma tespit edilen bir dięer madde ise 11 numaralı maddedir. Yapılan analizlerle, üç yöntemden ikisine (MH ve SIBTEST) göre B düzeyinde DMF gösteren bu maddenin erkeklere avantaj sağladığı tespit edilmiştir. Bu maddenin madde yanlılığı mı yoksa madde etkisinden ötürü mü DMF gösterdiğinin tespiti için uzmanların görüş bildirdiğı Uzman Görüşü Anket Formu incelenmiştir. 11 numaralı maddeye ait uzman görüşleri niceliksel olarak Tablo 21’de verilmiştir.

Tablo 21

11 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi

Uzman Görüşü	Madde Yanlılığı	Madde Etkisi	Toplam
Devlet Okulu Öğretmeni	8	-	8
Özel Okul Öğretmeni	8	-	8
Ölçme-Deęerlendirme Uzmanı	4	-	4
Toplam	20 (% 100)	0 (% 0)	20

Tablo 21’den de anlaşıldığı üzere, 11 numaralı madde için uzmanların % 100’ü madde yanlılığı yorumu yapmıştır. Bu yüzde de 11 numaralı maddenin bir grubun lehine yanlılık belirttiğı yorumunu yapmak için oldukça yeterli görülmüştür. Uzmanların bu madde hakkındaki yorumlarını dayandırdıkları olası yanlılık sebeplerine sonuçlar ve tartışma bölümünde yer verilmiştir.

Son olarak, cinsiyet deęişkenine göre yapılan analizlerle DMF gösterdiği tespit edilen maddelerden 16 numaralı madde, kadınlara avantaj sağlamıştır. Bu tespit sadece MH yöntemi ile elde edilmiştir. Fakat, 16 numaralı madde, SIBTEST yöntemine göre 0,055 ($0,059 \leq \beta < 0,088$ ise B Düzeyi) gibi yüksek bir deęerle A düzeyinde DMF göstermesinden ötürü Uzman Görüşü Anket Formuna eklenmiştir. Bu maddeye ait uzman görüşleri, niceliksel olarak Tablo 22’de verilmiştir.

Tablo 22

16 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi

Uzman Görüşü	Madde Yanlılığı	Madde Etkisi	Toplam
Devlet Okulu Öğretmeni	3	5	8
Özel Okul Öğretmeni	1	7	8
Ölçme-Deęerlendirme Uzmanı	2	2	4
Toplam	6 (% 30)	14 (% 70)	20

Tablo 22'den de anlaşıldığı üzere, 16 numaralı madde için uzmanların % 30'u madde yanlılığı yorumu yapmıştır. Bu yüzde de 16 numaralı maddenin bir grubun lehine yanlılık belirttiği yorumunu yapmak için yeterli görülmemiştir. Uzmanların bu madde hakkındaki yorumlarını dayandırdıkları olası yanlılık sebeplerine sonuçlar ve tartışma bölümünde yer verilmiştir.

Alt Problem 2' ye İlişkin Bulgular ve Yorumlar

Alt Problem 2. TEOG 2016 Nisan (II. Dönem) sınavının İngilizce alt testine ait maddeleri okul türüne göre yanlılık göstermekte midir?

İngilizce alt testi maddelerinin okul türü değişkenine göre DMF değerleri MH, LR ve SIBTEST yöntemleri kullanılarak ayrı ayrı hesaplandığından her yöntem analiz bulguları ayrı verilmiştir.

Okul türüne göre MH yöntemi analiz bulguları. İngilizce alt testine ait maddelerin okul türü değişkenine göre DMF değerlerinin MH tekniğine göre hesaplanması için EZDIF programından faydalanılmıştır. Analiz bulguları Tablo 23'te verilmiştir. Okul türüne göre İngilizce alt testi maddelerinin MH analizi için yapılan EZDIF program çıktısı EK-G'de verilmiştir.

Tablo 23

İngilizce Alt Testi Maddelerinin Okul Türüne Göre MH Analiz Sonuçları

Madde No	χ^2	p	α	ΔMH	SE ΔMH	DMF Düzeyi
1	27,259	0,000*	1,380	- 0,757	0,146	A
2	0,984	0,321	1,073	- 0,166	0,161	
3	12,724	0,000*	1,260	- 0,542	0,151	A
4	32,476	0,000*	0,608	0,168	0,204	B
5	16,107	0,000*	0,760	0,646	0,160	A
6	1,984	0,159	0,914	0,212	0,147	
7	6,769	0,009*	0,827	0,446	0,168	A
8	4,238	0,040	1,187	- 0,402	0,193	
9	7,352	0,007*	0,807	0,504	0,186	A
10	10,068	0,002*	0,692	0,864	0,266	A
11	16,780	0,000*	0,689	0,875	0,211	A
12	7,555	0,006*	0,813	0,486	0,174	A

Tablo 23

İngilizce Alt Testi Maddelerinin Okul Türüne Göre MH Analiz Sonuçları

Madde No	χ^2	p	α	ΔMH	SE ΔMH	DMF Düzeyi
13	9,522	0,002*	0,807	0,503	0,162	A
14	10,489	0,001*	0,763	0,636	0,194	A
15	1,271	0,260	1,078	- 0,176	0,152	
16	6,329	0,012	0,782	0,578	0,227	
17	1,316	0,251	0,916	0,206	0,174	
18	0,109	0,742	1,029	- 0,067	0,182	
19	3,146	0,076	0,876	0,311	0,171	
20	3,580	0,058	0,880	0,301	0,157	

Not: * $p < 0,01$ $|\Delta \alpha MH| < 1,0$ ise A Düzeyi, $1,0 \leq |\Delta \alpha MH| < 1,5$ ise B Düzeyi, $|\Delta \alpha MH| \geq 1,5$ ise C Düzeyi

MH analizine ait bulgular öncelikle 0,01 anlamlılık düzeyinde incelenerek DMF gösteren maddeler belirlenmiştir. Daha sonra, Tablo 23'de verilen ΔMH katsayı değerleri Zieky (1993)'nin sınıflama ölçütü kapsamında incelenmiş olup test maddelerinin DMF düzeyi tespit edilmiştir. Buna göre sadece 4. maddede B düzeyinde DMF belirlenmiştir. On maddeye (1,3,5,7,9,10,11,12,13 ve 14) ait DMF düzeyi ise A olarak belirlenmiş olup ihmal edilebilir seviyededir. B düzeyinde DMF gösterdiği belirlenen maddelerin avantaj sağladığı grup bilgisi Tablo 24'de verilmiştir.

Tablo 24

Okul Türüne Göre MH Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi

Madde No	ΔMH	DMF Düzeyi	Avantaj Sağladığı Grup
4	1,168	B	Odak grup (Özel okul)

Not: $\Delta MH < 0,00$ ise referans grup lehine, $\Delta MH > 0,00$ ise odak grup lehine

Maddelerin hangi gruba avantaj sağladığı bilgisine ulaşmak için ΔMH katsayısının işareti incelenmiştir. ΔMH katsayısının pozitif olması maddenin odak grubun lehine, negatif olması ise referans grubun lehine DMF gösterdiğini ifade etmektedir. Bu sınırlar kapsamında 4. maddenin özel okul öğrencilerine avantaj sağladığı görülmüştür.

Okul türüne göre LR yöntemi analiz bulguları. İngilizce alt testine ait maddelerin okul türü değişkenine göre DMF özelliğinin incelenmesi için LR analizi yapılmıştır. Okul türüne göre İngilizce alt testi maddelerinin LR analizi için yapılan EZDIF program çıktısı EK-Ğ'de verilmiştir. Program çıktısından elde edilen grup ve grup*toplam puan değişkenlerine ait β değerleri ile β değerine ait anlamlılık değerleri Tablo 25'te verilmiştir.

Tablo 25

İngilizce Alt Testi Maddelerinin Okul Türüne Göre LR Analizi β Estimate Değerleri

Madde No	Grup		Grup*Toplam Puan		DMF Biçimi
	β	p	β	p	
1	1,974711	0,000000*	-0,118009	0,000000*	TBO
2	3,170869	0,000000*	-0,229877	0,000000*	TBO
3	3,713378	0,000000*	-0,245842	0,000000*	TBO
4	-0,245178	0,295573	-0,019530	0,339990	
5	1,326985	0,000000*	-0,125365	0,000000*	TBO
6	3,168307	0,000000*	-0,223913	0,000000*	TBO
7	1,719287	0,000000*	-0,157247	0,000000*	TBO
8	0,695773	0,003078*	-0,044573	0,028732	TB
9	-0,450443	0,026139	0,029294	0,052502	
10	-0,337618	0,273916	0,013668	0,675257	
11	-0,171349	0,496618	-0,012315	0,592252	
12	1,631610	0,000000*	-0,149045	0,000000*	TBO
13	0,745158	0,000587*	-0,076607	0,000002*	TBO
14	0,702152	0,003108*	-0,088485	0,000018*	TBO
15	3,421875	0,000000*	-0,232933	0,000000*	TBO
16	-0,375914	0,154814	0,024689	0,331141	
17	1,112306	0,000001*	-0,099069	0,000000*	TBO
18	2,249403	0,000000*	-0,176439	0,000000*	TBO
19	3,670862	0,000000*	-0,272594	0,000000*	TBO
20	1,328805	0,000000*	-0,113099	0,000000*	TBO

Not: * $p < 0,01$

LR analizinde, grup ve grup*toplam puan etkileşimine ait β estimate değerlerinin 0,01 düzeyinde anlamlılığı incelenmiştir. DMF belirlenen maddelerin on dört tanesinin TBO ve bir tanesinin ise TB biçiminde DMF gösterdiği tespit edilmiştir.

DMF'nin etki büyüklüğü hakkında bilgi sağlamak için SPSS programında Zumbo (1999)'nun syntaxının (EK-E) kullanılarak elde edilen 2 serbestlik dereceli χ^2 değerine karşılık gelen Nagelkerge (R^2) değerleri yardımıyla ΔR^2 hesaplanmıştır. ΔR^2 değeri 3. Modele ait R^2 ile 1. Modele ait R^2 değerlerinin farkı (ΔR^2) olarak hesaplanır ve DMF'nin etki büyüklüğünü belirtir (Swaminathan ve Rogers, 1990). Okul türü değişkenine göre yapılan analiz için, İngilizce alt testinin maddelerine ait χ^2 değerinin anlamlılığı ve ΔR^2 hesaplamaları Tablo 26'da verilmiştir.

Tablo 26

İngilizce Alt Testi Maddelerinin Okul Türüne Göre LR Analiz Sonuçları

Madde No	R ² değerleri						DMF Düzeyi
	χ^2	p	Model-1	Model-2	Model-3	$\Delta R^2 (R_3^2 - R_1^2)$	
1	90,857	0,000*	0,397	0,399	0,407	0,010	A
2	194,931	0,000*	0,524	0,524	0,542	0,018	A
3	283,906	0,000*	0,418	0,419	0,447	0,029	A
5	81,533	0,000*	0,528	0,530	0,536	0,008	A
6	256,98	0,000*	0,440	0,442	0,465	0,025	A
7	96,448	0,000*	0,501	0,502	0,511	0,010	A
8	10,733	0,005*	0,474	0,474	0,475	0,001	A
12	72,770	0,000*	0,573	0,574	0,580	0,007	A
13	34,023	0,000*	0,510	0,511	0,513	0,003	A
14	29,030	0,000*	0,471	0,472	0,475	0,004	A
15	253,317	0,000*	0,448	0,448	0,472	0,024	A
17	32,371	0,000*	0,507	0,507	0,510	0,003	A
18	79,741	0,000*	0,618	0,618	0,625	0,007	A
19	229,742	0,000*	0,602	0,603	0,621	0,019	A
20	62,006	0,000*	0,500	0,500	0,506	0,006	A

Not: * $p < 0,01$ $\Delta R^2 < 0,035$ ise A Düzeyi, $0,035 \leq \Delta R^2 < 0,070$ ise B Düzeyi, $\Delta R^2 \geq 0,070$ ise C Düzeyi

DMF gösterdiği belirlenen maddelerin etki büyüklüğünün belirlenmesi amacıyla hesaplanan ΔR^2 değerleri, Jodoin ve Gierl (2001)'nin oluşturdukları sınıflama ile karşılaştırılmış ve DMF tespit edilen tüm maddelerin (15 madde) ihmal edilebilir düzeyde (A düzeyi) olduğu bulgusuna ulaşılmıştır.

Okul türüne göre SIBTEST yöntemi analiz bulguları. İngilizce alt testine ait maddelerin okul türü değişkenine göre DMF özelliğinin incelenmesi amacıyla

SIBTEST analizi yapılmıştır. Okul türüne göre İngilizce alt testi maddelerinin SIBTEST analizi için yapılan SIBTEST program çıktısı EK-H'de verilmiştir.

Tablo 27

İngilizce Alt Testi Maddelerinin Okul Türüne Göre SIBTEST Analiz Sonuçları

Madde No	p	β	SE β	DMF Düzeyi
1	0,000*	- 0,057	0,015	A
2	0,021	- 0,031	0,014	
3	0,004*	- 0,038	0,013	A
4	0,000*	0,067	0,013	B
5	0,000*	0,053	0,015	A
6	0,582	0,008	0,015	
7	0,517	0,009	0,015	
8	0,347	- 0,013	0,014	
9	0,000*	0,048	0,012	A
10	0,876	0,002	0,013	
11	0,003*	0,040	0,013	A
12	0,184	0,019	0,014	
13	0,001*	0,051	0,015	A
14	0,015	0,033	0,014	
15	0,279	- 0,016	0,015	
16	0,301	0,013	0,013	
17	0,101	- 0,022	0,014	
18	0,058	- 0,025	0,013	
19	0,862	- 0,002	0,014	
20	0,260	0,018	0,016	

Not: * $p < 0,01$ $\beta < 0,059$ ise A Düzeyi, $0,059 \leq \beta < 0,088$ ise B Düzeyi, $\beta \geq 0,088$ ise C Düzeyi

Tablo 27'de verilen β katsayısı değerleri Roussos ve Stout (1996)'nın yapmış olduğu sınıflama kapsamında incelenmiş olup maddelerin DMF düzeyleri belirlenmiştir. Buna göre sadece 4. madde B düzeyinde DMF göstermiştir. Altı maddeye (1, 3, 5, 9, 11 ve 13) ait DMF düzeyi ise A olarak belirlenmiş olup ihmal edilebilir seviyededir. DMF gösterdiği belirlenen maddelerin avantaj sağladığı grup bilgisi Tablo 28'de verilmiştir.

Tablo 28

Okul Türüne Göre SIBTEST Yöntemiyle DMF Belirlenen Maddelerin Avantaj Sağladığı Grup Bilgisi

Madde No	β	DMF Düzeyi	Avantaj Sağladığı Grup
4	0,067	B	Referans grup (Özel okul)

Not: $\beta < 0,00$ ise odak grup lehine, $\beta > 0,00$ ise referans grup lehine

DMF düzeyi B ve üstü olan maddelerin hangi gruba avantaj sağladığı bilgisine ulaşmak için β katsayısının işareti incelenmiştir. β katsayısının pozitif olması maddenin referans grubun lehine, negatif olması ise odak grubun lehine DMF gösterdiğini ifade etmektedir. Elde edilen değerler doğrultusunda 4. maddenin özel okul öğrencilere avantaj sağladığı görülmüştür.

Okul türüne göre DMF'li maddeler için yanlılığa ilişkin uzman görüşleri. Okul türü değişkenine göre yapılan analizlerle sadece 4 numaralı maddenin özel okul öğrencilerine avantaj sağlayacak şekilde DMF gösterdiği tespit edilmişti. Analizler; MH, LR ve SIBTEST olmak üzere üç yöntem ile gerçekleştirilmiştir. Maddenin bu üç yöntemden ikisine (MH ve SIBTEST) göre B düzeyinde DMF gösterdiği tespit edilmiştir. Maddenin DMF göstermesinin nedeninin uzmanlara göre madde yanlılığı mı yoksa madde etkisi mi olduğuna dair belirttikleri görüşleri, niceliksel olarak Tablo 29'da verilmiştir.

Tablo 29

4 Numaralı Maddeye Ait Uzman Görüşlerinin Niceliksel İfadesi

Uzman Görüşü	Madde Yanlılığı	Madde Etkisi	Toplam
Devlet Okulu Öğretmeni	7	1	8
Özel Okul Öğretmeni	7	1	8
Ölçme-Değerlendirme Uzmanı	4	-	4
Toplam	18 (%90)	2 (%10)	20

Tablo 29'dan da anlaşıldığı üzere, 4 numaralı madde için uzmanların % 90'ı madde yanlılığı yorumu yapmıştır. Bu yüzde de 4 numaralı maddenin bir grubun lehine yanlılık belirttiği yorumunu yapmak için yeterli görülmüştür. Uzmanların bu madde hakkındaki yorumlarını dayandırdıkları olası yanlılık sebeplerine sonuçlar ve tartışma bölümünde yer verilmiştir.

Alt Problem 3' e İlişkin Bulgular ve Yorumlar

Alt Problem 3. İngilizce alt testi maddelerinin MH, LR ve SIBTEST analiz yöntemleriyle elde edilen sonuçları uyumlu mudur?

İngilizce alt testi maddelerinin DMF analizlerinin yapılması amacıyla alan yazında yaygın kullanılan MH, LR ve SIBTEST yöntemlerinden yararlanılmıştır. Bu yöntemlerle elde edilen sonuçlar hem belirlenen DMF'li madde sayısı bakımından hem de DMF düzeyi bakımından karşılaştırılmıştır. DMF'li madde sayısı bakımından ortak maddelerin karşılaştırılması yapılmıştır. Yöntemlere göre belirlenen DMF düzeylerinin uyumu ise IBM SPSS Statistic (versiyon 22) programı ile spearman korelasyonu hesaplanarak tespit edilmiştir.

Çalışmada cinsiyet ve okul türü değişkenlerine göre kullanılan örneklemelerin farklı olması nedeniyle oluşturulan alt gruplara göre ayrı alt problemler yazılmış olup analizler de ayrı ayrı yapılmıştır.

Cinsiyete göre MH, LR ve SIBTEST analizi bulgularının uyumu. İngilizce alt testi maddelerinin cinsiyete göre MH, LR ve SIBTEST analizleri ile elde edilen sonuçlar, DMF belirlenen toplam madde sayısı ve bu maddelerin DMF düzeyleri bakımından karşılaştırılmış olup Tablo 30'da verilmiştir.

Tablo 30

Cinsiyete Göre MH, LR ve SIBTEST Yöntemlerine Ait Analiz Bulgularının Karşılaştırılması

DMF Belirleme Yöntemi	DMF Düzeyi			DMF Belirlenen Toplam Madde Sayısı
	A	B	C	
MH	5,6,7,8,14,15,17,20	9,11,16	-	11
LR	1,2,5,6,8,9,11,12,14,15,16,18,19	-	-	13
SIBTEST	7,8,14,15,16,17,19,20	5,9,11	-	11

Tablo 30'da yer verildiği üzere İngilizce alt testi maddelerinden MH yöntemiyle belirlenen DMF'li madde miktarı, sekiz adet A düzeyi ve üç adet B düzeyi olmak üzere toplam on bir adettir. LR yöntemiyle belirlenen DMF'li madde miktarı, ise hepsi A düzeyi olmak üzere on üç adettir. SIBTEST yöntemiyle belirlenen DMF'li madde miktarı da sekiz adet A düzeyi ve üç adet B düzeyi olmak üzere toplam on bir adettir. Her üç yöntem de C düzeyinde DMF'li madde tespit edilmemiştir.

MH yöntemiyle DMF göstermeyen ancak LR yöntemiyle DMF gösteren beş madde (1, 2, 12, 18 ve 19) bulunmuştur ve bu maddelerin hepsi de A düzeyindedir. LR yöntemiyle DMF göstermeyen ancak MH yöntemiyle DMF gösteren üç madde (7, 17 ve 20) bulunmuştur ve bu maddelerin de hepsi A düzeyindedir. Ayrıca her iki yöntemin B düzeyinde belirlediği ortak madde de bulunmamaktadır. 20 maddelik İngilizce alt testinde sekiz (5, 6, 8, 9, 11, 14, 15 ve 16) maddenin her iki yöntemle de ortak olarak DMF göstermesi nedeniyle iki yöntemin, belirlediği DMF'li madde sayısı bakımından uyumunun orta düzeyde olduğu söylenebilmektedir.

LR yöntemiyle DMF göstermeyen ancak SIBTEST yöntemiyle DMF gösteren üç madde (7, 17 ve 20) bulunmuştur ve bu maddelerin hepsi de A düzeyindedir. SIBTEST yöntemiyle DMF göstermeyen ancak LR yöntemiyle DMF gösteren beş madde (1, 2, 6, 12 ve 18) bulunmuştur ve bu maddelerin de hepsi A düzeyindedir. Ayrıca her iki yöntemin B düzeyinde belirlediği ortak madde de bulunmamaktadır. 20 maddelik İngilizce alt testinde sekiz (5, 8, 9, 11, 14, 15, 16 ve 19) maddenin her iki yöntemle de ortak DMF özelliği göstermesi nedeniyle iki yöntemin, belirlediği DMF'li madde sayısı bakımından uyumunun orta düzeyde olduğu söylenebilmektedir.

SIBTEST yöntemiyle de DMF göstermeyen ancak MH yöntemiyle DMF gösteren sadece 6 numaralı madde vardır ve bu madde A düzeyindedir. MH yöntemiyle DMF göstermeyen ancak SIBTEST yöntemiyle DMF gösteren tek madde ise 19 numaralı maddedir ve bu madde A düzeyindedir. Her iki yöntemle DMF belirlenen ortak maddeler (5, 7, 8, 9, 11, 14, 15, 16, 17 ve 20) ise on adettir. Ortak maddelerden 9 ve 11 numaralı maddeler her iki yöntemle göre de B düzeyindedir. 5 numaralı madde SIBTEST yöntemine göre B düzeyinde iken MH yöntemine göre A düzeyindedir. Benzer şekilde 16 numaralı madde de MH yöntemine göre B düzeyinde iken SIBTEST yöntemine göre A düzeyindedir. Diğer maddeler ise her iki yöntemle göre de A düzeyindedir. Her iki yöntemle göre de ortak olarak DMF özelliği gösterdiği tespit edilen madde sayısının on olması, iki yöntemin belirlediği DMF'li madde sayısı bakımından uyumunun orta düzeyde olduğunu göstermektedir.

Yöntemlerin DMF belirlediği maddelerin etki büyüklüğü bakımından uyumu ise 0,01 anlamlılık düzeyinde spearman-brown korelasyonu hesaplanarak tespit edilmiştir. Spearman rho katsayı değerleri Tablo 31'de verilmiştir.

Tablo 31

Cinsiyete Göre MH, LR ve SIBTEST Yöntemlerinin Spearman Rho Katsayı Değerleri

Yöntem	MH	SIBTEST	LR
MH	1,00		
SIBTEST	0,82	1,00	
LR	0,26	0,26	1,00

Yöntemler arası uyum ikili olarak değerlendirildiğinde $r_s = 0,26$ değerinde olan MH ve LR yöntemlerin arasındaki uyumun zayıf ($0,10 > r_s > 0,29$) olduğu yorumu yapılabilmektedir (Pallant, 2007). Benzer şekilde LR ve SIBTEST yöntemleri arasındaki uyumun da $r_s = 0,26$ değer olarak hesaplanmasıyla bu iki yöntem arasındaki uyumun zayıf olduğu yorumu yapılabilmektedir. MH ve SIBTEST yöntemlerinin ise hesaplanan $r_s = 0,82$ değeri ile aralarındaki uyumun yüksek ($0,50 > r_s > 1,0$) olduğu yorumu yapılabilmektedir (Pallant, 2007). Özetle cinsiyete göre yapılan analizlerde, parametrik olmayan yöntemlerden MH ve SIBTEST yöntemleri benzer bulgular verirken parametrik olan LR yönteminin bu yöntemlerle benzer bulgulara ulaşmadığı yorumu yapılabilmektedir.

Okul türüne göre MH, LR ve SIBTEST analiz bulgularının uyumu. İngilizce alt testi maddelerinin okul türüne göre MH, LR ve SIBTEST analizleri ile elde edilen sonuçlar, DMF düzeyi ve DMF belirlenen toplam madde sayısı olarak karşılaştırılmış olup Tablo 32’de verilmiştir.

Tablo 32

Okul Türüne Göre MH, LR ve SIBTEST Yöntemlerine Ait Analiz Bulgularının Karşılaştırılması

DMF Belirleme Yöntemi	DMF Düzeyi			DMF Belirlenen Toplam Madde Sayısı
	A	B	C	
MH	1,3,5,7,9,10,11,12,13,14	4	-	11
LR	1,2,3,5,6,7,8,12,13,14,15,17,18,19,20	-	-	15
SIBTEST	1,3,5,9,11,13	4	-	7

Tablo 32’de yer verildiği üzere İngilizce alt testi maddelerinden okul türüne göre MH yöntemiyle belirlenen DMF’li madde miktarı, on adet A düzeyi ve bir adet B düzeyi olmak üzere toplam on bir adettir. LR yöntemiyle ise belirlenen DMF’li

madde miktarı, hepsi A düzeyi olmak üzere on beş adettir. SIBTEST yöntemiyle belirlenen DMF'li madde miktarı da altı adet A düzeyi ve bir adet B düzeyi olmak üzere toplam yedi adettir. Her üç yöntem de C düzeyinde DMF'li madde tespit edilmemiştir.

MH yöntemiyle DMF göstermeyen ancak LR yöntemiyle DMF gösteren sekiz madde (2, 6, 8, 15, 17, 18, 19 ve 20) bulunmuştur ve bu maddelerin hepsi de A düzeyindedir. LR yöntemiyle DMF göstermeyen ancak MH yöntemiyle DMF gösteren dört madde (4, 9, 10 ve 11) bulunmuştur ve bu maddelerden 4 numaralı olan B düzeyinde iken diğerleri A düzeyindedir. Ayrıca her iki yöntemin B düzeyinde DMF belirlediği ortak madde de bulunmamaktadır. 20 maddelik İngilizce alt testinde yedi (1, 3, 5, 7, 12, 13 ve 14) maddenin her iki yöntemle de ortak olarak DMF göstermesi nedeniyle iki yöntemin belirlediği DMF'li madde sayısı bakımından uyumunun zayıf olduğu yorumu yapılmıştır.

LR yöntemiyle DMF göstermeyen ancak SIBTEST yöntemiyle DMF gösteren üç madde (4, 9 ve 11) bulunmuştur ve bu maddelerden 4 numaralı olan B düzeyinde iken diğer maddeler A düzeyindedir. SIBTEST yöntemiyle DMF göstermeyen ancak LR yöntemiyle DMF gösteren on bir madde (2, 6, 7, 8, 12, 14, 15, 17, 18, 19 ve 20) bulunmuştur ve bu maddelerin hepsi de A düzeyindedir. Her iki yöntemle DMF belirlenen ortak maddeler (1, 3, 5 ve 13) ise dört adettir. Ayrıca her iki yöntemin B düzeyinde DMF belirlediği ortak madde de bulunmamaktadır. Bu nedenle, bu iki yöntemin belirlediği DMF'li madde sayısı bakımından uyumunun çok zayıf olduğu görülmektedir.

SIBTEST yöntemiyle DMF göstermeyen ancak MH yöntemiyle DMF gösteren dört madde (7, 10, 12 ve 14) bulunmuştur ve bu maddeler A düzeyindedir. MH yöntemiyle DMF göstermeyen ancak SIBTEST yöntemiyle DMF gösteren hiç madde bulunmamıştır. Her iki yöntemle DMF belirlenen ortak maddeler (1, 3, 4, 5, 9, 11 ve 13) ise yedi adettir. Ortak maddelerden 4 numaralı madde, her iki yöntemle göre de B düzeyinde DMF gösterirken diğer maddeler A düzeyindedir. Bu iki yöntemin belirlediği DMF'li madde sayısı bakımından uyumunun zayıf olduğu görülmektedir.

Yöntemlerin belirlediği DMF'li maddelerin etki büyüklüğü bakımından uyumu 0,01 anlamlılık düzeyinde spearman-brown korelasyonu hesaplanarak tespit edilmiştir. Spearman rho katsayı değerleri Tablo 33'de verilmiştir.

Tablo 33

Okul Türüne Göre MH, LR ve SIBTEST Yöntemlerinin Spearman Rho Katsayı Değerleri

Yöntem	MH	SIBTEST	LR
MH	1,00		
SIBTEST	0,71	1,00	
LR	0,36	0,35	1,00

Yöntemler arası uyum ikili olarak değerlendirildiğinde 0,01 anlamlılık düzeyinde $r_s = 0,36$ ($p = 0,12$) değerinin anlamlı olmaması nedeniyle MH ve LR yöntemlerinin uyumlu sonuçlar vermediği yorumu yapılmıştır. Benzer şekilde LR ve SIBTEST yöntemlerinin de $r_s = 0,35$ ($p = 0,13$) değerine ulaşması nedeniyle bu iki yöntemin uyumlu sonuçlar vermediği yorumu yapılmıştır. MH ve SIBTEST yöntemleri için ise hesaplamalarda $r_s = 0,71$ ($p = 0,00$) değerine ulaşılması ile bu iki yöntem arasındaki uyumun yüksek ($0,50 > r_s > 1,0$) olduğu yorumu yapılmıştır (Pallant, 2007). Özetle okul türüne göre yapılan analizde, parametrik olmayan yöntemlerden MH ve SIBTEST benzer bulgular verirken parametrik olan LR yönteminin bu yöntemlerle benzer bulgulara ulaşmadığı yorumu yapılabilmektedir.

Bölüm 5

Sonuç, Tartışma ve Öneriler

Araştırmanın bu bölümünde de elde edinilen bulgular ve alan yazındaki bilgiler ışığında maddeler hakkındaki sonuçlara genel olarak yer verilmiştir. Ayrıca alt problemlere ait tartışmalara ise ayrı ayrı yer verilmiştir. Ek olarak bu alanda yapılabilecek araştırmalara yönelik önerilere de bu bölümde yer verilmiştir.

Sonuç

TEOG 2015 – 2016 Eğitim – Öğretim Yılı II. Dönem İngilizce alt testine ait maddelerin, öğrencilerin cinsiyet ve öğrenim gördüğü okul türüne (özel-devlet) göre DMF gösterip göstermediği incelenmiş olup elde edinilen bulgular doğrultusunda uzmanlara başvurularak bu maddelerin yanlışlık belirtip belirtmediği tespit edilmeye çalışılmıştır.

İngilizce alt testine ait maddeler cinsiyet değişkenine göre MH, LR ve SIBTEST yöntemleri ile analiz edildiğinde 5, 9, 11 ve 16 numaralı maddelerin B düzeyinde DMF gösterdiği tespit edilmiştir. Bu maddeler için uzman görüşleri ve alan yazındaki incelemeler göz önünde bulundurulduğunda 5 ve 16 numaralı maddelerin DMF göstermesinin madde etkisinden kaynaklandığı sonucuna ulaşılmıştır. 9 numaralı maddenin kadınlar lehine ve 11 numaralı maddenin ise erkekler lehine madde yanlışlığı gösterdiği sonucuna ulaşılmıştır.

İngilizce alt testine ait maddeler okul türü değişkenine göre MH, LR ve SIBTEST yöntemleri ile analiz edildiğinde ise sadece 4 numaralı maddenin B düzeyinde DMF gösterdiği tespit edilmiştir. Uzman görüşleri ve alan yazındaki incelemeler göz önünde bulundurulduğunda bu maddenin DMF göstermesinin madde yanlışlığından kaynaklı olduğu sonucuna ulaşılmıştır.

Bu çalışmada kullanılan analiz yöntemlerinin uyumu incelendiğinde ise hesaplanan spearman rho katsayısına göre MH ve SIBTEST yöntemlerinin birbiriyle uyumunun yüksek, LR yöntemi ile uyumunun ise düşük olduğu sonucuna ulaşılmıştır.

Alt Problem 1' e İlişkin Tartışmalar

İngilizce alt testine ait maddeler cinsiyet değişkenine göre MH, LR ve SIBTEST yöntemleri ile analiz edildiğinde 5, 9, 11 ve 16 numaralı maddelerin B düzeyinde DMF gösterdiği tespit edilmiştir.

İngilizce alt testine ait 5 numaralı madde, sadece SIBTEST yöntemine göre B düzeyinde DMF göstermiştir. MH ve LR yöntemine göre de DMF göstermiş olup A düzeyinde kalmıştır. Analiz sonuçlarına göre erkekler lehine DMF gösterdiği belirlenen ve yanlılığı hakkında karar vermek üzere uzman görüşüne başvuru olan bu madde Şekil 9.'da verilmiştir.

1-8. sorularda boş bırakılan yere uygun gelen kelime ya da ifadeyi işaretleyiniz.

5. Candy : - - - - ?
Sally : Chinese. And over 1 billion people speak it in the world.

A) Where would you like to visit
B) What is their official language
C) Where can a tourist stay there
D) What is the name of the country

Şekil 9. TEOG İngilizce alt testine ait 5 numaralı madde.

Bu madde, MEB Talim Terbiye Kurulu Başkanlığı tarafınca belirlenen öğretim programında sekizinci sınıf "The Internet" ünitesi kapsamında işlenmektedir. Maddede geçen "Chinese" kelimesi ise beşinci sınıf müfredatında "Foreign Language" ünitesi kapsamında işlenmektedir.

Uzmanların %75'i bu maddenin erkeklerin lehine DMF göstermesini madde etkisi ile açıklamıştır. Uzmanlar, bu maddeyi doğru cevaplandırmak için "Chinese" diye bir dil olduğunu bilmenin yeterli olduğunu ifade ederek maddenin DMF göstermesinin nedenini, yanlış yapanların bilgi ve beceri eksikliğinden kaynaklandığı şeklinde ifade etmiştir. Maddedeki farklılaşmayı madde yanlılığı ile ifade eden uzmanlar (uzmanların %25'i) ise düşük bir ihtimalle internet ve bilgisayar oyunlarına ilgisi olan erkek öğrencilerin oyunlarda "Chinese" kelimesi ile daha sık karşılaşacaklarından ötürü kalıcı öğrenme gerçekleştirebileceğini ifade etmiştir. Ayrıca erkeklerin spora olan ilgileri ve Dünya kupası etkinliklerini takip etmeleri

nedeniyle Dünya ülkelerinin özelliklerine daha aşına olmalarının bu madde için yanlılık oluşturabileceği de uzmanlar tarafından ifade edilmiştir.

Uzman görüşleri ve alan yazındaki çalışmalar dikkate alındığında TEOG İngilizce alt testinde yer alan 5 numaralı maddenin, madde yanlılığı göstermediği sonucuna ulaşılmıştır. Kelecioğlu, Karabay ve Karabay (2014)'ın çalışmalarında da bu çalışmaya benzer şekilde DMF analizleri ile farklılaşma tespit edilen maddelerin uzman görüşlerine göre madde yanlılığı belirtmeyeceğini ifade etmişlerdir. Yine Karakaya ve Kutlu (2012)'nin 2009 SBS Türkçe alt testine ait maddelerle yaptığı çalışmada da uzman görüşü sonrası maddelerde madde etkisinden kaynaklı farklılaşma olabileceği belirtilmiştir.

DMF tespit edilen bir diğer madde ise testte yer alan 9 numaralı maddedir. Bu madde hem MH hem de SIBTEST yöntemine göre B düzeyinde DMF göstermiştir. LR yöntemine göre de DMF göstermiş olup A düzeyinde kalmıştır. Analiz sonuçlarına göre madde, kadınlar lehine DMF göstermiştir. Maddenin yanlılığı hakkında karar vermek üzere uzman görüşüne başvuru bu madde Şekil 10.'da verilmiştir.

9. I. Hello, I'm Sarah.
II. Oh, sorry Sarah. This is Jim's father.
III. Could I ask who is calling?
IV. Hi, Sarah. Could I speak to Jim?

Yukarıda bir telefon konuşması karışık olarak verilmiştir.

Anlamli bir diyalog oluşturmak için numaralanmış cümleler nasıl sıralanmalıdır?

A) I, IV, III, II B) IV, II, III, I
C) I, II, IV, III D) IV, III, I, II

Şekil 10. TEOG İngilizce alt testine ait 9 numaralı madde.

Bu madde, MEB Talim Terbiye Kurulu Başkanlığı tarafınca belirlenen öğretim programında sekinci sınıf "On The Phone" ünitesi kapsamında işlenmektedir. Bu ünitenin kazanımları arasında karışık olarak verilen konuşmaların sıralanması da mevcuttur. Kavramlar incelendiğinde ise öğrencilere farklı gelebilecek karmaşık bir kavrama rastlanmamıştır.

Uzmanların % 65'i bu maddenin kadınlar lehine yanlılık belirttiğini ifade etmiştir. Uzmanlar, maddenin parçalardan anlamlı bütün oluşturma temelli olduğuna dikkat çekerek asıl ölçülenin sözel-dil becerisi olduğunu vurgulamıştır. Bu beceride de kadınların erkeklere göre daha önde olduğu ifade edilmiştir. Alan yazın incelendiğinde de uzmanların görüşünü destekleyen çalışmalar vardır (Azen, Bronner ve Gafni, 2002; Coley, 2001; Lietz, 2006; Lynn ve Mikk, 2009). Amrein ve Berliner (2002), yaptıkları çalışmalarıyla özellikle ilköğretim seviyesinde kadınların erkeklere oranla sözel-dil becerisine bağlı konularda daha başarılı olduğunu ortaya koymuştur. Yılmaz (2009) da çalışmasında benzer bir sonuca ulaşmıştır. Bir başka görüşe göre de kadınların sıralama gerektiren maddelerde daha başarılı olduğu ifade edilerek bu durum, kadınların yapılması gerekenleri daha düzenli bir şekilde adım adım yapmayı alışkanlık haline getirmesi ile açıklanmıştır.


Uzmanların % 35'i ise bu madde için kadınların daha başarılı olmasını, bilgi ve beceri üstünlüğü ile açıklayarak maddenin yapısından kaynaklı herhangi bir gruba yanlılık oluşturabilecek kusur taşımadığını ifade etmiştir.

TEOG İngilizce alt testine ait 9 numaralı madde, kavram bazında incelendiğinde de bilinmesi zor özel bir kavrama rastlanmamıştır. Maddenin iletişim konulu olması nedeniyle uzmanların geneli kadınların daha yüksek sözel dil becerisine sahip olduğunu ve bu becerinin ufak bir fark yaratabileceğini ifade etmiş olsa da maddenin DMF düzeyinin B seviyesinde kalması, testin yapı geçerliğini büyük ölçüde olumsuz yönde etkilemeyeceği yorumuna dayanak sağlayabilmektedir.

TEOG İngilizce alt testine ait 11 numaralı madde ise DMF tespit edilen bir başka maddedir. Bu madde MH ve SIBTEST yöntemine göre B düzeyinde DMF gösterirken LR yöntemine göre de A düzeyinde DMF göstermiştir. Analiz sonuçlarına göre madde, erkekler lehine DMF göstermiştir. Maddenin yanlılığı hakkında karar vermek üzere uzman görüşüne başvurulmuş bu madde Şekil 11.'de verilmiştir.

10 ve 11. sorularda verilen görsele göre boş bırakılan yere uygun gelen seçeneği işaretleyiniz.

11.



Mr. Carter

Mr. Carter - - - - when he was young.

A) enjoyed white-water rafting
B) hated doing bungee jumping
C) tried hang gliding several times
D) disliked performing parkour running

Şekil 11. TEOG İngilizce alt testine ait 11 numaralı madde.

Bu madde, MEB Talim Terbiye Kurulu Başkanlığı tarafınca belirlenen öğretim programında sekizinci sınıf "Adventures" ünitesi kapsamında işlenmektedir. Bu ünite de ekstrem sporlar konu başlığı ile maddede geçen spor türleri anlatılmaktadır. Madde, görsel bir maddedir. MEB tarafından dağıtılan kitaplar incelendiğinde de ekstrem sporların bu maddede olduğu gibi spor türlerinin görseline yer verilmiştir. Maddede görseli verilen spor, "hang-gliding" olarak bilinmektedir ve MEB tarafından uygun bulunan ders kitaplarında resimli olarak yer almaktadır.

Uzmanların % 100'ü yani tamamı, bu maddenin erkeklerin lehine çalışmasını madde yanlılığı ile açıklamıştır. Erkeklerin ekstrem sporlarla daha ilgili olduğunu savunan uzmanlar, spor isimlerinin erkeklerin aklında daha kalıcı yer ettiklerini ifade etmiştir. Hatta uzmanlardan biri, madde görselinde de sporu yapan kişinin erkek olarak seçilmesinin altını çizerek bu durumu, maddenin açık açık erkeklere avantaj sağladığının ispatı olarak yorumlamıştır. Stricker and Emmerich (1999) çalışmalarında ilgi alanlarının öğrencilerin maddeyi doğru cevaplama konusunda avantaj sağladığını ifade etmiştir. Bir başka uzman da ilgi alanlarının ve kültürlerinin bireylerin başarı düzeyine olumlu etkileri olduğunu ve bu sebepten ötürü spor konulu bu maddenin erkekler tarafından doğru yapılma olasılığının yüksek olduğu belirtilmiştir. Yeteneklerimizin ve algılarımızın bizi güdülediğini ifade eden Morgan (1991) da ilgi alanlarımızın kalıcı öğrenmeye etkisini vurgulayan bu yorumu

desteklemektedir. Alan yazın incelendiğinde de uzmanların görüşünü desteleyecek yönde birçok çalışma vardır (Benbow ve Stanley, 1982; Halpern, 1997; Ingels, 1990; Lohman, 1993). Yıldırım (2015), sekizinci sınıf öğrencilerinin katıldığı Seviye Belirleme Sınavı ile yaptığı çalışmasında erkek öğrencilerin spor konulu maddelerde daha avantajlı olduğunu ifade etmiştir. Educational Testing Service (1998), test geliştiricilerine madde yazımı sürecinde; askeri konular, ayrıntılı spor bilgileri, gibi testin amacı ile ilgisi olmayan gereksiz yere zor kelimelerden kaçınılması konusunda tavsiyede bulunmuştur. Bu konuda özellikle askeri ve spor konularının erkeklere avantaj sağlayabileceğinden olası DMF kaynağı olarak ifade edilmiştir.

TEOG İngilizce alt testine ait 11 numaralı madde, görsel özelliği nedeniyle de erkek öğrencilere avantaj sağlamış olabilir. Alan yazındaki çalışmalar incelendiğinde, görsel/uzamsal beceri gerektiren çoktan seçmeli maddelerde şekil, diyagram vb. verildiğinde erkek öğrencilerin kız öğrencilere göre daha başarılı olduğu bulgusuna rastlanmıştır (bkz. Hafner, Ingels, Schneider ve Stevenson, 1990; Ingels, 1990). Lui ve Wilson (2009)'ın PISA 2000 ve PISA 2003 matematik verileri ile yaptığı çalışmada görsellik içeren maddelerde erkek öğrencilerin dikkat çekecek düzeyde yüksek performans gösterdiğini belirtmiştir. Benzer şekilde Hamilton-Laura ve Snow (1998) da görsel içerikli maddelerin erkekler lehine çalıştığını tespit ettikleri bir çalışmaya sahiptirler. Bu çalışmada ayrıca okul dışında edinilen bilgi ve deneyimin de öğrencilerin kalıcı öğrenmelerine katkı sağladığı ifade edilmiştir.

Uzman görüşleri ve alan yazındaki çalışmalar dikkate alındığında TEOG İngilizce alt testinde yer alan 11 numaralı maddenin madde yanlılığı belirttiği fakat orta düzeyde DMF göstermesi nedeniyle yapı geçerliğine önemli boyutta olumsuz etki bırakmayacağı sonucuna ulaşılmıştır. Bu araştırma, test hazırlayan merciler için spor konulu madde yazarken tüm öğrencilerin daha aşına olduğu bir spor dalı kullanmaları konusunda öneri niteliği taşıyabilir.

TEOG İngilizce alt testinde cinsiyete göre DMF belirten 16 numaralı madde ise sadece MH yöntemine göre B düzeyinde DMF göstermiştir. SIBTEST ve LR yöntemine göre de DMF göstermiş olup A düzeyinde kalmıştır. Analiz sonuçlarına göre kadınlar lehine DMF gösterdiği belirlenen ve yanlılığı hakkında karar vermek üzere uzman görüşüne başvuru alan madde Şekil 12.'de verilmiştir.

15-17. soruları aşağıda verilen metinlere göre cevaplayınız.



Egypt is the centre of many civilizations and there are many historic buildings to visit. But it is really too hot in summers so I prefer going there in mild winter days.

Linda

I really enjoy being outdoors and on the water. That's why with its blue waters, Italy is a unique tourist destination for me. Also, Italy has mild weather the year around so it is really convenient for water sports. I would like to go there and enjoy its crystal blue waters.



Robert



I enjoy exploring tropical islands like Jamaica. Jamaica is well-known for its many traditional dishes and I like trying new meals. With its mild climate the year around, Jamaica is the best choice for my summer holidays.

Peter

16. Robert would like to visit Italy because - - - .

- A) it is usually cold in winters
- B) he wants to do water sports
- C) it has many historic buildings
- D) he would like to try new dishes

Şekil 12. TEOG İngilizce alt testine ait 16 numaralı madde.

Bu madde, MEB Talim Terbiye Kurulu Başkanlığı tarafınca belirlenen öğretim programında "Tourism" ünitesi kapsamında işlenmektedir. Bu ünite de tatil yapılabilecek yabancı ülkeler konu başlığında bu ülkelerin tercih edilme sebepleri; tarihi, ören yerleri, modern binaları, hava şartları, spor aktiviteleri, mimarisi, deniz tatili ve doğası gibi özelliklerinden bahsedilerek işlenmiş olup bu özelliklere ait kavramlara da yer verilmiştir.

Uzmanların %70'i bu maddenin kadınların lehine DMF göstermesini madde etkisi ile açıklamıştır. Maddenin yanlılık taşıyacak şekilde bir gruba yönelmediği ifade edilmiştir. Uzmanlardan biri maddenin birleşik olmasına dikkat çekerek aynı madde köküne bağlı olan madde 15 ve madde 17'nin de incelenmesini talep etmiştir. Eğer o maddeler de DMF gösterme eğilimindedir ise DMF belirtmesinin sebebinin madde kökünden kaynaklandığı ve maddenin yanlılık oluşturduğu yorumunun yapılabileceği savunulmuştur. Bu sebepten ötürü madde 15'e ait DMF analizleri tekrar incelendiğinde her üç yönteme göre de DMF düzeyinin A düzeyinde kaldığı görülmüştür. Madde 17'e ait analizler incelendiğinde ise, MH ve SIBTEST yöntemine göre A düzeyinde DMF tespit edilmişken LR yöntemine göre DMF tespit

edilmemiştir. Bu incelemeler sonrası madde 15 ve 17'nin kayda değer bir farklılaşma göstermediği yorumu yapılabilmektedir.

Maddenin DMF göstermesini madde yanlılığı olarak ifade eden uzmanların %30'u ise bu durumun dil becerisinden kaynaklanabileceğini söylemiştir. Madde kökünün uzun olması sözel-dil becerisi ile ilişkilendirilerek bu beceride daha başarılı olan gruba yani kadınlara, maddenin avantaj sağlamış olabileceği şeklinde yorum yapılmıştır. Bu yorumdan yola çıkılarak uzmanlardan biri, testteki madde kökü uzun olan diğer maddelerin de incelenmesi gerektiğini ifade etmiştir. Eğer o maddeler de DMF gösterme eğilimindedir ise farklılaşmanın sebebi madde kökünün uzunluğu olarak ele alınıp madde yanlılığı yorumunun daha kolay yapılabileceği ifade edilmiştir. Bu nedenle 15 ve 17 numaralı maddelere ek olarak 18, 19 ve 20 numaralı maddeler de tekrar incelenmiştir. Madde 18; MH ve SIBTEST yöntemine göre A düzeyinde DMF gösterirken LR yöntemine göre DMF göstermemektedir. Madde 19; MH ve LR yöntemine göre DMF gösterirken SIBTEST yöntemine göre DMF göstermemiştir. Madde 20 ise MH ve SIBTEST yöntemine göre A düzeyinde DMF gösterirken LR yöntemine göre DMF göstermemektedir. Bu nedenle testteki maddelerin madde yanlılığı oluşturacak kadar uzun madde köküne sahip olmadığı yorumu yapılmıştır. Ayrıca; 15, 16, 17, 18, 19 ve 20 numaralı maddelerin madde güçlükleri ve ayırıcılıkları (bkz. Sayfa 50) da incelendiğinde bu değerlerin birbirine çok yakın ve ortalama değerde olduğu görülmüştür.

Uzman görüşleri ve alan yazındaki çalışmalar dikkate alındığında TEOG İngilizce alt testinde yer alan 16 numaralı maddenin madde yanlılığı oluşturmadığı sonucuna ulaşılmıştır. Satıcı ve Özer-Özkan (2017) da TEOG 2014 (Kasım) sınavının maddeleri ile yaptıkları çalışmalarında bu maddede olduğu gibi DMF göstermesine rağmen cinsiyet bakımından yanlılık oluşturmayan maddeler tespit etmişlerdir.

Alt Problem 2' e İlişkin Tartışmalar

İngilizce alt testi maddeleri için okul türü değişkenine göre DMF analizi MH, LR ve SIBTEST yöntemleri kullanılarak yapılmıştır. Analiz sonucu, sadece 4 numaralı maddenin MH ve SIBTEST yöntemine göre B düzeyinde DMF tespit edilmiştir. Bu maddenin her iki yönetime göre de özel okul öğrencilerine avantaj

sağladığı sonucuna ulaşılmıştır. TEOG İngilizce alt testine ait 4 numaralı madde, Şekil 13.'te verilmiştir.

1-8. sorularda boş bırakılan yere uygun gelen kelime ya da ifadeyi işaretleyiniz.

4. Pelin : - - - - ?
Sue : Never. I prefer face-to-face interaction.

A) When do you go out with your parents
B) What are its technical specifications
C) What kind of films do you watch
D) How often do you chat online

Şekil 13. TEOG İngilizce alt testine ait 4 numaralı madde.

Uzmanların %90'ı bu maddenin özel okul öğrencilerine avantaj sağladığını ifade etmiştir. Uzmanlar genel olarak, maddede yer alan “face-to-face” ve “chat online” kavramlarının bilgisayar ve internet erişiminin daha kolay olan bireylere yanlılık sağlayabileceğini savunup özel okul öğrencilerinin bu maddeyi cevaplandırmada daha avantajlı olduğunu ifade etmiştir. Özel okullarda dil laboratuvarlarının bulunması, derslerde işitsel ve görsel araç gereçlerin kullanılması ve her öğrenciye bilgisayar tahsis edilebilir olması nedeniyle maddede geçen kavramların daha kolay öğrenilebileceği ifade edilmiştir. Devlet okullarında ise sınıf kalabalığı nedeniyle uygulamalı öğretim yapılmasının zor olması öğrencilerin bu kavramlara doğrudan maruz kalmalarını engellemektedir. Ayrıca, özel okula giden öğrencilerin evlerinde de bilgisayar ve internet kullanabilecek maddi yeterliliğe sahip olması nedeniyle öğrencilerin bu kavramları günlük hayatlarında da kullanmaları sonucu daha kalıcı öğrenme gerçekleştirebileceği ifade edilmiştir. Evde kullanılabilen bilgisayar gibi materyallerin, öğrencilerin okul öğrenmeleri üzerinde olumlu etkisi olduğunu ifade eden Yang (2003), bu görüşü desteklemektedir. Çalışmasında benzer sonuca ulaşan Bakan-Kalaycıoğlu (2008), incelediği maddede geçen “binici” kavramının İngilizce “rider” kavramından dilimize çevrildiğini belirterek maddenin İngilizce eğitiminin fazla olduğu bireylere (çalışmasında bu grup özel okul öğrencileri olarak belirtilmiş) avantaj sağladığını ifade etmiştir.

Testte yer alan 4 numaralı madde, MEB Talim Terbiye Kurulu Başkanlığı tarafınca belirlenen öğretim programında “On The Phone” ve “The Internet” üniteleri kapsamında işlenmektedir. Bu nedenle uzmanların %10'u bu maddede geçen

kavramların ders kapsamında kullanıldığını ve öğrencilerin öğrenmesi gerektiğini ifade etmiştir. Devlet okulu öğrencilerinin bu maddeyi doğru cevaplandırma oranının fazla olmaması ise bu uzmanlar tarafından öğrencilerin yeterli öğrenme gerçekleştirilmesine bağlanmıştır. Maddenin DMF göstermesini ise madde etkisi ile ifade etmişlerdir.

Uzmanların görüşlerine göre ortak kanı, bu maddenin özel okul öğrencilerine yanlılık göstermesinin sebebi öğrencilerin sadece okullarında aldıkları eğitim değil, sahip oldukları sosyo-ekonomik düzeyin etkisidir. Bu çalışmayla benzer sonuca ulaşan Berberoğlu, Demirtaşlı, Güzel ve diğerleri (2010), öğrencilerin sosyo-ekonomik düzeylerini dikkate alarak yaptıkları çalışmalarında; özel okullarda okuyan öğrencilerin aileleri ile devlet okullarında okuyan öğrencilerin aileleri arasında hem maddi imkân hem eğitime verdikleri önem hem de bilinç olarak oldukça farkın olduğu bu nedenle de özel okullardaki öğrencilerin daha başarılı olabildiklerinin altını çizmişlerdir.

Uzmanların görüşleri doğrultusunda; her iki okul türünün de aynı çerçeve programı kapsamında öğretim sürecini gerçekleştirilmesinin farklılaşma oluşturmada tek başına etkili olmadığı, özel okul öğrencilerinin ve ebeveynlerinin sahip olduğu sosyo-ekonomik düzeyin, bu öğrencilerin daha fazla yaşantı kazanmasına olanak sağladığı ve bu durumun da yaparak-yaşayarak öğrenmeyi önemli kıldığı sonucuna ulaşılmıştır. Yurdugül ve Aşkar (2004a), öğrencilerin özelde edindiği bilgilerin yani tecrübelerin ışığında maddeyi doğru cevaplama olasılıklarının yükselmesini madde etkisi ile ifade etmişlerdir. Böyle bir durumda, maddenin içeriğinden ya da formatından ötürü bir kusur olmayacağını ve madde yanlılığı çerçevesinde değerlendirilmemesi gerektiğini savunmuşlardır.

Sonuç olarak TEOG İngilizce alt testine ait 4 numaralı madde, uzmanların %90'ı tarafından madde yanlılığı oluşturuyor diye ifade edilse de alan yazındaki bilgiler de değerlendirildiğinde, DMF düzeyinin B seviyesinde kalması testin yapı geçerliğini olumsuz etkilemeyeceği şeklinde yorumlanabilir. Bu şekliyle yapılabilecek iyileştirmeler sayesinde testteki yerini koruyabilir.

Alt Problem 3' e İlişkin Tartışmalar

Bu çalışmada analizlerde kullanılan üç yöntemin uyumu, hem cinsiyet hem de okul türüne göre ayrı ayrı bakıldığında benzer sonuca ulaşılmıştır.

Alan yazında, LR yöntemi ile yapılan analizlerde MH ve SIBTEST yöntemine oranla daha fazla DMF tespit edilmesinin beklenildiği ifade edilmiştir (Gierl, Khalig ve Boughton, 1999; Hidalgo ve Lopez-Pina, 2004). Bu çalışmada da 20 maddelik İngilizce alt testi ile cinsiyet değişkenine göre yapılan analizde LR (DMF gösteren madde sayısı:13) yöntemiyle belirlenen DMF'li madde sayısının MH (DMF gösteren madde sayısı:11) ve SIBTEST (DMF gösteren madde sayısı:11) yöntemine oranla daha fazla olduğu sonucuna ulaşılmıştır. Benzer şekilde okul türü değişkenine göre yapılan analizde de LR (DMF gösteren madde sayısı:15) yöntemiyle belirlenen DMF'li madde sayısı MH (DMF gösteren madde sayısı:11) ve SIBTEST (DMF gösteren madde sayısı:7) yöntemine oranla daha fazla olduğu sonucuna ulaşılmıştır. Erdem (2015), TEOG 2015 sınavının alt testlerini farklı kitapçık türüne göre incelediği çalışmasında SIBTEST yöntemiyle MH ve LR testine oranla daha fazla DMF tespit ederek bu çalışmayla çelişkili bir sonuca ulaşmıştır. Benzer bir sonuca ulan Gök, Kelecioğlu ve Doğan (2010), MH ve LR yöntemleri ile belirlenen DMF'li madde sayısının uyumsuz olduğunu ifade etmişlerdir. Tespit edilen DMF'li madde sayısı bakımından yöntemler arasındaki uyumsuzluğun nedeni; grup ortalamaları, maddelerin ayıt edicilik ve güçlüklerindeki farklılıklar veya farklı yetenek düzeyleri olabilmektedir (Clauser, Mazor ve Hambleton, 1991; Narayanan ve Swaminathan, 1996; Swaminathan ve Rogers, 1990).

MH, LR ve SIBTEST analiz yöntemleri ile geniş ölçekli sınavlarda yapılan DMF araştırmalarında genellikle orta (B) düzeyde DMF belirlenmiştir (Narayanan ve Swaminathan, 1994; Rogers ve Swaminathan, 1993). Bu çalışmada da MH ve SIBTEST yöntemiyle yapılan analizlerde B düzeyinde DMF tespit edilmişken, LR yöntemi ile tespit edilen DMF'li maddelerin hepsi A düzeyinde kalmıştır. Farklı yöntemlere göre maddelerin DMF düzeylerinin farklı çıkması, yöntemlerin DMF düzeylerine ilişkin sınıflamalara ait değer aralıklarının farklı olmasından kaynaklanabilir (Arıkan-Akın, 2015). Bu durum araştırmacılar arasında tartışma konusu olan LR yönteminin DMF düzeyini belirleme kriteri hakkında hem fikir olmayıp da yeni bir sınıflama sistemi oluşturan Bakan Kalaycıoğlu ve Kelecioğlu (2011)'nin çalışması gözetilerek incelendiğinde LR yöntemiyle de B düzeyinde (cinsiyete göre 9. madde, okul türüne göre 1, 2, 3, 6, 7, 15 ve 19. madde) maddeler belirlenmektedir. Fakat bu maddelerin diğer yöntemlerle uyumlu olmaması ve alan

yazında daha çok tercih edilen sınıflamanın Jodoin ve Gierl (2001)'in sınıflaması olmasından ötürü bu sınıflama kullanılmıştır.

Cinsiyet değişkenine göre B düzeyinde DMF gösteren maddeler ise sırasıyla; madde 9 ve madde 11, hem MH hem de SIBTEST yöntemi ile madde 5 sadece SIBTEST yöntemi ile ve madde 16 ise sadece MH yöntemi ile belirlenen maddeler olarak tespit edilmiştir. Okul türü değişkenine göre ise sadece madde 4 hem MH hem de SIBTEST yöntemiyle B düzeyinde DMF'li madde olarak tespit edilmiştir. Bu sonuçlara ek olarak hesaplanan spearman rho katsayısına göre yöntemlerin etki büyüklüğü bakımından uyumu incelendiğinde MH ve SIBTEST yöntemlerinin birbiriyle uyumunun yüksek, LR yöntemi ile uyumunun ise düşük olduğu sonucuna ulaşılmıştır. Alan yazın incelendiğinde de benzer şekilde MH ve SIBTEST yöntemlerinin birbiriyle uyumunun yüksek, LR yöntemiyle ise uyumunun düşük olduğu çalışmalara rastlanmaktadır (Akın-Arık, Uğurlu ve Atar, 2016; Çepni, 2011; Kelecioğlu, Karabay ve Karabay, 2014; Toprak ve Yakar, 2017). Ayrıca Fidalgo, Ferreres ve Muniz (2004), MH ve SIBTEST yöntemlerinin özelliklerini tespit etmek amacıyla yaptıkları simülasyon çalışmasında da bu iki yöntem arasındaki uyumun yüksek olduğunu görmüşlerdir.

DMF çalışmalarının birçoğunda birlikte kullanılan MH ve LR yöntemleri genellikle birbirinden uyumsuz sonuçlar vermektedir (Ayan, 2011; Doğan ve Öğretmen, 2008; Zheng, Gierl ve Cui, 2007). Bertrand ve Boiteau (2003), da TIMMS verilerini kullanarak yaptıkları çalışmalarında MH ve LR yöntemlerinin DMF belirleme uyumlarının düşük olduğunu tespit etmişlerdir. Bu durumun nedeni MH yönteminin parametrik olmayan verilerle LR yönteminin ise parametrik verilerle çalışması olabilir.

Bu çalışmalardan farklı olarak; Gierl, Jodoin ve Ackerman (2000)' in DMF belirleme yöntemlerinin farklı örneklem büyüklüklerinde DMF belirleme gücünü araştırdıkları çalışmalarında ise MH ve LR yöntemlerinin uyumlu sonuçlar verdiği görülmektedir.

Son olarak bu çalışmayla çelişkili sonuçlara ulaşan Suna (2012) ve Demir (2013), çalışmalarında, MH ve LR yöntemlerinin SIBTEST yöntemine göre daha tutarlı sonuçlar verdiğini öne sürmüşlerdir.

Alan yazında, bu çalışmanın sonucuyla uyumlu ve çelişkili çalışmaların olması DMF yöntemlerinden herhangi birinin en doğru sonucu verdiğine dair genel bir yargının olmadığı kanıtıdır (Gierl, Khalig ve Boughton, 1999). Bu nedendir ki, bu çalışmada da olduğu gibi birçok çalışmada birden fazla yöntem kullanılarak analizler gerçekleştirilmektedir.

Öneriler

Bu çalışmada öneriler, test uygulayıcılarına ve test geliştirenlere yönelik öneriler ve araştırmacılara yönelik öneriler olmak üzere iki başlık altında verilmiştir.

Araştırmacılara yönelik öneriler. Bu araştırma TEOG İngilizce alt testine ait maddeler üzerinde cinsiyet ve okul türü değişkeni dikkate alınarak çalışılmıştır. Alan yazın incelendiğinde ülkemizde yapılan DMF ve madde yanlılığı çalışmalarında cinsiyet değişkeninin çok araştırıldığı fakat okul türü değişkeninin yeterince araştırılmadığı görülmektedir. Bu nedenle bu çalışmada da okul türüne göre yanlılık belirten bir maddenin tespit edildiği göz önünde bulundurularak ülkemizde uygulanan diğer merkezi sınavlara ait maddelerin okul türü değişkeni dikkate alınarak araştırılması araştırmacılara önerilebilir. Ayrıca bu çalışmada okul türü değişkeni devlet okulu – özel okul ayrımı yapılarak alt gruplara ayrılmıştır. Başka çalışmalar için bu alt gruplara imam hatip ortaokulları ve yatılı bölge ortaokullarının da dâhil edilebilmesi önerilebilir.

Bu çalışmada okul türü değişkenine göre DMF gösteren maddeler uzman görüşüne sunulduğunda, farklılaşmanın nedeni çoğunlukla, öğrencilerin sahip olduğu sosyoekonomik düzey farklılığı ile açıklanmıştır. Uzmanların bu görüşte olması, ülkemizde uygulanan sınavların bireylerin sosyoekonomik düzeylerine göre madde yanlılığı açısından incelenmesi için araştırmacılara önerilebilir.

Ülkemizde uygulanan merkezi sınavlarda maddelerin sıralamaları değiştirilerek farklı soru kitapçıkları hazırlanmaktadır. Bu durum alan yazında, yanıtlayıcıları güdülemesi nedeniyle geçerliği ve dolayısıyla güvenilirliği etkileyen bir durum olarak belirtilmektedir (Doğan, 2007; Tippets ve Benson, 1989). Bu durum dikkate alınarak bu çalışmada sadece A kitapçığını işaretleyen öğrencilere ait veriler kullanılmıştır. Bu nedenle farklı kitapçık türlerinin kullanıldığı durumlarda maddelerin test içindeki konumlarının DMF analiz sonuçlarına etkisinin araştırılması araştırmacılara önerilebilir.

Yanlılık çalışmalarının istatistiksel ayağında kullanılan DMF yöntemlerinin etkililiği ile ilgili çalışmalar çok fazla olmasına rağmen uzman görüşlerinin etkililiği ile ilgili çalışmalara pek rastlanmamaktadır. Bu çalışma kapsamında da yanlılık kavramına hakim uzmanlara ulaşma konusunda sıkıntı yaşanmıştır. Bu nedenle uzman görüşlerini daha etkili hale getirmek adına, çalışmayla ilgili fakat çeşitli uzmanlık alanlarına sahip kişiler tercih edilip bu kişilere madde yanlılığı ve DMF konularında seminerler verdikten sonra görüşlerinin alındığı çalışmaların yapılması araştırmacılara önerilebilir.

Alan yazında yer alan tüm DMF ve yanlılık çalışmaları taranıp DMF gösteren ve DMF göstermeyen maddelerin içeriği, formatı, madde kökünde yer alan negatif kelimeler, resimler, tablolar, madde ile ilişkili bilişsel boyutlar gibi olası kaynakların derlenip bir araya getirildiği bir çalışma araştırmacılara önerilebilir.

Alan yazında DMF ve madde yanlılığı ile ilgili çalışmalarda merkezi sınav verilerinin ilk olarak kullanıldığı çalışmalardan günümüzdekilerine kadar yıllarına göre sıralanıp gelişimi incelenebilir. DMF ve madde yanlılığı araştırmalarının artmasının uygulanan merkezi sınavların geçerliğine olumlu yönde etkisi olmuş mu diye incelenmesi araştırmacılara önerilebilir.

Bu çalışmada alan yazında en çok tercih edilen yöntemlerden MH, LR ve SIBTEST yöntemleri ile DMF analizleri gerçekleştirilmiş ve LR yönteminin diğer yöntemlerle uyumunun düşük olduğu tespit edilmiştir. Diğer DMF belirleme yöntemleri kullanılarak yapılan çalışmalarla bu yöntemlerin özellikleri geniş kapsamda tespit edilebilir. Bu sebeple DMF belirleme yöntemlerinin etkililiğinin incelenmesi adına farklı DMF analiz yöntemlerinin kullanılması araştırmacılara önerilebilir.

Test uygulayıcılarına ve test geliştirenlere yönelik öneriler. Alan yazında merkezi sınav verilerinin kullanıldığı çalışmalarda yabancı dil testlerinin kullanımına çok rastlanmamaktadır. Özellikle ebeveynlerin yabancı dil eğitimi konusunda özel okulları tercih etmeleri, merkezi sınavlarda yabancı dil testlerine yer verilmesi konusunda yanlılık oluşturur algısını yaratması nedeniyle bu testlerin merkezi sınavlarda olmalı mı olmamalı mı konusunun ilgili kurumlarca araştırılması önerilebilir.

Merkezi sınavların betimsel istatistiksel özellikleri her sınav sonrası araştırılıp kamoyu ile paylaşılmaktadır. Bu sınavların arařtırmalarına DMF ve yanlılık arařtırmalarının da eklenmesi bu sayede ilgili kurumların madde yanlılıđı oluřturacak olası sebepleri belirleyerek soru hazırlayanların dikkatine sunmaları önerilebilir. Ayrıca soru yazarlarına, maddelerin yanlılık oluřturabileceđi etkenlere dikkat edilebilmesi için DMF'nin olası kaynaklarına iliřkin eđitimlerin verilmesi önerilebilir.

Kadın ve erkek öğrencilerin biliřsel ve duyuřsal olarak farklı geliřime sahip oldukları göz önünde bulundurularak farklılık oluřturabilecek kaynaklar belirlenip soru yazarlarının dikkatine sunulmalıdır. Özellikle yabancı dil testleri, dil geliřimi aynı zaman aralıđında olmayan kadın ve erkek öğrenciler için hazırlanırken seçilen kavramlara ya da maddelerin formatına dikkat edilmesi önerilebilir. Benzer řekilde kadın ve erkek öğrencilerin farklı ilgili alanlarının olduđu ve ilgilerin kalıcı öğrenmeye olumlu etkisinin olduđu göz önünde bulundurularak test hazırlanma sürecinde bu durumun dikkate alınması ilgili kurumlara önerilebilir. Örneđin maddede mutlaka sporla ilgili kavramlar yer alacaksa hem kadınların hem de erkeklerin ařına olabileceđi spor dallarının tercih edilmesi önerilebilir. Hiç biri yapılamıyorsa da test hazırlanırken kadın ve erkek öğrencilere yanlılık oluřturabilecek madde sayılarının eřit tutulması önerilebilir.

Kaynaklar

- Abbott, M. L. (2007). A confirmatory approach to differential item functioning on an ESL reading assessment. *Language testing*, 24(1), 7-36.
- Ackerman, T. A., & Evans, J. A. (1992a). *An investigation of the relationship between reliability, power, and the type 1 error rate of the Mantel-Haenszel and simultaneous item bias detection procedures*. Paper presented at the 32nd Annual Meeting of the National Council on Measurement in Education. Retrieved from <http://files.eric.ed.gov/fulltext/ED344937.pdf>
- Ackerman, T. A. (1992b). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, 29(1), 67-91.
- Akın-Arık, Ç., Uğurlu, S. ve Atar, B. (2016). MIMIC, SIBTEST, Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle gerçekleştirilen DMF ve yanlılık çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(1), 34-52.
- Amrein, A.L. & Berliner, D.C. (2002, March 28). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). 1-74. Retrieved from <https://doi.org/10.14507/epaa.v10n18.2002>
- Angoff, W. H. (1993). Perspectives on differential item functioning methodology. In P.W. Holland ve H. Wainer (Eds.), *Differential item functioning* (pp. 3-23). Hillside, NJ: Lawrence Erlbaum.
- Akın-Arık, A. (2015). Değişen madde fonksiyonu belirlemede MTK-Olabilirlik Oranı, Ordinal Lojistik Regresyon ve Poly-Sıbtest yöntemlerinin karşılaştırılması. *E-Uluslararası Eğitim Araştırmaları Dergisi*, 6(1), 1-16. doi: 10.19160/e-ijer.24504
- Asil, M. ve Gelbal, S. (2012). PISA öğrenci anketinin kültürler arası eşdeğerliği. *Eğitim ve Bilim*, 37(166), 236-249.
- Atalay-Kabasakal, K., Arsan, N., Gök, B., ve Kelecioğlu, H. (2014). Değişen madde fonksiyonunun belirlenmesinde MTK Olabilirlik Oranı, SIBTEST ve Mantel Haenszel yöntemlerinin performanslarının (I. tip hata ve güç) karşılaştırılması. *Kuram ve Uygulamada Eğitim Bilimleri*, 14(6), 2175-2193.

- Atalay, K., Gök, B., Kelecioğlu, H., ve Arsan, N. (2012). Değişen madde fonksiyonunun belirlenmesinde kullanılan farklı yöntemlerin karşılaştırılması: Bir simülasyon çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 43, 270-281.
- Ayala, R. J. (2009). *Methodology in the social sciences: The theory and practice of item response theory*. New York: The Guilford Press.
- Ayan, C. (2011). *PISA 2009 fen okuryazarlığı alt testinin değişen madde fonksiyonu açısından incelenmesi* (Yüksek Lisans Tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Azen, R., Bronner, S., & Gafni, N. (2002). Examination of gender bias in university admissions. *Applied Measurement in Education*, 15(1), 75-94.
- Bakan-Kalaycıoğlu, D. (2008). *Öğrenci Seçme Sınavı'nın madde yanlılığı açısından incelenmesi* (Doktora Tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Bakan-Kalaycıoğlu, D. ve Kelecioğlu, H. (2011). Öğrenci Seçme Sınavı'nın madde yanlılığı açısından incelenmesi. *Eğitim ve Bilim*, 36(161), 3-12.
- Baykul, Y. (2010). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması* (2. Baskı). Ankara: Pegem Akademi.
- Benbow, C. P., & Stanley, J. C. (1982). Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal perspective. *American Educational Research Journal*, 19(4), 598-622.
- Berberoğlu, G., Demirtaşlı, N., Güzel Ç. İ., Arıkan, S., ve Tuncer, Ç. Ö. (2010). Okul dışı etmenlerin öğrenci başarısı ile ilişkisi. *Cito Eğitim: Kuram ve Uygulama*, 7, 27-36.
- Bertrand, R. & Boiteau, N. (2003). Comparing the stability of IRT-Based and non IRTbased DIF methods in different cultural contexts using TIMSS data. Retrieved from <https://files.eric.ed.gov/fulltext/ED476924.pdf>
- Buzick, H. ve Stone, E. (2011). Recommendations for conducting differential item functioning (DIF) analyses for students with disabilities based on previous DIF studies. ETS Research Report Series, 2, 1-26. Retrieved from <https://doi.org/10.1002/j.2333-8504.2011.tb02270.x>

- Büyüköztürk. Ş., Çokluk, Ö. ve Köklü, N. (2011). *Sosyal Bilimler İçin İstatistik* (7. Baskı). Ankara: Pegem Akademi.
- Camilli, G. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational measurement* (pp. 221-256). Westport: American Council on Education & Praeger Publishers.
- Camilli, G. & Shepard, L. A. (1994). *Methods for identifying biased test items*. California: Sage Publications
- Clauser, B. E., ve Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*, 17(1), 31-44. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1745-3992.1998.tb00619.x/pdf>
- Coley, R. J. (2001). *Differences in the gender gap: Comparison across racial/ethnic groups in education and work*. Retrieved from <https://www.ets.org/Media/Research/pdf/PICGENDER.pdf>
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando: Holt, Rinehart and Winston.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational Measurement* (pp. 443-507). Washington, D.C.: American Council on Education.
- Cronbach, L. J. (1984). *Essentials of psychological testing*. New York: Harper.
- Çelen, Ü. (2008). Klasik test kuramı ve madde tepki kuramı yöntemleriyle geliştirilen iki testin geçerlilik ve güvenilirliğinin karşılaştırılması. *İlköğretim Online*, 7(3), 758- 768.
- Çepni, Z. (2011). Değişen madde fonksiyonlarının SIBTEST, Mantel Haenzsel, Lojistik Regresyon ve Madde Tepki Kuramı yöntemleriyle incelenmesi (Doktora Tezi). Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Çokluk, O., Şekercioğlu, G. ve Büyüköztürk, S. (2010). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve Lisrel uygulamaları*. Ankara: Pegem A Yayıncılık.

- Demir, S. (2013). PISA 2009 matematik okuryazarlığı alt testinde bulunan maddelerinin Mantel-Haenszel, SIBTEST ve Lojistik Regresyon yöntemleri ile değişen madde fonksiyonunun incelenmesi (Yüksek Lisans Tezi). Abant İzzet Baysal Üniversitesi, Eğitim Bilimleri Enstitüsü, Bolu.
- Demir, S. ve Köse, İ. A. (2014). Mantel-Haenszel, SIBTEST ve Lojistik Regresyon yöntemleri ile değişen madde fonksiyonu analizi. *International Journal of Human Sciences*, 11(1), 700-714.
- Demircioğlu, G. (2011). Ölçme ve Değerlendirme (4. Baskı). In Emin Karip (Ed.), *Geçerlik ve Güvenirlik* (pp. 90-120). Ankara: Pegem Akademi.
- Doğan, N. (2007). Eğitimde Ölçme ve Değerlendirme (2.Baskı). In Hakan Atılğan (Ed.), *Çoktan Seçmeli Testler* (pp. 224-258). Ankara: Anı Yayıncılık.
- Doğan, N. ve Kılıç, A. F. (2017). Küreselleşen Dünyada Eğitim. In Özcan Demirel ve Serkan Dinçer (Ed.), *Madde Tepki Kuramı yetenek ve madde parametre kestirimlerinin değişmezliğinin incelenmesi* (pp. 297-314). Ankara: Pegem Akademi.
- Doğan, N. ve Öğretmen, T. (2005). Test ve madde yanlılığı. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 5 (1), 89-103.
- Doğan, N. ve Öğretmen, T. (2008). Değişen madde fonksiyonunu belirlemede Mantel-Haenszel, Ki-kare ve Lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 33(148), 100-112.
- Doğan, N. ve Tezbaşaran, A. A. (2003). Klasik test kuramı ve örtük özellikler kuramının örneklem bağlamında karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 25, 58-67.
- Doğan-Başokçu, Ö. ve Doğan, N. (2005). Akademik benlik kavramı ölçeğinin ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavını yordama geçerliği. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 29, 53-62.
- Doolittle, A. E., & Cleary, T. A. (1987). Gender-based differential item performance in mathematics achievement items. *Journal of Educational Measurement*, 24(2), 157-166.
- Dorans, N. J. & Holland, P. W. (1992). DIF detection and description: Mantel-haenszel and standardization 1, 2. *ETS Research Report Series*, 1, 1-40.

- Dorans, N. J., & Kulick, E. (1986). Demonstrating The Utility of the Standardization Approach to Assessing Unexpected Differential Item Performance On the Scholastic Aptitude Test. *Journal of Educational Measurement*, 23(4), 355–368. doi:10.1111/j.1745-3984.1986.tb00255.x
- ETS, Fairness Review Steering Committee (1998). *Overview: ETS Fairness Review*. NJ: Princeton.
- Ellis, B. B. & Raju, N. S. (2003). *Test and item bias: What they are, what they aren't, and how to detect them*. Retrieved from <https://files.eric.ed.gov/fulltext/ED480042.pdf>
- Ercikan, K. (1998). Translation effects in international assessments. *International Journal of Educational Research*, 29(6), 543-553.
- Erdem, B. (2015). Ortaöğretime geçişte kullanılan ortak sınavların değişen madde fonksiyonu açısından kitapçık türlerine göre farklı yöntemlerle incelenmesi (Yüksek Lisans Tezi). Hacettepe Üniversitesi, Ankara.
- Erdem-Keklik, D. (2014). Değişen madde fonksiyonu belirlemede mantel-haenszel ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(2), 12-25.
- Erkuş, A., Sünbül, Ö., Sünbül, S. Ö., Yormaz, S. ve Aşiret, S. (2017). Psikolojide ölçme ve ölçek geliştirme-II: Ölçme araçlarının psikometrik nitelikleri ve ölçme kuramları. Pegem: Ankara.
- Fennema, E. & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal*, 14, 51-71.
- Fidalgo, A. M., Ferreres, D., & Muñiz, J. (2004). Liberal and conservative differential item functioning detection using Mantel-Haenszel and SIBTEST: Implications for type I and type II error Rates. *The Journal of Experimental Education*, 73(1), 23-39. doi: 10.3200/JEXE.71.1.23-40
- Fraenkel, J., & Wallen, N. (2007). *How To Design And Evaluate Research in Education* (6.Baskı). New York, NY: McGraw-Hill.

- Gelbal, S. (1994). p Madde güçlük indeksi ile rasch modelinin b parametresi ve bunlara dayalı yetenek ölçüleri üzerine bir karşılaştırma. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 10(10).
- Gelbal, S. (2013). *Ölçme ve değerlendirme*. Eskişehir: Açıköğretim Fakültesi Yayını No: 1802
- Gierl, M., Khaliq, S., & Boughton, K. (1999). *Gender differential item functioning in mathematics and science: Prevalence and policy implications*. Paper presented at the annual meeting of the Canadian Society for the Study of Education, Sherbrooke, Quebec.
- Gómez-Benito, J. & Navas-Ara, M. J. (2000). A comparison of χ^2 , RFA and IRT based procedures in the detection of DIF. *Quality and Quantity*, 34(1), 17-31.
- Gotzmann, A. J., & Boughton, K. A. (2004). *A comparison of type I error and power rates for the Mantel-Haenszel and SIBTEST procedures when the group differences are large and unbalanced*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Gök, B., Kelecioğlu, H. ve Doğan, N. (2010). Değişen madde fonksiyonunu belirlemede Mantel-Haenszel ve Lojistik Regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 35(156), 3-16.
- Hafner, A., Ingels, S., Schneider, B., & Stevenson, D. (1990). *A profile of the American eighth grader: NELS:88 student descriptive summary*. Washington, DC: National Center for Education Statistics
- Halpern, D. F. (1997). Sex differences in intelligence: Implications for education. *American Psychologist*, 52, 1091-1102.
- Hambleton, R. K. & Jones, R. W. (1993). An NCME instructional module on: 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. & Rogers, H. J. (1989). Detecting potentially biased test items: Comparison of IRT area and Mantel-Haenszel methods. *Applied Measurement in Education*, 2(4), 313-334.

- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: Principles and application*. Kluwer: Kluwer Academic Publisher Group.
- Hambleton, R.K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of item response theory*. California: SAGE Publications.
- Hamilton Laura S. & Snow R. E. (1998). *Exploring differential item functioning on science achievement tests*. CSE Technical Report No: 483.
- Hamilton, L. S. (1999). Detecting gender-based differential item functioning on a constructed-response science test. *Applied measurement in Education*, 12(3), 211-235.
- Harris, A. M., & Carlton, S. T. (1993). Patterns of gender differences on mathematics items on the scholastic aptitude test. *Applied Measurement in Education*, 6(2), 137–151.
- Henderson, D. L. (2001). *Prevalence of gender DIF in mixed format high school exit examinations*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Hesapçioğlu, M. ve Nohutçu, A. (1999). Velilerin özel okul tercihlerini etkileyen faktörler ve özel okulların reklam stratejileri. *M.Ü. Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 11,183-202.
- Holland, P. W. & Thayer, D. T. (1986). Differential item functioning and the Mantel-Haenszel procedure. *ETS Research Report Series*, 2, 1-24.
- Holland, W. P., & Thayer, D. T. (1988). Differential item performance and the Mantel Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129–145). Hillsdale: LEA.
- Holland, P.W. Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum
- Holmes Finch, W. & French, B. F. (2007). Detection of crossing differential item functioning: A comparison of four methods. *Educational and Psychological Measurement*, 67(4), 565-582.

- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Ingels, S. J. (1990, April). *Findings from the NELS: 88 Base Year Student Survey*. Paper presented at the annual meeting of the American Educational Research Association, Boston.
- Ironson, G. H. (1982). Use of chi-square and latent trait approaches for detecting item bias. In R. A. Berk (Ed.), *Handbook of methods for detecting item bias* (pp. 117-155). Baltimore: Johns Hopkins University Press.
- Ironson, G. H. & Craig, R. (1982). *Item bias techniques when amount of bias is varied and score differences groups are presented*. University of South Florida, Tampa. (ERIC Document Reproduction Service No. ED 227 146)
- Jodoin, M. G. & Gierl, M. J. (2001). Evaluating type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14(4), 329-349.
- Kalaycıođlu, D. B. ve Keleciođlu, H. (2011). Öğrenci Seçme Sınavı'nın madde yanlılıđı açısından incelenmesi. *Eđitim ve Bilim*, 36(161), 3-13.
- Kan, A. (2007a). Eđitimde Ölçme ve Deđerlendirme (2.Baskı). In Hakan Atılđan (Ed.), *Ölçme Araçlarında Bulunması Gereken Nitelikler* (pp. 24-80). Ankara: Anı Yayıncılık.
- Kan, A. (2007b). Test yansızlıđı: Yabancı dil muafiyet sınavının cinsiyete ve bölümlere göre DMF analizi. *Eurasian Journal of Educational Research*, 29, 45-58.
- Kan, A., Sünbül, Ö. ve Ömür, S. (2013). 6.-8. Sınıf Seviye Belirleme sınavları alt testlerinin çeşitli yöntemlere göre deđişen madde fonksiyonlarının incelenmesi. *Mersin Üniversitesi Eđitim Fakóltesi Dergisi*, 9(2), 207-222.
- Karakaya, İ. (2012). Seviye belirleme sınavındaki fen ve teknoloji ile matematik alt testlerinin madde yanlılıđı açısından incelenmesi. *Kuram ve Uygulamada Eđitim Bilimleri*, 12(1), 215-229.
- Karakaya, İ. ve Kutlu, Ö. (2012). Seviye belirleme sınavındaki Türkçe alt testlerinin madde yanlılıđının incelenmesi. *Eđitim ve Bilim*, 37(165), 348-362.

- Karasar, N. (1999). *Bilimsel Araştırma Yöntemi* (9. Baskı). Ankara: Nobel Yayın Dağıtım.
- Kelecioğlu, H. (2001). Örtük özellikler teorisindeki b ve a parametreleri ile klasik test teorisindeki p ve r istatistikleri arasındaki ilişki. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 20(20), 104-110.
- Kelecioğlu, H. ve Göçer-Şahin, S. (2014). Geçmişten günümüze geçerlik. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(2), 1-11.
- Kelecioğlu, H., Karabay, B. ve Karabay, E. (2014). Seviye Belirleme Sınavı'nın madde yanlılığı açısından incelenmesi. *İlköğretim Online*, 13(3), 934-953.
- Kelloway, K. E. (1998). *Using Lisrel for Structural Equation Modeling: A Researcher's Guide*. London: Sage.
- Kliene, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. (Second Edition). NY: Guilford Publications, Inc.
- Koyuncu, İ., Aksu, G. ve Kelecioğlu, H. (2018). Mantel-Haenszel, Lojistik Regresyon ve Olabilirlik Oranı değişen madde fonksiyonu inceleme yöntemlerinin farklı yazılımlar kullanılarak karşılaştırılması. *İlköğretim Online*, 17(2), 909-925.
- Köse, İ. A. (2015). PISA 2009 öğrenci anketi alt ölçeklerinde (Q32-Q33) bulunan maddelerin değişen madde fonksiyonu açısından incelenmesi. *Kastamonu Eğitim Dergisi*, 23(1), 227-240.
- Kristanjansson, E., Aylesworth, R., McDowell, I. & Zumbo, B. D. (2005). A comparison of four methods for detecting differential item functioning in ordered response model. *Educational and Psychological Measurement*, 65(6), 935-953.
- Lietz, P. (2006). A meta-analysis of gender differences in reading achievement at the secondary school level. *Studies in Educational Evaluation* 32(4), 317–344.
- Linn, R. L. & Drasgow, F. (1987). Implications of the golden rule settlement for test construction. *Educational Measurement: Issues and Practice*, 6, 13-17.

- Linn, R. L., Levine, M. V., Hastings, C. N. & Wardrop, J. L. (1981). Item bias in a test of reading comprehension. *Applied Psychological Measurement*, 5(2), 159-173.
- Lohman, D. F. (1993). *Spatially gifted, verbally inconvenienced*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.126.8184&rep=rep1&type=pdf>
- Lord, M.F & Novick, R. M. (1968). *Statistical theories of mental test scores*. New York: Addison-Wesley Publishing Company.
- Lui, O. L. & Wilson, M. (2009). Gender differences in large scale math assessments: PISA trend 2000 and 2003. *Applied Measurement in Education*, 22, 164-184.
- Lynn, R. & Mikk, J. (2009). Sex differences in reading achievement. *Trames*, 13, 3-13.
- Mantel, N. & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719- 748. Retrieved from <https://doi.org/10.1093/jnci/22.4.719>
- MEB (2016a). *2015-2016 2. dönem ortak sınavlar sayısal bilgiler*. 2 Mart 2018 tarihinde <http://odsgm.meb.gov.tr/test/analizler/docs/2015-2016OrtakSinavlar2.DonemSayisalBilgiler.pdf> sayfasından erişilmiştir.
- MEB (2016b). 2015-2016 öğretim yılı ortak sınavlar e-kılavuzu. 2 Mart 2018 tarihinde http://www.meb.gov.tr/sinavlar/dokumanlar/2015/kilavuz/OrtakSinavlar_E_Klavuz2015_2016.pdf sayfasından erişilmiştir.
- MEB (2016c). 2015-2016 öğretim yılı ortak sınavlar II. dönem ortak sınav test ve madde İstatistikleri. 2 Mart 2018 tarihinde <http://odsgm.meb.gov.tr/test/analizler/docs/2015-2016-ortak-sinav-2-donem-madde-istatistikleri.pdf> sayfasından erişilmiştir.
- Mendes-Barnett, S., & Ercikan, K. (2006). Examining sources of gender DIF in mathematics assessments using a confirmatory multidimensional model approach. *Applied Measurement in Education*, 19(4), 289–304.

- Millsap, R. E. & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing measurement bias. *Applied psychological measurement, 17*(4), 297-334. doi: 10.1177/014662169301700401
- Nandakumar, R. (1993). Simultaneous DIF amplification and cancellation: Shealy-Stout's test for DIF. *Journal of Educational Measurement, 30*(4), 293-311.
- Narayanan, P. & Swaminathan, H. (1994). Performance of the Mantel-Haenszel and simultaneous item bias procedures for detecting differential item functioning. *Applied Psychological Measurement, 18*(4), 315-328.
- Narayanan, P. & Swaminathan, H. (1996). Identification of items that show nonuniform DIF. *Applied Psychological Measurement, 20*(3), 257-274.
- Nartgün, Ş. ve Kaya, A. (2016). Özel okul velilerinin beklentileri doğrultusunda okul imajı oluşturma. *Eğitim ve Öğretim Araştırmaları Dergisi, 5*(2), 153-167.
- O'Neill, K. A. & McPeck, W. M. (1993). Item and test characteristics that are associated with differential item functioning. In P. W. Holland ve H. Wainer, (Eds.), *Differential item functioning* (pp. 255-276). Hillsdale, NJ: Erlbaum.
- Osterlind, S. J. (1983). *Test item bias*. Beverly Hills: Sage Publications.
- Osterlind, S. J. & Everson, H. T. (2009). *Differential item functioning*. Londra: Sage Publications.
- Öğretmen, T. ve Doğan, N. (2004). OKÖSYS matematik alt testine ait maddelerin yanlılık analizi. *İnönü Üniversitesi Eğitim Fakültesi Dergisi, 5*(8), 61-75.
- Özçelik, D. A. (2010). *Ölçme ve Değerlendirme* (3. Baskı). Ankara: Pegem Akademi.
- Özdemir, D. (2003). Çoktan seçmeli testlerde iki kategorik ve önsel ağırlıklı puanlamanın diferansiyel madde fonksiyonuna etkisi ile ilgili bir araştırma. *Eğitim ve Bilim, 28*(129), 37-43.
- Özmen, D. T. (2014). PISA 2009 okuma testi maddelerinin yanlılığı üzerine bir çalışma. *Eğitim Bilimleri ve Uygulama, 13*(26), 147-165.
- Pae, T. I. & Park, G. P. (2006). Examining the relationship between differential item functioning and differential test functioning. *Language testing, 23*(4), 475-496.

- Pallant, J. (2007). *SPSS Survival manual: A step by step guide to data analysis using SPSS for windows* (3. Baskı). Berkshire, England: Open University Press.
- Phillips, A. & Holland, P. W. (1987). Estimators of the variance of the Mantel-Haenszel log-odds-ratio estimate. *Biometrics*, 43, 425-431.
- Potenza, M. T., & Dorans, N. J. (1995). DIF assessment for polytomously scored items: A framework for classification and evaluation. *Applied psychological measurement*, 19(1), 23-37.
- Raju, N. S. (1988). The area between two item characteristic curves. *Psychometrika*, 53(4), 495-502.
- Rodney, G. L. & Drasgow, F. (1990). Evaluation of two methods for estimating item response theory parameters when assessing differential item functioning. *Journal of Applied Psychology*, 75(2), 164-174.
- Rogers, H. J. & Swaminathan, H. (1993). A comparison of logistic regression and Mantel-Haenszel procedures for detecting differential item functioning. *Applied Psychological Measurement*, 17(2), 105-116.
- Ross, S. J. & Okabe, J. (2006). The subjective and objective interface of bias detection on language tests. *International Journal of Testing*, 6(3), 229-253.
- Roussos, L. A. & Stout, W. F. (1996a). Simulation studies of the effects of small sample size and studied item parameters on SIBTEST and Mantel-Haenszel type I error Performance. *Journal of Educational Measurement*, 33(2), 215-230.
- Roussos, L. A. & Stout, W. F. (1996b). A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement*, 20, 355-371.
- Satıcı, D. K., ve Özkan, Y. Ö. (2017). Temel Eğitimden Ortaöğretime Geçiş Sınavının (2014-Kasım) cinsiyet açısından madde yanlılığının incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 13(1), 254-274.
- Shealy, R. & Stout, W. F. (1993). A model-based standardization approach that separates true bias/ DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika*, 58, 159–194.

- Shepard, L.A., Camilli, G. & Averill, M. (1981). Comparison of procedures for detecting test item bias with both internal and external ability criteria. *Journal of Educational Statistics*, 6, 317-375.
- Shepard, L. A., Camilli, G. & Williams, D. M. (1984). Accounting for statistical artifacts in item bias research. *Journal of Educational Statistics*, 9(2), 93-128.
- Shepard, L. A., Camilli, G. & Williams, D. M. (1985). Validity of approximation techniques for detecting item bias. *Journal of Educational Measurement*, 22(2), 77-105.
- Stout, W. (2005). *SIBTEST Software*. Institute for Measurement: Assessment System Corporations.
- Stricker, L. J. (1981). *A new index of differential subgroup performance: Application to the GRE aptitude test*. Princeton, NJ: Educational Testing Service.
- Stricker, L. J., & Emmerich, W. (1999). Possible determinants of differential item functioning: Familiarity, interest, and emotional reaction. *Journal of Educational Measurement*, 36, 347–366
- Suna, H. E. (2012). *TIMMS 2007 Fen Bilimleri testindeki maddelerin dil ve cinsiyet yanlılığı açısından incelenmesi* (Yüksek Lisans Tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Sünbül, Ö. ve Erkuş, A. (2013). Madde parametrelerinin değişmezliğinin çeşitli boyutluluk özelliği gösteren yapılarda madde tepki kuramına göre incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 9(2), 378-398.
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- Şenferah, S. (2015). 2010 seviye belirleme sınavı matematik alt testi için değişen madde fonksiyonlarının ve madde yanlılığının incelenmesi (Yüksek Lisans Tezi). Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Tippets, E. & Benson, J. (1989). The effect of item arrangement on test anxiety. *Applied Measurement in Education*, 2(4), 289-296.

- Toprak, E. ve Yakar, L. (2017). SBS Türkçe alt testindeki maddelerin deęişen madde fonksiyonu açısından farklı yöntemlerle incelenmesi. *International Journal of Eurasia Social Sciences*, 8(26), 220-231.
- Turgut, M. F. ve Baykul, Y. (2010) *Eđitimde Ölçme ve Deęerlendirme* (2.Baskı). Ankara: Pegem Akademi.
- Turgut, M.F. (1995). *Eđitimde Ölçme ve Deęerlendirme Metotları* (10.Baskı). Ankara: Yargıcı Matbaası.
- Ulutaş, S. (2012). *PISA 2006 Fen okuryazarlığı testindeki maddelerin yanlılık bakımından araştırılması* (Doktora Tezi). Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Uttaro, T. & Millsap, R. E. (1994). Factors influencing the Mantel-Haenszel procedure the detection of differential item functioning. *Applied Psychological Measurement*, 18(1), 15-25.
- Uyar, Ş. ve Kaya-Uyanık, G. (2016). PISA 2012 bilişsel maddelerinin kültüre göre deęişen madde fonksiyonu bakımından incelenmesi. *Eđitim ve Öğretim Araştırmaları Dergisi*, 5(3), 230-240.
- Uzun, N. B. ve Gelbal, S. (2017). PISA fen başarı testinin madde yanlılığının kültür ve dil açısından incelenmesi. *Kastamonu Eğitim Dergisi*, 25(6), 2427-2446.
- Waller, N. G. (1998). EZDIF: A computer program for detecting uniform and non-uniform differential item functioning with the Mantel-Haenszel and logistic regression procedures. *Applied Psychological Measurement*, 22(4), 391-391. Retrieved from <https://doi.org/10.1177/014662169802200409>
- Yakar, L. ve Yavuz, S. (2014). *Soru kitapçıklarına göre deęişen madde fonksiyonları*. IV. Ulusal Eğitimde ve Psikolojide Ölçme ve Deęerlendirme Kongresi Ankara, 9 - 13 Haziran 2014.
- Yang, Y. (2003). Dimensions of socio economic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian Journal of Educational Research*, 47, 21-41.
- Yıldırım, H. (2015), 2012 Seviye belirleme sınavı matematik alt testinin madde yanlılığı açısından incelenmesi (Yüksek Lisans Tezi). Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

- Yıldırım, S. (2008). Farklı işleyen maddelerin belirlenmesinde sınırlandırılmış faktör çözümlemesinin olabilirlik-oranı ve Mantel-Haenszel yöntemleriyle karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 34, 297-307.
- Yılmaz, H. B. (2009). On the development and measurement of spatial ability. *International Electronic Journal of Elementary Education*, 1(2), 83-96.
- Yurdugül, H. ve Aşkar, P. (2004a). Ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavının, öğrencilerin yerleşim yerlerine göre, diferansiyel madde fonksiyonu açısından incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 27, 268-275.
- Yurdugül, H. ve Aşkar, P. (2004b). Ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavının, cinsiyete göre, madde yanlılığı açısından incelenmesi. *Eğitim Bilimleri ve Uygulama*, 3(5), 3-20.
- Zenisky, A. L., Hambleton, R. K. & Robin, F. (2004). DIF detection and interpretation in large-scale science assessments: Informing item writing practices. *Educational Assessment*, 9(1-2), 61-78.
- Zheng, Y., Gierl, M. J., & Cui, Y. (2007). *Using real data to compare DIF detection and effect size measures among Mantel-Haenszel, SIBTEST and Logistic Regression procedures*. Paper presented at NCME 2007, Chicago
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P.W. Holland ve H. Wainer (Eds.), *Differential item functioning* (pp.337-347). Hillsdale NJ: Erlbaum.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF)*. Ottawa: National Defense Headquarters.
- Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4(2), 223-233.
- Zwick, R. (1990). When do item response function and Mantel-Haenszel definitions of differential item functioning coincide? *Journal of Educational Statistics*, 15(3), 185-197. doi: 10.3102/10769986015003185

- Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of educational measurement*, 26(1), 55-66.
- Zwick, R., Donoghue, J. R. & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30(3), 233-251.

EK-A: TEOG Sınavına Ait İngilizce Alt Testi



A

2015-2016 EĞİTİM - ÖĞRETİM YILI

2. MERKEZİ ORTAK SINAV

İNGİLİZCE

1. Bu testte 20 soru vardır.
2. Cevaplarınızı, cevap kağıdına işaretleyiniz.

1-8. sorularda boş bırakılan yere uygun gelen kelime ya da ifadeyi işaretleyiniz.

1. Amy : Who uses the Internet most in your family?

Bernard : - - - - , my sister. She uses the Internet about 6 hours a day.

- A) I mean B) I guess
C) You're right D) You're welcome

2. Lucy : Hello, this is Lucy.

Emma : Hi, Lucy. - - - - ?

Lucy : Fine, thanks.

- A) Where are you going
B) Why are you asking
C) How is it going
D) Who is calling

3. Kate : I argued with my best friend last week.

Sophie : Oh really? - - - - .

- A) I'm sorry to hear that
B) I guess you can come
C) I hope to see you there
D) I'm sure you are available

4. Pelin : - - - - ?

Sue : Never. I prefer face-to-face interaction.

- A) When do you go out with your parents
B) What are its technical specifications
C) What kind of films do you watch
D) How often do you chat online

5. Candy : - - - - ?

Sally : Chinese. And over 1 billion people speak it in the world.

- A) Where would you like to visit
B) What is their official language
C) Where can a tourist stay there
D) What is the name of the country

6. Bill : How should we cook the vegetables?

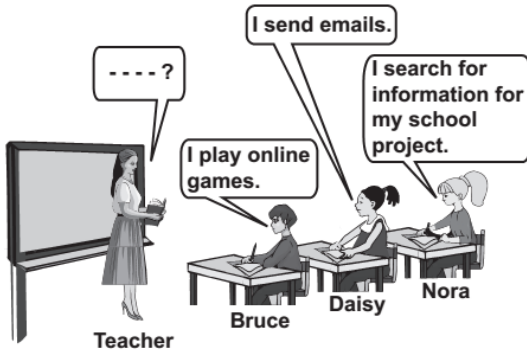
Nancy : You - - - - .

- A) should have all the ingredients
B) must serve them immediately
C) should steam them
D) must wash them

7. Mike : Are you doing anything on Sunday?
Sally : Yes, - - - - .

- A) we are getting on well with classmates
B) we are going to meet with friends
C) I really like sci-fi movies
D) I'm good at fixing bikes

8.



- A) How many hours a day do you use the Net
B) How often do you do homework on the Net
C) What do you usually do on the Internet
D) Who spends much time on the Internet

9. I. Hello, I'm Sarah.

II. Oh, sorry Sarah. This is Jim's father.

III. Could I ask who is calling?

IV. Hi, Sarah. Could I speak to Jim?

Yukarıda bir telefon konuşması karışık olarak verilmiştir.

Anlamalı bir diyalog oluşturmak için numaralanmış cümleler nasıl sıralanmalıdır?

- A) I, IV, III, II
B) IV, II, III, I
C) I, II, IV, III
D) IV, III, I, II

10 ve 11. sorularda verilen görsele göre boş bırakılan yere uygun gelen seçeneği işaretleyiniz.

10.



I prefer - - - - because I like exploring the blue waters.

- A) bungee jumping
B) parachute diving
C) hang gliding
D) scuba diving

11.



Mr. Carter

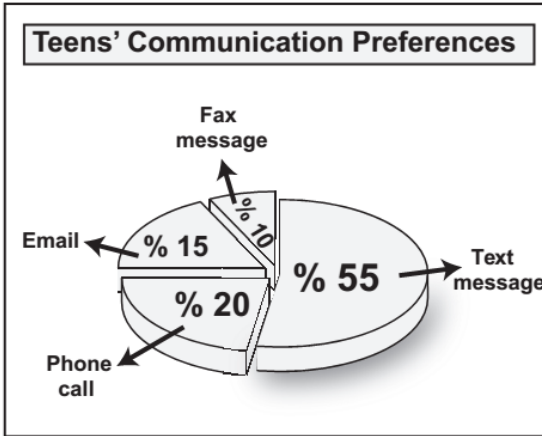
Mr. Carter - - - - when he was young.

- A) enjoyed white-water rafting
B) hated doing bungee jumping
C) tried hang gliding several times
D) disliked performing parkour running

12. Your friend invites you to a party. But you have another plan. What do you say to refuse his invitation?

- A) I hope you feel better soon.
- B) Sounds good but I'm busy.
- C) It will be great to join.
- D) Sure, I would like to.

13. soruyu aşağıda verilen grafiğe göre cevaplayınız.



13. Teens - - - - .

- A) make phone calls most
- B) never prefer sending fax messages
- C) enjoy sending text messages the least
- D) send text messages more than emails

14. "I like mopping the floor. It's my favorite chore." ifadesini anlatan görsel aşağıdakilerden hangisidir?

A)



B)



C)



D)



15-17. soruları aşağıda verilen metinlere göre cevaplayınız.



Egypt is the centre of many civilizations and there are many historic buildings to visit. But it is really too hot in summers so I prefer going there in mild winter days.

Linda

I really enjoy being outdoors and on the water. That's why with its blue waters, Italy is a unique tourist destination for me. Also, Italy has mild weather the year around so it is really convenient for water sports. I would like to go there and enjoy its crystal blue waters.



Robert



I enjoy exploring tropical islands like Jamaica. Jamaica is well-known for its many traditional dishes and I like trying new meals. With its mild climate the year around, Jamaica is the best choice for my summer holidays.

Peter

15. Egypt - - - .

- A) is the home of different cultures
- B) is a great place for water sports
- C) has a mild climate in summers
- D) has many wonderful dishes

16. Robert would like to visit Italy because - - - .

- A) it is usually cold in winters
- B) he wants to do water sports
- C) it has many historic buildings
- D) he would like to try new dishes

17. Peter - - - .

- A) thinks Jamaica is a historic place
- B) prefers going to Jamaica in winter
- C) never visits countries with mild climate
- D) would rather spend his holiday in Jamaica

18-20. soruları aşağıda verilen metne göre cevaplayınız.

Hi, I'm Jeremy. In our family everyone has some responsibilities. We always share the daily chores at home. I must water the flowers in the garden. I like it because I love being outdoors. My brother Robert is responsible for taking out the garbage. He hates doing it but he always does it. Doing the grocery shopping is my father's duty. My mother cooks the meals and does the laundry. Also, she has to wash the dishes and my father helps her with that. I think we all should help each other with the chores because sharing the responsibilities means respecting each other.

18. Jeremy spends some time in the garden because he - - - .

- A) must take care of the plants
- B) hates taking out the garbage
- C) loves going shopping for food
- D) is responsible for washing the dishes

19. Robert always takes the rubbish out but he - - - .

- A) always helps his father
- B) never does his tasks
- C) likes being outdoors
- D) dislikes doing it

20. Jeremy's father helps mom in the kitchen because - - - .

- A) he loves being alone
- B) they respect each other
- C) he likes doing the laundry
- D) they enjoy watering the plants

TEST BİTTİ.
CEVAPLARINIZI KONTROL EDİNİZ.

EK-B: Uzman Görüşü Anket Formu

UZMAN GÖRÜŞÜ ANKET FORMU

Sayın Uzman,

Hacettepe Üniversitesi Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalında Prof. Dr. Nuri DOĞAN danışmanlığında yürütmekte olduğum yüksek lisans tez çalışmam kapsamında, TEOG 2015–2016 Eğitim–Öğretim Yılı II. Dönem İngilizce alt testine ait maddelerin, testi alan öğrencilere cinsiyet ve öğrenim gördüğü okul türüne (özel-devlet) göre yanlılık taşıyor mu sorusuna yanıt bulmayı amaçlamaktayım. Bu kapsamda Değişen Madde Fonksiyonu gösteren maddeler tarafımda belirlenmiş olup bu maddelerin alt gruplardan herhangi birine avantaj sağlayıp sağlamadığı ve sağlıyorsa olası kaynaklarının ne olabileceği hakkında görüşünüze ihtiyaç duymaktayım. Aşağıda sunduğum açıklamalar Değişen Madde Fonksiyonu, Madde Etkisi ve Madde Yanlılığı hakkında size ön bilgi vermek amaçlıdır.

Değişen Madde Fonksiyonu (DMF), maddenin ölçmek istediği yetenek düzeyinde eşleşen fakat cinsiyet, sosyo-ekonomik düzey, kültür gibi farklı alt gruplardan bireylerin maddedeki başarı olasılıklarının farklılaşmasından kaynaklanır (Swaminathan ve Rogers, 1991; Zumbo, 1999). Bu farklılaşma ya madde yanlılığından ya da gerçek bilgi ve beceri farklılığından (madde etkisi) kaynaklanıyor olabilir. Bu nedenle maddenin, DMF göstermesinin olası kaynakları bakımından incelenmesi gerekmektedir. DMF'nin olası kaynakları aşağıda verilmiştir (Doolittle ve Clearly, 1987; Berberoğlu, 1995; Li, Cohen ve Ibarra, 2004; Kalaycıoğlu ve Berberoğlu, 2010; Kelecioğlu vd., 2014; Yıldırım ve Berberoğlu, 2009).

- Sosyo-ekonomik farklılıklar
- Bilgiye ulaşma imkânları (internet, bilgisayar, dersane vb.)
- Müfredat kapsamındaki farklılıklar
- Kullanılan öğretim metotları
- Öğrencilerin kavramlara olan aşinalık düzeyi
- Öğrencilerin ilgi alanları (Örneğin; Bireyler aynı yetenek düzeyinde olmalarına rağmen bir matematik sorusunun konusu basketbol, futbol, araba yarışları veya avcılık gibi daha çok erkeklerin ilgisini çeken konulardan seçildiğinde, maddeleri erkeklerin doğru cevaplama olasılığı artmaktadır. Benzer şekilde, kız öğrencilerin daha aşina olduğu konuların yer aldığı maddelerde ise kız öğrencilerin maddeleri doğru cevaplama olasılığı artmaktadır)

- Soru çözümünde kullanılan stratejiler
- Bilişsel düşünme biçimleri
- Madde içeriği ve formatı; Maddenin okul dışı bilgi ve deneyim gerektirmesi, günlük yaşam problemlerine dayalı olması; algoritmik işlemlere ve hesaplama dayanması; aritmetik ve cebirsel ifadelerin kullanılması; soyut içeriklere sahip olması; tablo, şekil, grafik içermesi, hız, uzunluk, zaman, para, yüzde ve alan içerikli olması; görsellik içermesi; görsel ve uzamsal zeka gerektirmesi vb.

Madde etkisi; maddenin gruplara göre farklılaşmasının nedenini, maddenin ölçmek istediği yetenek düzeyinde eşleşen gruplar arasındaki gerçek bilgi veya beceri farklılığı olarak ortaya koymasındır.(Zumbo, 1999).

Madde yanlılığı ise maddenin, testte ölçülmek istenen özellik dışındaki bazı özelliklere sahip olan bireylere avantaj sağlarken, diğerlerine dezavantaj sağlamasıdır (Osterlind ve Everson, 2009). Bu durumun daha iyi ifade edilebilmesi için daha önceki çalışmalarda yanlı olduğu tespit edilen bir madde örnek olarak sunulmuştur.

Madde Yanlılığı Örneği;

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 tetiklemeindedir. 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, nesnelere etrafına sarılabilir ya da sıkılıp top durumuna getirilebilir. Bilim adamları, onun toptan üretiminin ucuz olacağını da ileri sürmektedirler.

Soru 1: GİYSİLER

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

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

N. Bilge Başusta (2013), “PISA 2006 Fen Başarı Testinin Madde Yanlılığının Kültür ve Dil Açısından İncelenmesi” adlı tez çalışmasında maddelerin farklı alt kültürlerle (İngiltere – Türkiye) göre herhangi bir gruba yanlılık gösterip göstermediğini incelemeyi amaçlamış ve çalışmasında yukarıdaki maddenin İngiltere lehine avantaj sağladığını belirtmiştir.

Uzmanlarla yaptığı çalışmalar sonrasında bu maddenin Türkiye'nin aleyhine çalışmasının nedenini; Türkiye'de ilköğretim düzeyindeki okullarda uygun şartlarda laboratuvarlarının bulunmaması, bulunsa bile birçok okulda yeterince deney malzemesine ulaşamadığı ve ayrıca iletkenliğinin test edilmesine ilişkin ülkemiz müfredatında benzer soru tiplerinin kullanılmayışının etkili olabileceğini ifade etmiştir.

Değerli uzman, içtenlikle vereceğiniz cevaplar çalışmam için çok önemlidir. İlginiz ve desteğiniz için şimdiden teşekkür ederim.

Saygılarımla,

Münevver ARSLAN
Hacettepe Üniversitesi
Eğitimde Ölçme Değerlendirme A.B.D.
Yüksek Lisans Öğrencisi

İletişim: 0545 958 97 07

e-mail: munevverarsalan@yandex.com

DMF BELİRLENEN TEST MADDELERİ

MADDE:

1-8. sorularda boş bırakılan yere uygun gelen kelime ya da ifadeyi işaretleyiniz.

4. Pelin : - - - - ?

Sue : Never. I prefer face-to-face interaction.

- A) When do you go out with your parents
- B) What are its technical specifications
- C) What kind of films do you watch
- D) How often do you chat online

Yapılan DMF analizleri sonucu yandaki maddenin, İngilizce alt testini alan aynı yetenek düzeyindeki öğrenci gruplarından özel okul öğrencilerine avantaj sağladığı tespit edilmiştir.

Bu maddenin okul türüne göre DMF göstermesinin nedeni ne olabilir?

(1) madde etkisi

(2) madde yanlılığı

Eğer madde yanlılığından kaynaklandığını düşünüyorsanız olası nedenleri nelerdir?

Lütfen görüşlerinizi aşağıya yazınız.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

MADDE:

9. I. Hello, I'm Sarah.
II. Oh, sorry Sarah. This is Jim's father.
III. Could I ask who is calling?
IV. Hi, Sarah. Could I speak to Jim?

Yukarıda bir telefon konuşması karışık olarak verilmiştir.

Anlamli bir diyalog oluřturmak iin numaralanmıř cümleler nasıl sıralanmalıdır?

- A) I, IV, III, II B) IV, II, III, I
C) I, II, IV, III D) IV, III, I, II

Yapılan DMF analizleri sonucu yandaki maddenin, İngilizce alt testini alan aynı yetenek düzeyindeki öğrenci gruplarından **kadınlara** avantaj sağladığı tespit edilmiştir.

Bu maddenin cinsiyete göre DMF göstermesinin nedeni ne olabilir?

(1) madde etkisi

(2) madde yanlılığı

Eğer madde yanlılığından kaynaklandığını düşünüyorsanız olası nedenleri nelerdir?

Lütfen görüşlerinizi aşağıya yazınız.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

MADDE:

10 ve 11. sorularda verilen görsele göre boş bırakılan yere uygun gelen seçeneği işaretleyiniz.

11.



Mr. Carter

Mr. Carter - - - - when he was young.

- A) enjoyed white-water rafting
- B) hated doing bungee jumping
- C) tried hang gliding several times
- D) disliked performing parkour running

Yapılan DMF analizleri sonucu yandaki maddenin, İngilizce alt testini alan aynı yetenek düzeyindeki öğrenci gruplarından **erkekler**e avantaj sağladığı tespit edilmiştir.

Bu maddenin cinsiyete göre DMF göstermesinin nedeni ne olabilir?

- (1) madde etkisi
- (2) madde yanlılığı

Eğer madde yanlılığından kaynaklandığını düşünüyorsanız olası nedenleri nelerdir?

Lütfen görüşlerinizi aşağıya yazınız.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

MADDE:

15-17. soruları aşağıda verilen metinlere göre cevaplayınız.



Egypt is the centre of many civilizations and there are many historic buildings to visit. But it is really too hot in summers so I prefer going there in mild winter days.

Linda

I really enjoy being outdoors and on the water. That's why with its blue waters, Italy is a unique tourist destination for me. Also, Italy has mild weather the year around so it is really convenient for water sports. I would like to go there and enjoy its crystal blue waters.



Robert



I enjoy exploring tropical islands like Jamaica. Jamaica is well-known for its many traditional dishes and I like trying new meals. With its mild climate the year around, Jamaica is the best choice for my summer holidays.

Peter

16. Robert would like to visit Italy because - - - - .

- A) it is usually cold in winters
- B) he wants to do water sports
- C) it has many historic buildings
- D) he would like to try new dishes

Yapılan DMF analizleri sonucu yandaki maddenin, İngilizce alt testini alan aynı yetenek düzeyindeki öğrenci gruplarından **kadınlara** avantaj sağladığı tespit edilmiştir.

Bu maddenin cinsiyete göre DMF göstermesinin nedeni ne olabilir?

(1) madde etkisi

(2) madde yanlılığı

Eğer madde yanlılığından kaynaklandığını düşünüyorsanız olası nedenleri nelerdir?

Lütfen görüşlerinizi aşağıya yazınız.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

EK-C: İngilizce Alt Testine Ait Madde Faktör Yükleri

Component Matrix^a

	Component	
	1	2
VAR00018	.718	-.097
VAR00019	.710	-.273
VAR00012	.683	-.027
VAR00005	.656	-.122
VAR00002	.654	-.257
VAR00013	.638	-.030
VAR00020	.636	-.031
VAR00017	.628	.139
VAR00007	.627	-.034
VAR00011	.613	.292
VAR00015	.595	-.328
VAR00006	.588	-.295
VAR00004	.581	.248
VAR00016	.579	.445
VAR00014	.577	.128
VAR00008	.570	.230
VAR00003	.568	-.334
VAR00001	.556	-.245
VAR00010	.551	.461
VAR00009	.360	.438

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Rotated Component Matrix^a

	Component	
	1	2
VAR00019	.725	.231
VAR00002	.671	.209
VAR00015	.669	.116
VAR00003	.653	.095
VAR00006	.643	.138
VAR00018	.621	.373
VAR00005	.588	.315
VAR00001	.587	.156
VAR00012	.550	.406
VAR00013	.517	.375
VAR00020	.516	.373
VAR00007	.511	.366
VAR00016	.173	.709
VAR00010	.141	.704
VAR00011	.296	.611
VAR00009	.007	.566
VAR00004	.298	.557
VAR00008	.301	.536
VAR00017	.403	.501
VAR00014	.370	.461

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.^a

a. Rotation converged in 3 iterations.

**EK-Ç: Cinsiyete Göre İngilizce Alt Testi Maddelerinin MH Analizi İçin Yapılan
EZDIF Program Çıktısı**

Mantel-Haenszel and Logistic Regression Analysis of
DIFFERENTIAL ITEM FUNCTIONING

Programmed by Niels G. Waller

EZDIF deneme

Reference Group: D:\erkek.prn

Focal Group: D:\kadın.prn

Number of Cases in Reference Group: 4.864

Number of Cases in Focal Group: 4.244

Conditioning Levels

0	5	9	13	16	19
4	8	12	15	18	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	MH D-DIF	SE (MH D-DIF)
A 8	0,911	3,111	0,078	0,220	0,123
A 44	0,927	1,635	0,201	0,179	0,136
A 48	1,064	1,231	0,267	-0,145	0,128
A 52	0,911	1,972	0,160	0,218	0,152
A 57***	1,439	38,219	0,000	-0,856	0,139
A 84**	1,162	7,194	0,007	-0,353	0,130
A 95***	0,827	10,293	0,001	0,447	0,138
A 99***	0,708	27,558	0,000	0,812	0,154
B 106***	0,629	52,411	0,000	1,090	0,151
A 108	1,105	1,577	0,209	-0,234	0,181
B 111***	1,756	67,070	0,000	-1,323	0,162
A 115	0,987	0,032	0,859	0,030	0,145
A 136*	1,125	3,991	0,046	-0,276	0,137
A 140**	0,835	7,837	0,005	0,423	0,150
A 173***	1,332	25,434	0,000	-0,673	0,133
B 193***	0,626	42,220	0,000	1,100	0,169
A 207***	0,836	8,789	0,003	0,422	0,141
A 208	1,028	0,154	0,695	-0,064	0,151
A 209*	1,179	6,595	0,010	-0,388	0,149
A 227***	0,832	10,306	0,001	0,433	0,134

Number of Items Purged in Pass 1: 0

Item Numbers:

**EK-D: Cinsiyete Göre İngilizce Alt Testi Maddelerinin LR Analizi İçin Yapılan
EZDIF Program Çıktısı**

Logistic Regression Procedure.

Note: Items purged from previous step

Madde 1	Item Number: 8	Estimate	SE Estimate	Z	p-value
	Intercept	-3,225213	0,078852	-40,902077	0,000000
	Trait	0,258914	0,005449	47,515324	0,000000
	Group	0,599091	0,155269	3,858399	0,000121
	Trait x Group	-0,049153	0,010898	-4,510223	0,000007
Madde 2	Item Number: 44	Estimate	SE Estimate	Z	p-value
	Intercept	-3,980671	0,092826	-42,882954	0,000000
	Trait	0,337823	0,006626	50,982773	0,000000
	Group	0,848831	0,177058	4,794077	0,000002
	Trait x Group	-0,071070	0,013252	-5,362787	0,000000
Madde 3	Item Number: 48	Estimate	SE Estimate	Z	p-value
	Intercept	-3,382011	0,080329	-42,101810	0,000000
	Trait	0,269845	0,005554	48,587250	0,000000
	Group	0,432104	0,158519	2,725878	0,006525
	Trait x Group	-0,025227	0,011108	-2,271125	0,023351
Madde 4	Item Number: 52	Estimate	SE Estimate	Z	p-value
	Intercept	-2,640013	0,110111	-23,975929	0,000000
	Trait	0,358640	0,008647	41,477772	0,000000
	Group	-0,396465	0,173655	-2,283065	0,022636
	Trait x Group	0,035012	0,017293	2,024603	0,043174
Madde 5	Item Number: 57	Estimate	SE Estimate	Z	p-value
	Intercept	-3,832144	0,093685	-40,904564	0,000000
	Trait	0,346943	0,006826	50,823067	0,000000
	Group	0,608886	0,174514	3,489041	0,000506
	Trait x Group	-0,017314	0,013653	-1,268116	0,205052
Madde 6	Item Number: 84	Estimate	SE Estimate	Z	p-value
	Intercept	-3,889592	0,086053	-45,200031	0,000000
	Trait	0,287575	0,005826	49,360497	0,000000
	Group	0,945163	0,173412	5,450394	0,000000
	Trait x Group	-0,055055	0,011652	-4,724924	0,000003
Madde 7	Item Number: 95	Estimate	SE Estimate	Z	p-value
	Intercept	-3,229563	0,092962	-34,740715	0,000000
	Trait	0,331660	0,006951	47,715660	0,000000
	Group	0,095957	0,164835	0,582144	0,560601
	Trait x Group	-0,024487	0,013902	-1,761448	0,078468
Madde 8	Item Number: 99	Estimate	SE Estimate	Z	p-value
	Intercept	-2,574692	0,115362	-22,318422	0,000000
	Trait	0,359048	0,009063	39,617275	0,000000
	Group	0,148715	0,178272	0,834202	0,404366
	Trait x Group	-0,052456	0,018126	-2,893958	0,003887
Madde 9	Item Number:106	Estimate	SE Estimate	Z	p-value
	Intercept	-0,586756	0,086367	-6,793771	0,000000
	Trait	0,191643	0,006434	29,786671	0,000000
	Group	-1,034291	0,148821	-6,949903	0,000000
	Trait x Group	0,056554	0,012868	4,394999	0,000012
Madde 10	Item Number:108	Estimate	SE Estimate	Z	p-value
	Intercept	-2,893696	0,169240	-17,098198	0,000000
	Trait	0,477987	0,013793	34,653671	0,000000
	Group	-0,118726	0,226758	-0,523578	0,600688
	Trait x Group	0,034350	0,027587	1,245180	0,213357

Madde 11	Item Number:111	Estimate	SE Estimate	Z	p-value
	Intercept	-3,289329	0,127455	-25,807787	0,000000
	Trait	0,423585	0,010130	41,814045	0,000000
	Group	0,526613	0,195646	2,691667	0,007228
	Trait x Group	0,009013	0,020260	0,444852	0,656523
Madde 12	Item Number:115	Estimate	SE Estimate	Z	p-value
	Intercept	-3,977756	0,104723	-37,983612	0,000000
	Trait	0,385827	0,007857	49,106033	0,000000
	Group	0,620810	0,186025	3,337236	0,000877
	Trait x Group	-0,054345	0,015714	-3,458411	0,000566
Madde 13	Item Number:136	Estimate	SE Estimate	Z	p-value
	Intercept	-3,428247	0,090879	-37,723198	0,000000
	Trait	0,331799	0,006719	49,385853	0,000000
	Group	0,013629	0,165488	0,082354	0,934382
	Trait x Group	0,011201	0,013437	0,833577	0,404718
Madde 14	Item Number:140	Estimate	SE Estimate	Z	p-value
	Intercept	-2,632579	0,109799	-23,976326	0,000000
	Trait	0,351410	0,008568	41,011864	0,000000
	Group	0,347057	0,173837	1,996454	0,046155
	Trait x Group	-0,053801	0,017137	-3,139469	0,001742
Madde 15	Item Number:173	Estimate	SE Estimate	Z	p-value
	Intercept	-3,972056	0,087780	-45,249886	0,000000
	Trait	0,294929	0,005942	49,633926	0,000000
	Group	1,220927	0,176801	6,905652	0,000000
	Trait x Group	-0,065693	0,011884	-5,527794	0,000000
Madde 16	Item Number:193	Estimate	SE Estimate	Z	p-value
	Intercept	-2,764220	0,140509	-19,672909	0,000000
	Trait	0,421902	0,011306	37,315685	0,000000
	Group	-0,916000	0,202236	-4,529351	0,000007
	Trait x Group	0,055677	0,022613	2,462226	0,013976
Madde 17	Item Number:207	Estimate	SE Estimate	Z	p-value
	Intercept	-3,100698	0,096832	-32,021290	0,000000
	Trait	0,342282	0,007370	46,445564	0,000000
	Group	-0,337199	0,166497	-2,025258	0,043106
	Trait x Group	0,016653	0,014739	1,129868	0,258803
Madde 18	Item Number:208	Estimate	SE Estimate	Z	p-value
	Intercept	-4,617799	0,113379	-40,728779	0,000000
	Trait	0,421724	0,008442	49,957485	0,000000
	Group	0,991561	0,205113	4,834218	0,000002
	Trait x Group	-0,081380	0,016883	-4,820138	0,000002
Madde 19	Item Number:209	Estimate	SE Estimate	Z	p-value
	Intercept	-5,074634	0,106663	-47,576294	0,000000
	Trait	0,395967	0,007471	52,999229	0,000000
	Group	1,064947	0,209699	5,078464	0,000000
	Trait x Group	-0,066116	0,014942	-4,424692	0,000011
Madde 20	Item Number:227	Estimate	SE Estimate	Z	p-value
	Intercept	-3,463662	0,088885	-38,968105	0,000000
	Trait	0,321450	0,006481	49,598698	0,000000
	Group	0,127626	0,164625	0,775251	0,438375
	Trait x Group	-0,024583	0,012962	-1,896541	0,058177

EK-E: Zumbo (1999)'nun Syntaxı

GET

FILE='binary.sav'.

EXECUTE .

compute item= item2.

compute total= scale.

compute grp= group.

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

* of uniform and non-uniform DIF).

LOGISTIC REGRESSION VAR=item

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

/CONTRAST (grp)=Indicator

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

execute.

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

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

* attributed to the uniform DIF.

LOGISTIC REGRESSION VAR=item

/METHOD=ENTER total /METHOD=ENTER grp

/CONTRAST (grp)=Indicator

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

execute.

EK-F: Cinsiyete Göre İngilizce Alt Testi Maddelerinin SIBTEST Analizi İçin Yapılan SIBTEST Program Çıktısı

name of input parameter file = sib.in
number of items on test = 20
name of file for Ref. grp. scores = C:\kadın1.prn
name of file for Focal grp. scores = C:\erkek1.prn
minimum no. of examinees per matching score cell = 2
number of runs for this data set = 20
number of examinees in Reference group = 4.244
number of examinees in Focal group = 4.864

Examinee Test Score Summary Statistics

Reference Group: Mean = 14,47
Standard deviation = 5,23
Focal Group: Mean = 13,04
Standard deviation = 5,59

Standardized Score Difference = 0,26

Item Statistics

= item number

p = proportion right on item (Classical Test Theory p-value)

r = point biserial (item score-test score correlation)

#:	1	2	3	4	5	6	7	8	9	10
p:	0,565	0,604	0,564	0,790	0,640	0,523	0,690	0,795	0,835	0,857
r:	0,567	0,656	0,582	0,575	0,655	0,598	0,627	0,565	0,372	0,539

#:	11	12	13	14	15	16	17	18	19	20
p:	0,789	0,674	0,667	0,784	0,528	0,827	0,719	0,649	0,563	0,644
r:	0,604	0,677	0,637	0,571	0,603	0,567	0,621	0,710	0,709	0,635

p-value notation:

R denotes p-value for test of DIF/DBF against Ref. group

F denotes p-value for test of DIF/DBF against Foc. group

E denotes p-value for test of DIF/DBF against either the
Ref. or Foc. group.

NOTES:

MS/SSD = Matching Subtest Standardized Score Difference.
Standardized difference in mean observed scores
between Reference group and Focal group on the
matching subtest.

p-elim = proportion of Reference (R) and Focal (F) groups
eliminated (not used) in SIBTEST calculations.

Positive Beta estimate indicates DIF/DBF favoring Ref. grp.
Negative Beta estimate indicates DIF/DBF favoring Foc. grp.

FLAG = error flag indicator. FLAG=0 indicates a normal
successful completion of a SIBTEST run. All other values
of FLAG come with short error messages.

SIBTEST-pooled weighting

Run No.	Suspect Item	Subtest Numbers	Beta estimate	standard error	p-value	p-elim		MS SSD	F L A G
						R	F		G
1	1		0,013	0,011	0,257 E	0,28	0,22	0,26	0
2	2		0,018	0,010	0,095 E	0,27	0,21	0,26	0
3	3		-0,010	0,011	0,374 E	0,27	0,21	0,27	0
4	4		0,015	0,010	0,121 E	0,27	0,21	0,26	0
5	5		-0,065	0,010	0,000 E	0,27	0,21	0,27	0
6	6		-0,028	0,011	0,011 E	0,28	0,21	0,27	0
7	7		0,035	0,010	0,001 E	0,27	0,21	0,26	0
8	8		0,048	0,010	0,000 E	0,26	0,20	0,26	0
9	9		0,065	0,009	0,000 E	0,27	0,22	0,25	0
10	10		-0,019	0,008	0,025 E	0,26	0,20	0,26	0
11	11		-0,082	0,009	0,000 E	0,26	0,21	0,27	0
12	12		0,000	0,010	0,966 E	0,26	0,21	0,26	0
13	13		-0,019	0,011	0,076 E	0,27	0,21	0,27	0
14	14		0,027	0,010	0,006 E	0,26	0,20	0,26	0
15	15		-0,050	0,011	0,000 E	0,28	0,21	0,27	0
16	16		0,055	0,009	0,000 E	0,26	0,20	0,25	0
17	17		0,032	0,010	0,002 E	0,27	0,20	0,26	0
18	18		-0,006	0,010	0,539 E	0,26	0,21	0,26	0
19	19		-0,026	0,010	0,007 E	0,27	0,20	0,27	0
20	20		0,034	0,011	0,002 E	0,27	0,21	0,26	0

**EK-G: Okul Türüne Göre İngilizce Alt Testi Maddelerinin MH Analizi İçin
Yapılan EZDIF Program Çıktısı**

Mantel-Haenszel and Logistic Regression Analysis of
DIFFERENTIAL ITEM FUNCTIONING

Programmed by Niels G. Waller

EZDIF deneme

Reference Group: D:\devlet.prn

Focal Group: D:\ozel.prn

Number of Cases in Reference Group: 5.140

Number of Cases in Focal Group: 3.968

Conditioning Levels

0	5	9	13	16	19
4	8	12	15	18	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	MH D-DIF	SE (MH D-DIF)
A 8***	1,380	27,259	0,000	-0,757	0,146
A 44	1,073	0,984	0,321	-0,166	0,161
A 48***	1,260	12,724	0,000	-0,542	0,151
B 52***	0,608	32,476	0,000	1,168	0,204
A 57***	0,760	16,107	0,000	0,646	0,160
A 84	0,914	1,984	0,159	0,212	0,147
A 95**	0,827	6,769	0,009	0,446	0,168
A 99*	1,187	4,238	0,040	-0,402	0,193
A 106**	0,807	7,352	0,007	0,504	0,186
A 108***	0,692	10,068	0,002	0,864	0,266
A 111***	0,689	16,780	0,000	0,875	0,211
A 115**	0,813	7,555	0,006	0,486	0,174
A 136***	0,807	9,522	0,002	0,503	0,162
A 140***	0,763	10,489	0,001	0,636	0,194
A 173	1,078	1,271	0,260	-0,176	0,152
A 193*	0,782	6,329	0,012	0,578	0,227
A 207	0,916	1,316	0,251	0,206	0,174
A 208	1,029	0,109	0,742	-0,067	0,182
A 209	0,876	3,146	0,076	0,311	0,171
A 227	0,880	3,580	0,058	0,301	0,157

Number of Items Purged in Pass 1: 0

Item Numbers:

**EK-Ğ: Okul Türüne Göre İngilizce Alt Testi Maddelerinin LR Analizi İçin
Yapılan EZDIF Program Çıktısı**

Logistic Regression Procedure.

Note: Items purged from previous step

Madde 1	Item Number: 8	Estimate	SE Estimate	Z	p-value
	Intercept	-3,805439	0,110515	-34,433678	0,000000
	Trait	0,289330	0,006959	41,579002	0,000000
	Group	1,974711	0,218218	9,049276	0,000000
	Trait x Group	-0,118009	0,013917	-8,479378	0,000000
Madde 2	Item Number: 44	Estimate	SE Estimate	Z	p-value
	Intercept	-4,770098	0,138578	-34,421688	0,000000
	Trait	0,381966	0,008920	42,820332	0,000000
	Group	3,170869	0,262032	12,101068	0,000000
	Trait x Group	-0,229877	0,017840	-12,885187	0,000000
Madde 3	Item Number: 48	Estimate	SE Estimate	Z	p-value
	Intercept	-4,444856	0,127867	-34,761608	0,000000
	Trait	0,326049	0,007921	41,162617	0,000000
	Group	3,713378	0,251208	14,782080	0,000000
	Trait x Group	-0,245842	0,015842	-15,518519	0,000000
Madde 4	Item Number: 52	Estimate	SE Estimate	Z	p-value
	Intercept	-2,453301	0,143924	-17,045807	0,000000
	Trait	0,350349	0,010229	34,250076	0,000000
	Group	-0,245178	0,234279	-1,046523	0,295573
	Trait x Group	-0,019530	0,020458	-0,954641	0,339990
Madde 5	Item Number: 57	Estimate	SE Estimate	Z	p-value
	Intercept	-3,922888	0,124443	-31,523584	0,000000
	Trait	0,351911	0,008289	42,453983	0,000000
	Group	1,326985	0,228801	5,799723	0,000000
	Trait x Group	-0,125365	0,016578	-7,561932	0,000000
Madde 6	Item Number: 84	Estimate	SE Estimate	Z	p-value
	Intercept	-4,540526	0,125630	-36,142010	0,000000
	Trait	0,319124	0,007748	41,186466	0,000000
	Group	3,168307	0,248812	12,733722	0,000000
	Trait x Group	-0,223913	0,015497	-14,449116	0,000000
Madde 7	Item Number: 95	Estimate	SE Estimate	Z	p-value
	Intercept	-3,611901	0,132416	-27,276836	0,000000
	Trait	0,356315	0,008954	39,794872	0,000000
	Group	1,719287	0,234164	7,342244	0,000000
	Trait x Group	-0,157247	0,017908	-8,781034	0,000000
Madde 8	Item Number: 99	Estimate	SE Estimate	Z	p-value
	Intercept	-2,846503	0,142169	-20,021898	0,000000
	Trait	0,373822	0,010175	36,738979	0,000000
	Group	0,695773	0,234493	2,967131	0,003078
	Trait x Group	-0,044573	0,020350	-2,190313	0,028732
Madde 9	Item Number:106	Estimate	SE Estimate	Z	p-value
	Intercept	-0,596917	0,110884	-5,383250	0,000000
	Trait	0,193124	0,007545	25,596888	0,000000
	Group	-0,450443	0,202223	-2,227457	0,026139
	Trait x Group	0,029294	0,015090	1,941304	0,052502
Madde 10	Item Number:108	Estimate	SE Estimate	Z	p-value
	Intercept	-2,712909	0,220634	-12,295957	0,000000
	Trait	0,466640	0,016307	28,615704	0,000000
	Group	-0,337618	0,308414	-1,094689	0,273916
	Trait x Group	0,013668	0,032614	0,419067	0,675257

Madde 11	Item Number:111	Estimate	SE Estimate	Z	p-value
	Intercept	-2,985889	0,159227	-18,752386	0,000000
	Trait	0,403621	0,011493	35,117649	0,000000
	Group	-0,171349	0,251959	-0,680069	0,496618
	Trait x Group	-0,012315	0,022987	-0,535748	0,592252
Madde 12	Item Number:115	Estimate	SE Estimate	Z	p-value
	Intercept	-4,261548	0,144647	-29,461634	0,000000
	Trait	0,404234	0,009848	41,049274	0,000000
	Group	1,631610	0,254577	6,409105	0,000000
	Trait x Group	-0,149045	0,019695	-7,567638	0,000000
Madde 13	Item Number:136	Estimate	SE Estimate	Z	p-value
	Intercept	-3,491374	0,118256	-29,523951	0,000000
	Trait	0,335316	0,007983	42,002388	0,000000
	Group	0,745158	0,216078	3,448556	0,000587
	Trait x Group	-0,076607	0,015966	-4,798043	0,000002
Madde 14	Item Number:140	Estimate	SE Estimate	Z	p-value
	Intercept	-2,693901	0,145554	-18,507854	0,000000
	Trait	0,356333	0,010262	34,723434	0,000000
	Group	0,702152	0,236886	2,964096	0,003108
	Trait x Group	-0,088485	0,020524	-4,311306	0,000018
Madde 15	Item Number:173	Estimate	SE Estimate	Z	p-value
	Intercept	-4,709012	0,129375	-36,398136	0,000000
	Trait	0,332155	0,007971	41,669121	0,000000
	Group	3,421875	0,256175	13,357557	0,000000
	Trait x Group	-0,232933	0,015942	-14,610867	0,000000
Madde 16	Item Number:193	Estimate	SE Estimate	Z	p-value
	Intercept	-2,724738	0,172952	-15,754334	0,000000
	Trait	0,418844	0,012696	32,989105	0,000000
	Group	-0,375914	0,264020	-1,423810	0,154814
	Trait x Group	0,024689	0,025393	0,972291	0,331141
Madde 17	Item Number:207	Estimate	SE Estimate	Z	p-value
	Intercept	-3,396661	0,131244	-25,880444	0,000000
	Trait	0,361147	0,009047	39,916828	0,000000
	Group	1,112306	0,228992	4,857413	0,000001
	Trait x Group	-0,099069	0,018095	-5,474925	0,000000
Madde 18	Item Number:208	Estimate	SE Estimate	Z	p-value
	Intercept	-5,108825	0,158622	-32,207561	0,000000
	Trait	0,452148	0,010706	42,231468	0,000000
	Group	2,249403	0,283852	7,924555	0,000000
	Trait x Group	-0,176439	0,021413	-8,239860	0,000000
Madde 19	Item Number:209	Estimate	SE Estimate	Z	p-value
	Intercept	-5,860854	0,159899	-36,653404	0,000000
	Trait	0,441623	0,010214	43,236561	0,000000
	Group	3,670862	0,304169	12,068513	0,000000
	Trait x Group	-0,272594	0,020428	-13,344047	0,000000
Madde 20	Item Number:227	Estimate	SE Estimate	Z	p-value
	Intercept	-3,710690	0,118475	-31,320475	0,000000
	Trait	0,333798	0,007859	42,474457	0,000000
	Group	1,328805	0,220216	6,034092	0,000000
	Trait x Group	-0,113099	0,015718	-7,195717	0,000000

EK-H: Okul Türüne Göre İngilizce Alt Testi Maddelerinin SIBTEST Analizi İçin Yapılan SIBTEST Program Çıktısı

name of input parameter file = sib.in
number of items on test = 20
name of file for Ref. grp. scores = C:\OZELOKUL.prn
name of file for Focal grp. scores = C:\DEVLETOKUL.prn
minimum no. of examinees per matching score cell = 2
number of runs for this data set = 20
number of examinees in Reference group = 3.968
number of examinees in Focal group = 5.140

Examinee Test Score Summary Statistics

Reference Group: Mean = 17,08
Standard deviation = 3,97
Focal Group: Mean = 11,11
Standard deviation = 5,03

Standardized Score Difference = 1,30

Item Statistics

= item number
p = proportion right on item (Classical Test Theory p-value)
r = point biserial (item score-test score correlation)

#:	1	2	3	4	5	6	7	8	9	10
p:	0,565	0,604	0,564	0,790	0,640	0,523	0,690	0,795	0,835	0,857
r:	0,567	0,656	0,582	0,575	0,655	0,598	0,627	0,565	0,372	0,539

#:	11	12	13	14	15	16	17	18	19	20
p:	0,789	0,674	0,667	0,784	0,528	0,827	0,719	0,649	0,563	0,644
r:	0,604	0,677	0,637	0,571	0,603	0,567	0,621	0,710	0,709	0,635

p-value notation:

R denotes p-value for test of DIF/DBF against Ref. group
F denotes p-value for test of DIF/DBF against Foc. group
E denotes p-value for test of DIF/DBF against either the Ref. or Foc. group.

NOTES:

MS/SSD = Matching Subtest Standardized Score Difference.
Standardized difference in mean observed scores
between Reference group and Focal group on the
matching subtest.

p-elim = proportion of Reference (R) and Focal (F) groups
eliminated (not used) in SIBTEST calculations.

Positive Beta estimate indicates DIF/DBF favoring Ref. grp.
Negative Beta estimate indicates DIF/DBF favoring Foc. grp.

FLAG = error flag indicator. FLAG=0 indicates a normal
successful completion of a SIBTEST run. All other values
of FLAG come with short error messages.

SIBTEST-pooled weighting

Run No.	Suspect Item	Subtest Numbers	Beta estimate	standard error	p-value	p-elim		MS SSD	F L A G
						R	F		
1	1		-0,057	0,015	0,000 E	0,47	0,09	1,30	0
2	2		-0,031	0,014	0,021 E	0,44	0,09	1,29	0
3	3		-0,038	0,013	0,004 E	0,45	0,08	1,29	0
4	4		0,067	0,013	0,000 E	0,44	0,09	1,29	0
5	5		0,053	0,015	0,000 E	0,44	0,09	1,29	0
6	6		0,008	0,015	0,582 E	0,45	0,09	1,28	0
7	7		0,009	0,015	0,517 E	0,44	0,09	1,29	0
8	8		-0,013	0,014	0,347 E	0,43	0,09	1,30	0
9	9		0,048	0,012	0,000 E	0,45	0,10	1,29	0
10	10		0,002	0,013	0,876 E	0,44	0,09	1,30	0
11	11		0,040	0,013	0,003 E	0,44	0,09	1,30	0
12	12		0,019	0,014	0,184 E	0,44	0,09	1,29	0
13	13		0,051	0,015	0,001 E	0,44	0,09	1,29	0
14	14		0,033	0,014	0,015 E	0,44	0,09	1,29	0
15	15		-0,016	0,015	0,279 E	0,45	0,09	1,29	0
16	16		0,013	0,013	0,301 E	0,43	0,09	1,30	0
17	17		-0,022	0,014	0,101 E	0,44	0,09	1,30	0
18	18		-0,025	0,013	0,058 E	0,44	0,09	1,30	0
19	19		-0,002	0,014	0,862 E	0,44	0,09	1,29	0
20	20		0,018	0,016	0,260 E	0,44	0,09	1,29	0