



**UZMAN GÖRÜŞÜNE GÖRE YANLI OLDUĞU BELİRLENEN
MADDELERİN DEĞİŞEN MADDE FONKSİYONU BELİRLEME
YÖNTEMLERİ İLE TUTARLILIĞININ İNCELENMESİ**

Nurcan Ceylan

YÜKSEK LİSANS TEZİ

EĞİTİMDE ÖLÇME DEĞERLENDİRME ANABİLİM DALI

**GAZİ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

EKİM, 2022

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU

Bu tezin tüm hakları saklıdır. Kaynak göstermek koşuluyla tezin teslim tarihinden itibaren 12 (on iki) ay sonra tezden fotokopi çekilebilir.

YAZARIN

Adı : Nurcan

Soyadı : Ceylan

Bölümü : Eğitim Bilimleri

İmza :

Teslim Tarihi :

TEZİN

Türkçe Adı : Uzman Görüşüne Göre Yanlı Olduğu Belirlenen Maddelerin Değişen Madde Fonksiyonu Belirleme Yöntemleri İle Tutarlılığının İncelenmesi

İngilizce Adı : Examining The Consistency Of The Items Determined To Be Biased According To Expert Opinions With Differential Item Functioning

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik ilkelere uyduđumu, yararlandıđım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttiđimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduđunu beyan ederim.

Yazarın Adı Soyadı : Nurcan Ceylan

İmza :

JÜRİ ONAY SAYFASI

Nurcan Ceylan tarafından hazırlanan " Uzman Görüşüne Göre Yanlı Olduğu Belirlenen Maddelerin Değişen Madde Fonksiyonu Belirleme Yöntemleri İle Tutarlılığının İncelenmesi " adlı tez çalışması aşağıdaki jüri tarafından oy birliği ile Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda Yüksek Lisans tezi olarak kabul edilmiştir.

Danışman : Prof. Dr. Bayram ÇETİN

(Gaziantep Eğitim Fakültesi, Gaziantep Üniversitesi)

Başkan : Prof. Dr. Nuri Doğan

(Eğitimde Ölçme Ve Değerlendirme Ana Bilim Dalı, Hacettepe Üniversitesi)

Üye : Prof. Dr. Hakan Yavuz Atar

(Eğitimde Ölçme Ve Değerlendirme Ana Bilim Dalı, Gazi Üniversitesi)

Tez Savunma Tarihi : 14.09.2022

Bu tezin Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda Yüksek Lisans tezi olması için şartları yerine getirdiğini onaylıyorum.

Prof. Dr. Şaban Çetin

Eğitim Bilimleri Enstitüsü Müdürü



Canım Aileme

TEŞEKKÜR

Tüm yoğunluğuna rağmen her danıştığım beni asla reddetmeyen ve yardımcı olan, öğrencisi olmaktan gurur duyduğum kıymetli hocam Prof. Dr. Bayram ÇETİN' e, çalışmam ve ders dönemim süresince bana olan katkısı ve yol göstericiliğinden dolayı teşekkürlerimi sunuyorum.

Tez savunma jürisinde bulunmasından onur duyduğum kıymetli hocam Prof. Dr. Nuri Doğan'a ve hem jürimde bulunarak beni onurlandıran hem de ders dönemimde gelişime büyük katkı sağlayan Prof. Dr. Hakan Yavuz Atar'a katkılarından dolayı çok teşekkür ederim.

Yüksek lisans eğitimim boyunca ders aldığım Prof. Dr. Mehtap Çakan, Prof. Dr. İsmail Karakaya, Prof. Dr. Dilara Bakan Kalaycıoğlu hocalarıma ve lisans eğitimimde ve sonraki süreçte bana kattıklarından dolayı canım hocam Hatice Gonca Usta'ya bana kattıkları her şey için çok teşekkür ederim.

Parçası olmaktan mutluluk duyduğum Sivas Ölçme Değerlendirme Merkezindeki, müdürüm Banu Boğazlıyan, müdür yardımcım Mesut Çelik ve ekip üyesi arkadaşlarıma yüksek lisansa başlama sürecimde beni teşvik etmelerinden ve süreçte bana destek vermelerinden dolayı teşekkür ederim, onlar olmasa yüksek lisans sürecine başlama cesareti ve bitirme azmi gösteremezdim.

Yüksek lisans eğitimi sürecinde en büyük kazanımım olan canım arkadaşlarım Ebru Verim Kocalar, Naime Sözen ve Ömer Topraktepe'ye desteklerinden dolayı teşekkür ederim.

Hayatım boyunca yaptığım her işte bana güvenlerini sonuna kadar hissettiren ve verdikleri bu destekle hayata karşı dik durabilmemi sağlayan, varlıklarının bedelini asla ödeyemeyeceğim babam Mustafa Çetin, annem Fatma Çetin ve kardeşlerim Gülden ve Kubilay Çetin'e varlıklarından ve hayatımda olmalarından dolayı teşekkür ederim. İyi ki benim ailemsiniz.

Ve en büyük teşekkürüm ise bu süreçte başladığımda hayatıma dâhil olan, yardımını ve desteğini asla esirgemeyen, her umutsuzluğa düştüğümde beni ayağa kaldıran, hayattaki en büyük şansım canım eşim Yasin Ceylan'a, desteğinden ve hayatımda olduğundan dolayı çok teşekkür ediyorum.



**UZMAN GÖRÜŞÜNE GÖRE YANLI OLDUĞU BELİRLENEN
MADDELERİN DEĞİŞEN MADDE FONKSİYONU BELİRLEME
YÖNTEMLERİ İLE TUTARLILIĞININ İNCELENMESİ**

(Yüksek Lisans Tezi)

Nurcan Ceylan

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Ekim, 2022

ÖZ

Araştırmada, yerleşim bölgesine göre (kırsal ve kent merkezi) yanlı olarak yazılmış ve uzmanlarca da yanlılığı onaylanmış maddeleri bulunan bir testte, DMF belirleme yöntemlerinden Mantel-Haenszel (MH), Lojistik Regresyon (LR) ve Simultaneous Item Bias Test (SIBTEST) yöntemlerinin belirtilen maddeleri tespit etme tutarlılıkları incelenmiştir. Analizler için kullanılan veriler kırsal ve kent merkezinde bulunan 1722 tane 5.sınıf öğrencisinin yanıt verdiği 20 soruluk başarı testinden elde edilmiştir. DMF analizi öncesinde testin tek boyutluluğu incelenmiş ve teste ait betimsel istatistikler, güvenilirlik analizleri ve madde istatistikleri hesaplanmıştır. Araştırmada madde yanlılığı süreci ters yönde işletilmiş ve test maddelerinin tamamı öncelikle uzman görüşüne gönderilmiştir. Uzmanlarca testin yanlı yazılan dört maddesinin yanlı madde olduğu onaylanmış ve daha sonra testin DMF analizleri R programında “difR” paketi ile yapılarak uzman görüşüncü yanlı olduğu onaylanan maddeler ve DMF analizleri karşılaştırılmıştır. Çalışmanın

sonucunda uzmanların yanlı olduğunu belirttiđi dört maddede, LR yönteminin maddelerin birini tespit edebildiđi, MH yönteminin maddelerden iki tanesini, SIBTEST yönteminin ise maddelerden üç tanesini tespit edebildiđi görölmüştür. Ayrıca uzmanlarca yanlı olarak görölmeyen maddelerde de MH ve SIBTEST yöntemlerinin B ve C düzeyinde DMF tespit ettiđi ortaya çıkmıştır.



Anahtar Kelimeler : Deđişen madde fonksiyonu, madde yanlılıđı , Mantel-Haenszel, Lojistik Regresyon, SIBTEST

Sayfa Adeti : xvi+85

Danışman : Prof. Dr. Bayram ÇETİN

**EXAMINING THE CONSISTENCY OF THE ITEMS DETERMINED
TO BE BIASED ACCORDING TO EXPERT OPINIONS WITH
DIFFERENTIAL ITEM FUNCTIONING**

(M.S. Thesis)

Nurcan Ceylan

GAZI UNIVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

October, 2022

ABSTRACT

In the study, the consistency of detecting the specified items of the DIF determination methods Mantel-Haenszel (MH), Logistic Regression (LR) and Simultaneous Item Bias Test (SIBTEST) was examined in a test that had items biased according to the settlement area and approved by experts. The data used for the analyzes were obtained from the achievement test of 20 questions answered by 1722 5th grade students in country and city centers. Before the DIF analysis, the unidimensionality of the test was examined and the descriptive statistics, reliability analyzes and item statistics of the test were calculated. In the research, the item bias process was operated in the opposite direction and all of the test items were first sent to the expert opinion. It was approved by the experts that the four incorrectly written items of the test were biased, and then the DIF analyzes of the test were performed with the “difR” package in the R program, and the items that were confirmed to be biased by the expert opinion and the DIF analyzes were compared. As a result of the study, it was seen that LR method could detect one of the four items, which experts stated to be biased, the MH method could detect two of the items, and the SIBTEST method could detect three of the

items. In addition, it was revealed that the MH and SIBTEST methods detected DIF at B and C levels in items that were not seen as biased by experts.



Key Words : Differential item functioning, item bias, Mantel-Haenszel, Logistic Regression, SIBTEST

Page Number : xvi+85

Supervisor : Prof. Dr. Bayram Çetin

İÇİNDEKİLER

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU	i
ETİK İLKELERE UYGUNLUK BEYANI	ii
JÜRİ ONAY SAYFASI.....	iii
TEŞEKKÜR.....	v
ÖZ	vii
ABSTRACT	ix
İÇİNDEKİLER.....	xi
TABLolar LİSTESİ.....	xiv
ŞEKİLLER LİSTESİ.....	xvi
BÖLÜM I	1
GİRİŞ.....	1
1.1. Problem Durumu	1
1.2. Araştırmanın Amacı	5
1.3. Araştırma Problemi	5
1.4. Araştırmanın Önemi.....	5
1.5. Sayıtlar	8
1.6. Sınırlılıklar	8
BÖLÜM II.....	9
KURAMSAL ÇERÇEVE	9

2.1. Değişen Madde Fonksiyonu	9
2.2. Değişen Madde Fonksiyonu Belirlemede Kullanılan Yöntemler	12
2.2.1. Mantel-Haenszel Yöntemi	14
2.2.2. Lojistik Regresyon Yöntemi.....	16
2.2.3. Simultaneous Item Bias Test (SIBTEST).....	19
2.3. Madde Yanlılığı	21
2.4. DMF Kaynakları	23
2.5. DMF Çalışmalarının Gelişimi.....	25
2.6. İlgili Araştırmalar	25
BÖLÜM III	29
YÖNTEM.....	29
3.1. Araştırmanın Modeli	29
3.2. Çalışma Grubu	30
3.3. Veri Toplama Araçları	31
3.3.1. Başarı Testi	32
3.3.2. Yanlılık Anketi.....	33
3.4. Verilerin Elde Edilmesi ve Analizi	34
3.4.1. Taslak Testin Verilerinin Elde Edilmesi	34
3.4.2. Taslak Testin Analizi	35
3.4.3. Nihai Testin Verilerinin Elde Edilmesi	38
3.4.4. Uzmanlarca Yanlı Olduğu Düşünülen Maddelerin Belirlenmesi	39
3.4.5. Nihai Testin Analizi.....	41
BÖLÜM IV	49
BULGULAR VE YORUM	49
4.1. Araştırma Problemi 1'e Yönelik Bulgular.....	49

4.2. Araştırma Problemi 2'ye Yönelik Bulgular.....	51
4.3. Araştırma Problemi 3'e Yönelik Bulgular.....	53
4.4. Araştırma Problemi 4'e Yönelik Bulgular.....	55
BÖLÜM V	58
SONUÇ, TARTIŞMA VE ÖNERİLER	58
5.1. Sonuç ve Tartışma.....	58
5.2. Öneriler	62
5.2.3. Araştırma Sonuçlarıyla İlgili Öneriler	62
5.2.3. Araştırmacılara Yönelik Öneriler	62
KAYNAKLAR.....	64
EKLER	74
Ek 1. Veri Elde Edilmesinde Kullanılan Başarı Testi	75
Ek 2. Madde Yanlılığı Anketi	79
Ek 3. Gazi Üniversitesi Etik Komisyonu Görüşü ve Onayı	81
Ek 4. Sivas Milli Eğitim Müdürlüğü Onayı.....	82
Ek 5. Gönüllü Katılım Formu.....	83
Ek 6. Tez Benzerlik Raporu	85

TABLÖLAR LİSTESİ

Tablo 1. <i>KTK ve MTK'ya Dayalı DMF Belirleme Yöntemleri</i>	13
Tablo 2. <i>Referans ve Odak Grubun Cevap Dağılımı</i>	15
Tablo 3. <i>MH Yönteminde Etki Büyüklüğünün Yorumlama Ölçüleri</i>	16
Tablo 4. <i>ΔR^2 Değeri Yorumlanması için Kullanılabilecek Sınıflamalar</i>	18
Tablo 5. <i>SIBTEST Yönteminin Değerlendirilmesinde Önerilen Sınıflama Ölçütleri</i>	21
Tablo 6. <i>Cinsiyet ve Yerleşim Bölgelerine Göre Dağılım (Taslak Test)</i>	30
Tablo 7. <i>Cinsiyet ve Yerleşim Bölgelerine Göre Dağılım (Asıl Uygulama)</i>	31
Tablo 8. <i>Taslak Testin Madde Analizleri</i>	36
Tablo 9. <i>Taslak Testin Betimsel İstatistikleri</i>	37
Tablo 10. <i>Nihai Test Uygulanan Okullar ve Elde Edilen Veri Sayısı</i>	38
Tablo 11. <i>Uzmanlarca Yanlı Olduğu Düşünülen Maddeler</i>	39
Tablo 12. <i>Nihai Testin Madde Analizleri</i>	42
Tablo 13. <i>Nihai Testin Betimsel İstatistikleri</i>	43
Tablo 14. <i>Nihai Testin Betimsel İstatistikleri (Kırsal)</i>	44
Tablo 15. <i>Nihai Testin Betimsel İstatistikleri (Kent Merkezi)</i>	44
Tablo 16. <i>Kaiser- Meyer-Olkin (KMO) Testi Sonucu</i>	45
Tablo 17. <i>Maddelerin Faktörlerdeki Yük Değerleri</i>	46
Tablo 18. <i>DFA Model Veri Uyumu İstatistikleri</i>	48
Tablo 19. <i>MH Yöntemi İle DMF Analizi Sonuçları</i>	50
Tablo 20. <i>LR Analizi β Değerleri ve DMF Türü</i>	51
Tablo 21. <i>LR Analiz Sonuçları ve DMF Düzeyleri</i>	52

Tablo 22. <i>SIBTEST Yöntemi İle DMF Analiz Sonuçları</i>	53
Tablo 23. <i>SIBTEST Yöntemi DMF Düzeyleri</i>	54
Tablo 24. <i>Uzman Görüşleri ve Yöntemlerce Bulunan DMF Düzeyleri</i>	56



ŞEKİLLER LİSTESİ

Şekil 1. Tek biçimli değişen madde fonksiyonu.	11
Şekil 2. Tek biçimli olmayan değişen madde fonksiyonu.....	11
Şekil 3. Yanlı madde örneği	22
Şekil 4. Yamaç birikinti grafiği.....	47

BÖLÜM I

GİRİŞ

Bu bölümde araştırma problemi, araştırmanın amacı, araştırma problemi, araştırmanın önemi, sayıtlar ve sınırlılıklara yer verilmiştir.

1.1. Problem Durumu

Sistem, bir şeye ulaşmak amacıyla tutarlı biçimde organize olmuş bağlantılı bileşenler anlamına gelir (Meadows, 2008). Bir yapının sistem olarak nitelendirilebilmesi için bir amacı olmalı ve birbiriyle etkileşimi olan parçaları bulunmalıdır. Sistemlerde bulunan bu parçalar karşımıza girdiler, süreç, çıktılar ve geri bildirimler olarak çıkar (Tecim, 2004).

Eğitim kavramı genellikle ‘insan davranışlarını değiştiren bir sistem’ olarak tanımlanır (Özkan, 2019). Görüldüğü üzere eğitim bir ‘sistem’ olarak tanımlanır ve bu yüzden eğitimin de diğer sistemler gibi amacı ve kendine ait girdileri, çıktısı, süreci ve geribildirimi vardır. Eğitim sistemindeki temel amaç bireylere önceden tanımlanmış belli davranışların kazandırılmasıdır. Eğitim sürecinin ne derece etkili olduğunun ve bu davranışların ne düzeyde kazandırılabilirdiğinin kontrolünün yapılması eğitim sisteminin amacına ulaşabilmesi açısından önemlidir (Suna, 2012). Bu işleyişin kontrolü eğitimde değerlendirme vasıtası ile yapılır (Özkan, 2019). Eğitim sisteminde değerlendirme sistemin

kontrol mekanizması olarak kullanılır. Değerlendirme ile sistemin tüm öğelerinin işbirliğinin kontrolü ve eksiklerin belirlenmesi sağlanır (Tan, 2020). Değerlendirme eğitim sistemi içindeki öğelerin iyileşme durumlarını da ortaya koyar (Baykul, 2015). Değerlendirme sayesinde aksayan ve eksik yönlerin tespiti yapılarak, sistemin kendini kontrol etmesi ve eksiklerini gidermesi yani ‘geri bildirim’ aşaması sağlanır (Tan, 2020).

Eğitimde, öğretim metotlarının etkinliğini belirleme, öğrenme güçlüklerini tespit etme, bireylerin başarısını saptama, eğitim programının sağlamlığını anlama gibi amaçlarla yapılan bu değerlendirmelerin tamamı bir ölçme sonucuna dayanır (Baykul, 2015). Ölçme, eğitim sisteminin varlığını devam ettirebilmesi için en temel öğelerden biridir (Özkan, 2019).

Eğitim, sosyal bilimlerin diğer dallarında da olduğu gibi operasyonel olarak tanımlaması zor olan değişkenlerden oluşur. Eğitimde yapılan ölçmeler bu yüzden zorlaşır ve ölçmelerde kolaylık sağlanması adına ölçme araçları kullanılır (Atılğan, 2014). Bu ölçme araçlarıyla elde edilen puanlar belirli kararlara imkân tanıdığı için elde edilen puanların hatasız olması beklenir. Bu beklentinin karşılanabilmesi için ise ölçme araçlarının sahip olması gereken yapısal özellikler vardır. Geçerlik, bu özelliklerin en önemlilerinden birisidir (Baykul, 2010). Tekin (1979)’in “Bir ölçme aracının ölçmeyi amaçladığı özelliği, başka herhangi bir özellik karıştırmadan doğru biçimde ölçebilmesi” tanımıyla karşımıza çıkan geçerliğin sağlanabilmesi için testin ölçmek istenilen özelliği ölçmede kullanılması ve elde ettiğimiz ölçümün ilgili amaç doğrultusunda yorumlanması gerekir (Akbaş, 2019). Yani geçerlik, testlerin önerildiği kullanım alanları doğrultusunda elde edilmiş puanların yorumlanabilirliğini sağlayan teori ve kanıtların derecesidir (Crocker ve Algina, 1986). Bu kanıtların sağlanması yani testin geçerliğinin yüksek olması ise, test doğrultusunda bireylere ilişkin alınan kararların doğruluğunu desteklemek açısından önemli görülmektedir (Atılğan, 2014).

Ölçme yaparken mümkün olduğunca hatasız ölçme yapılmaya çalışılır ancak eğitimde ölçme türünden kaynaklı olarak ölçmeye farklı hata türleri karışması kaçınılmazdır (Yurdugül,

2003). Ölçmede yaptığımız ölçümlere karışacak bu hatalar sabit, sistematik ve tesadüfi hatalar olarak karşımıza çıkar. Sabit hatalar her ölçüme aynı miktar ve yönde karışırken, tesadüfi hatalar kaynak ve yönleri kesin bilinmeyen hatalardır (Akbaş, 2019). Sistematik hata ise ölçülen yapı ile alakası olmadan gruplar arasında ortaya çıkan tutarlı farklılıklar olarak nitelendirilebilir (Yıldırım, 2015). Eğitimde ölçmeye karışan bu sistematik hatalar genellikle ‘yanlılık’ olarak adlandırılır (Yurdugül, 2003).

Zumbo (1999), madde yanlılığı ifadesini test koşulları veya maddenin farklı özelliklerinden dolayı bazı grupların maddeleri cevaplama olasılıklarının diğer gruplara kıyasla fazla ya da az olması olarak ifade etmiştir (Zumbo’dan aktaran Kelecioğlu, 2014). Bir maddeyi doğru yanıtlama olasılığının yetenek düzeyine bağlı olması beklenir ancak yanlı olan maddelerde bu olasılık öğrencinin bulunduğu gruba bağlıdır (Osterlind, 1983). Ackerman (1991)’a göre araştırmacılar madde yanlılığının ana nedenini gizil yeteneklerin yanlış belirlenmesi olarak öne sürer ancak yanlılık maddenin çok boyutlu yetenek dağılımına sahip iki farklı katılımcı grubun yetenek seviyeleri arasında ayrım yapmasıdır. Yanlılık, yapılan ölçmenin hata içerdiğini gösterir. Ölçümümüz bireylerin bulundukları gruba (farklı cinsiyet, farklı okul türü, farklı yerleşim yeri vb.) göre yarar sağlamıştır denilebilir.

Angoff (1993)’a göre yanlılık istatistiksel ve sosyal olmak üzere iki farklı anlamda kullanılmaktadır (Angoff’tan aktaran Ayva Yörü, 2019). Çünkü madde yanlılığı hem istatistiksel hem de uzman kanılarına dayalı süreçleri gerektiren iki farklı süreçten oluşur (Kelecioğlu & Karabay, 2014).

Madde yanlılığı belirlemenin ilk aşamasında değişen madde fonksiyonu (DMF) analizi yapılır. DMF çalışmaları madde yanlılığının istatistiksel bir göstergesidir. DMF çalışmaları sayesinde testi alan grupların madde yanıt örüntüleri arasında bulunan farklılıklar belirlenerek istatistiksel açıdan yanlı maddelerin belirlenmesi sağlanır (Zumbo & Gelin, 2005). İkinci aşamada ise DMF belirlenen maddelerdeki ortaya çıkan farklılıklarının yetenek farkından mı grup üyeliklerinden mi kaynaklandığı belirlenir yani DMF ile ortaya çıkan farklılığın kaynağı uzmanlardan yardım alınarak araştırılır (Karami, 2012). DMF gösteren

maddelerin yanlı olup olmadığını belirlemek için en sık kullanılan yöntem anket yoluyla bu maddeler için uzman görüşü almak ve sonrasında bu maddeler hakkında yorumda bulunmaktır (Ateşok Deveci, 2008; Çepni, 2011; Kalaycıoğlu, 2008; Karakaya & Kutlu, 2012).

Millsap (2005)'a göre geçerli kabul edilebilecek bir ölçme aracından elde edilen puanların, bireyin bağlı bulunduğu alt gruplardan bağımsız olduğu kabul edilir ve bu ölçme aracını yanıtlayan grup üyelerinin her biri için değişmez olması beklenir. Eğer ölçme aracının ölçtüğü özellikler farklı gruplar için değişmez değilse bu ölçme aracı sonucunda verilen kararların geçerliliği ve ölçme aracına göre yapılan yorumlamalar ve verilen kararlar sorgulanabilir (Uzun & Gelbal, 2017).

Madde yanlılığı bir geçerlik problemidir. Eğer bir testte yanlı madde varsa testin geçerliği ve dolayısıyla da güvenirliği olumsuz etkilenir. Bu durum da testin tarafsızlığını tehdit eder (Kristanjansson, Aylesworth, McDowell ve Zumbo, 2005). Güvenirliği ve geçerliği düşük bir testin sonucuna göre yapılan seçme ve yerleştirmeler ve bu test sonucuna göre alınan kararlar gerçeği yansıtamaz (Bakan Kalaycıoğlu, 2008).

Ülkemizde sınav sonuçlarına diğer ülkelerde olduğundan daha fazla itibar edilmektedir. Bu yüzden de ölçme araçlarının geliştirilmesinde daha titiz davranılması gerekir (Çetin, 2019). Bu nedenle özellikle son yıllarda madde yanlılığı ve değişen madde fonksiyonu (DMF) konularının araştırılmasında büyük bir artış görülmüştür (Uzun & Gelbal, 2017).

Alan yazında DMF belirlemek için birden fazla yöntem vardır. Bu yöntemler eşleştirme için farklı kriterler kullandıklarından tam bir uyum içinde değildir. Bu nedenle yöntemlerin birbirlerine göre avantaj ve dezavantajları göz önünde bulundurularak birden fazla yöntemin kullanılması önerilmektedir (Holland & Wainer, 1993).

Ancak çok sayıda DMF belirleme yönteminin hangisinin daha işlevsel ve doğru sonuç verdiğine ilişkin araştırmalar henüz kesin bir sonuç göstermemekte ve 'hangi yöntem?' sorusuna henüz net bir cevap alınamamaktadır.

1.2. Araştırmanın Amacı

Bu araştırmanın amacı; yerleşim bölgesine göre yanlı olarak yazılmış ve uzmanlarca da yanlılığı onaylanmış maddeleri bulunan bir testte, DMF belirleme yöntemlerinden Mantel-Haenszel (MH), Lojistik Regresyon (LR) ve Eş Zamanlı Madde Yanlılığı Testi (SIBTEST) yöntemlerinin belirtilen maddeleri tespit etme tutarlılıklarını karşılaştırmaktır.

1.3. Araştırma Problemi

Araştırma amacı doğrultusunda;

1. Yerleşim bölgesine göre yanlı olarak yazılmış ve uzman görüşü ile yanlılığı doğrulanmış maddeleri bulunan testteki yanlı maddeler ile Mantel-Haenszel yöntemine göre DMF gösteren maddeler tutarlı mıdır?
 2. Yerleşim bölgesine göre yanlı olarak yazılmış ve uzman görüşü ile yanlılığı doğrulanmış maddeleri bulunan testteki yanlı maddeler ile Lojistik Regresyon yöntemine göre Tek Biçimli ve Tek Biçimli Olmayan DMF gösteren maddeler tutarlı mıdır?
 3. Yerleşim bölgesine göre yanlı olarak yazılmış ve uzman görüşü ile yanlılığı doğrulanmış maddeleri bulunan testteki yanlı maddeler ile SIBTEST yöntemine göre Tek Biçimli ve Tek Biçimli Olmayan DMF gösteren maddeler tutarlı mıdır?
 4. Yerleşim bölgesine göre yanlı olarak yazılmış ve uzman görüşü ile yanlılığı doğrulanmış maddeleri bulunan testteki yanlı maddeler ile MH, LR ve SIBTEST yöntemleri karşılaştırıldığında hangi yöntem daha tutarlı sonuç vermektedir?
- sorularına cevap aranacaktır.

1.4. Araştırmanın Önemi

Geçerlik, bir ölçme aracında bulunması gereken en önemli özelliklerden biridir (Baykul, 2000; French & Finch, 2013; Li & Zumbo, 2009; Thissen vd., 1988). Seçme ve yerleştirme amacıyla kullanılan sınavlar bireyler hakkında önemli kararlar almaya neden olduğu için bu

sınavların geçerliğı çok 6nemlidir. Test maddeleri eęer test ile 6l6len yapıyla alakalı olmadan alt gruplardan herhangi birine avantaj saęlıyorsa bu testin geçerliğı olumsuz y6nde etkilenir ve test puanları ile alınan kararlar sorgulanabilir. Bu y6zden test maddelerinin bireylerin herhangi bir 6zellięine g6re yanlılık ierip iermedięine iliřkin incelemelerin yapılması ve kanıtların sunulması gerekir.

Yanlılık arařtırma s6releri incelendięinde, yanlılık iin ilk adım olan DMF analizleri yapıldıktan sonra bu maddelerin yanlılık durumlarının incelenmesi iin DMF tespit edilen maddeler uzman g6r6řlerine sunularak maddelerin belirli bir gruba g6re yanlılık ierip iermedięi sorulmuřtur. Bu s6recin ortaya ıkarabileceęi iki olumsuz durum olduęu d6ř6n6lmektedir. Bunlardan ilki DMF olduęu d6ř6n6len maddelerin uzman g6r6ř6ne sunulması sonucunda uzmanların bu maddeler hakkında incelemelerde bulunurken maddelere yanlı davranması, yani maddelerin analizlerine g6re beklenen cevapları verme eęiliminde olarak madde etkisini g6z ardı etme ihtimalleridir. Geniř 6lekli sınavlar 6zerinde yapılan yanlılık alıřmaları incelendięinde uzmanların oęu zaman maddelerin yanlı olma ihtimaline karřı olası kaynaklara deęindięi g6r6lmektedir. DMF g6sterdięi belirlenmiř maddelere genellikle kesin olarak yanlı madde denilmemiřtir ancak maddelerde ortaya ıkan DMF nedeniyle bu farklılıęa neden olabilecek olası ihtimaller 6zerinde durulmuřtur (Yıldırım, 2015). Bu durumun da bahsedilen olumsuzluęun bir parası olduęu d6ř6n6lmektedir. Ortaya ıkabilecek ikinci olumsuz durum ise uzmanlara g6re yanlı olduęu d6ř6n6lebilecek ancak analizlerde eřitli nedenlerden dolayı (6ęrencilerin kopya ekmesi, 6ęrencilere yardım edilmesi vb) DMF varlıęı tespit edilememiř maddelerin g6z ardı edilmesi ve bu maddelerin testin geçerlięini tehdit etmeye devam etmesidir.

Alan yazında yapılan alıřmalar incelendięinde, alıřmaların biroęunda genellikle ulusal ve uluslararası geniř 6lekli sınavların belirli alt testlerine y6nelik yanlılık alıřmaları yapıldıęı ve yanlı olduęu belirlenen maddelerin ortaya ıkarıldıęı g6r6lmektedir. alıřmalar genellikle uygulanmıř bir sınavın maddelerine ait DMF belirlemeye y6nelik analizlerin yapılması ve bu analizlerin uzman g6r6ř6ne sunulması 6zerinedir. oęu zamansa

bahsedildiği gibi DMF gösteren maddelerde net bir yanlılık çıkmamış, maddelerde DMF çıkmasına neden olabilecek olası sebepler üzerinde durulmuştur. DMF belirlemede kullanılan istatistiksel yöntemleri karşılaştırmak amacıyla yapılan çalışmalar bulunmaktadır (Akın Arıkan, Uğurlu & Atar, 2015; Atalay Kabasakal, Arsan, Gök & Kelecioğlu, 2014; Awuor, 2008; Cohen, Kim & Wollack, 1996; Dainis, 2008; Demir & Köse, 2014; Doğan, Hambleton, Yurtçu & Yavuz, 2018; Finch & French, 2007; Jodoin & Gierl, 2001; Erdem Keklik, 2012; Erdem Keklik, 2015; Kim, 2010; Özdemir, 2015; Özmen, 2013; Tunç, 2016;). Ancak bu çalışmalar genellikle geniş ölçekli sınavlarda bulunan madde yanlılıklarını iki farklı yöntem üzerinden karşılaştırmaya yöneliktir ve çalışmalar incelendiğinde her yöntemin maddelerdeki DMF'yi tespit etme durumlarında farklar olduğu göze çarpmıştır. Doğan vd. (2018), yaptıkları çalışmada diğer çalışmalardan farklı olarak bu çalışmanın konusuna benzer şekilde uzman görüşleri ve istatistiksel yöntemlerin karşılaştırılmasını esas almıştır.

Bu çalışmada, içerisinde yerleşim bölgesine göre yanlı maddelerin bulunduğu bir çoktan seçmeli testte madde yanlılığı süreci ters yönde işletilmiştir. Süreçte öncelikle Sosyal Bilgiler dersinden başarı testi geliştirilmiş ve 20 sorudan oluşan bu teste yerleşim bölgesine göre yanlı olacak şekilde 4 madde yazılmıştır. Daha sonra maddelerin tamamı uzman görüşüne sunulmuş ve uzmanlar da yanlı olduğu düşünülen bu 4 maddenin yerleşim bölgesine göre yanlı olduğunu onaylamıştır. Sonrasında ise çalışmanın istatistiksel yönü devreye girmiş ve uzmanların yanlı olduğunu düşündükleri maddeler ile istatistiksel olarak DMF içeren maddeler Mantel-Haenszel, Lojistik Regresyon ve SIBTEST yöntemleri ile incelenmiştir. Böylece uzman görüşleriyle en tutarlı sonuç veren yöntemi belirlemek amaçlanmıştır. Bu nedenle bu araştırmanın sonuçlarının madde yanlılığı tespitinde Mantel-Haenszel, Lojistik Regresyon ve SIBTEST yöntemlerinden hangisinin işe koşulmasının uzmanlarla daha tutarlı sonuçlar vereceğinin bilinmesinde yarar sağlayacağı düşünülmektedir. Alan yazında yapılan çalışmalar iki türdür: Birinci tür ve en yaygın olan çalışma türü uygulanmış ve verileri elde edilmiş büyük ölçekli sınavlarda yanlı madde

bulunup bulunmadığına yönelik çalışmalardır. İkinci tür çalışma grubu ise yöntem karşılaştırma çalışmalarıdır ve genellikle ya büyük ölçekli sınavların verileri üzerinde ya da simülasyon verileri ile yürütülmüştür. Uzman görüşüne dayalı kestirimler ve istatistiksel kestirimlerin karşılaştırılmasına dayalı çalışmalar ise oldukça az sayıdadır (Doğan vd., 2018; Gierl, Rogers & Klinger, 1999; Roth, Oliveri, Sandilands, Lyons-Thomas ve Ercikan, 2013). Bu çalışmada amaç yerleşim bölgesine göre özellikle yanlı olarak yazılmış maddeleri bulunan bir testte hangi yöntemin uzmanlarca yanlı olduğuna karar verilen maddeleri tespit etmede daha işlevsel olduğunu tespit etmektir. Çalışmada olduğu gibi yanlılık sürecinin ters yönde yürütüldüğü, öncelikle uzman görüşlerine bağlı olarak madde yanlılıklarının incelendiği çalışmalar mevcuttur ancak özellikle yanlı maddeler üretilerek, bu yanlı maddeleri içeren bir testin uygulandığı ve çalışmaların bu test sonucuna dayalı olarak yürütüldüğü herhangi bir çalışmaya rastlanmamıştır.

1.5. Sayıtlar

1. Öğrencilerin bu teste verdiği cevaplar gerçek bilgi düzeylerini yansıtmaktadır.
2. Araştırmada öğrenciler testteki tüm maddeleri ciddiyetle ve duyarlılıkla cevaplamıştır.

1.6. Sınırlılıklar

1. Araştırmada ele alınan koşullar çoktan seçmeli madde biçimi ve 20 maddelik test uzunluğu ile sınırlandırılmıştır.
2. Bu araştırma Mantel- Haenszel, Lojistik Regresyon ve SIBTEST yöntemleri ile sınırlandırılmıştır.
3. Bu araştırma iki tane Sosyal Bilgiler ve üç tane Ölçme ve Değerlendirme alan uzmanı olmak üzere toplam beş uzman ile sınırlandırılmıştır.

BÖLÜM II

KURAMSAL ÇERÇEVE

Bu bölümde değişen madde fonksiyonu ve değişen madde fonksiyonu belirleme yöntemleri, madde yanlılığı, olası DMF kaynakları, DMF çalışmalarının gelişimi ve ilgili araştırmalara yer verilmiştir.

2.1. Değişen Madde Fonksiyonu

Başarı testlerinde ölçtüğümüz yapıyla alakalı geçerli sonuçlar elde etmek uygulanan testin geçerli olmasına bağlıdır. Testlerin geçerlikleri ise testi oluşturan maddelere bağımlıdır. Bu yüzden geçerlik çalışmaları için madde analizleri gerekli ve önemlidir (Akcan, 2018). Madde analizleri ile bireylerin maddelerdeki performansları incelenir. DMF analizleri de bu analizlerin önemli bir parçasını oluşturmaktadır (Wiberg, 2007).

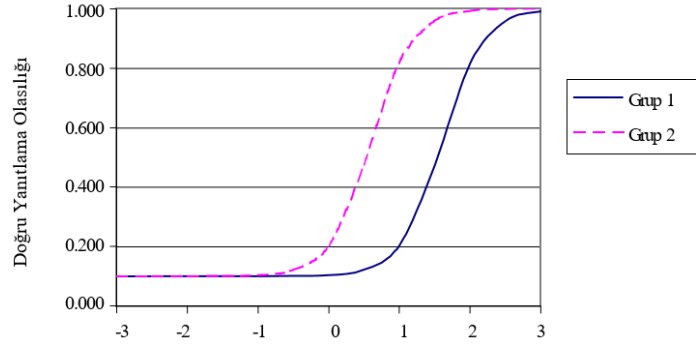
DMF, maddenin ölçtüğü iddia edilen yetenek veya nitelik açısından gruplar eşleştirildikten sonra maddede ortaya çıkan farklılaşmaları ifade eder ve maddenin ölçtüğü özellik bakımından benzer olduğu varsayılan sınav grupları arasındaki beklenmedik farktır (Dorans & Holland, 1993). Bir maddede DMF olması gizil özellik haricinde bazı durumların (yerleşim bölgesi, cinsiyet vb.) madde puanını etkilemiş olabileceğine işarettir (Ryan, 2008). Ölçme aracındaki DMF varlığı ise bu aracın tüm bireylerde aynı yapıyı ölçemediğinin bir

göstergesidir. Ölçme aracının uygulandığı gruplardan biri belli özelliklerinden dolayı maddelere daha aşına ise testten yüksek puan alma olasılığı diğer gruba göre daha yüksektir. Bu durum ise test sonucunda elde edilen bireylere ait bilişsel düzeylerin gerçekliğinin sorgulanmasına neden olur (Walker,2011).

DMF analizlerinde maddelerin herhangi bir grup lehine işleyip işlemediğine karar verilir (Douglas, Roussos & Stout, 1996). Bunun için öncelikle ölçülecek özellik belirlenerek bireyler bu özelliklerine göre gruplara ayrılır. Daha sonra ise gruplar ölçülecek özellik açısından aynı yetenek düzeyine getirilir (Holland & Wainer, 1993). Bu eşleştirmelerin yapıldığı gruplar ‘odak grup’ ve ‘referans grup’ olarak adlandırılır. DMF analizleri odak ve referans olarak betimlenen bu iki grubun performansları arasında karşılaştırma yapma sürecidir (Walker, 2011). Referans grup karşılaştırma için kullanılan ve avantajlı konumda olduğu düşünülen grup, odak grup ise ilgili madde yönünden dezavantajlı konumda ve risk altında olan grup olarak tanımlanır (French & Finch, 2013; Holland & Wainer, 1993). DMF, odak ve referans gruplar arasında maddenin istatistiksel özellikleri tutarsız olduğunda meydana gelir (Sireci & Rios, 2013). Analiz sürecinde maddenin referans grubun lehine işleyip işlemediği hipotezi test edilir ve bu hipotezin doğrulanması maddenin odak grup aleyhinde çalıştığını gösterir (Holland & Thayer, 1988). Bu durum θ yetenek seviyesinde grup üyeliği ile madde yanıtlama performansı arasındaki ilişkinin bir kanıtıdır (Holland & Wainer, 1993). Bu ilişki karşımıza tek biçimli (TB DMF) ve tek biçimli olmayan (TBO DMF) olarak iki farklı şekilde çıkabilir.

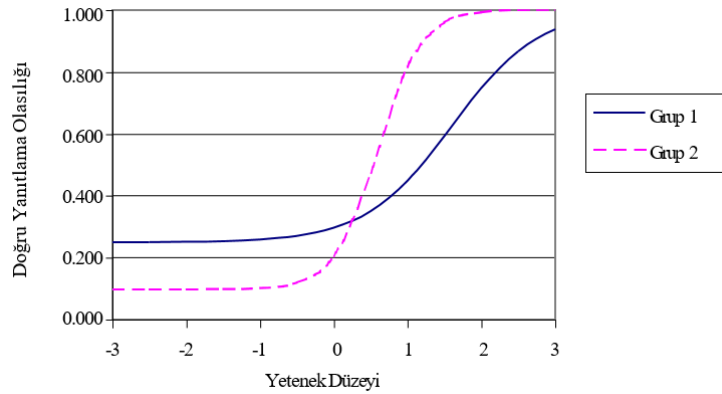
Bir maddenin doğru yanıtlanma ihtimalinin belli bir grup lehine tüm yetenek düzeylerinde tutarlı olması TB DMF olarak tanımlanır. TB DMF’de bir grubun üyeleri tüm yetenek düzeyleri için daha avantajlıdır. TBO DMF ise maddenin doğru yanıtlanma olasılığının aynı grubun tüm yetenek düzeylerinde tutarlılık göstermemesi durumu olarak tanımlanır. (Zumbo, 1999). Ayrıca eğer odak ve referans gruplar arasında yalnızca madde güçlüğü farklılık oluştursa tek biçimli, madde güçlüğü ile birlikte madde ayırt ediciliği de gruplar

arasında farklı ise tek biçimli olmayan değişen madde fonksiyonundan bahsedilebilir (Camilli & Shepard, 1994).



Şekil 1. Tek biçimli değişen madde fonksiyonu.

Şekil 1 tek biçimli DMF gösteren bir madde için madde karakteristik eğrisini (MKE) göstermektedir (Zumbo, 1999). MKE'de eğriler kesişmemiş ve θ yetenek düzeyleri için fonksiyonun iki farklı görüntüsü oluşmuştur. Yani aynı yetenek düzeyindeki bireyler farklı alt gruplardan geldikleri için maddeyi doğru yanıtlama olasılıkları $P(\theta)$ farklılaşmıştır. Madde 2. grup için tüm yetenek düzeylerinde daha avantajlıdır bu yüzden tek biçimli DMF varlığından söz edilebilir.



Şekil 2. Tek biçimli olmayan değişen madde fonksiyonu.

Şekil 2 ise tek biçimli olmayan DMF gösteren bir madde için MKE'yi göstermektedir. (Zumbo, 1999). θ yetenek düzeyleri için fonksiyonun yine iki farklı görüntüsü oluşmuştur, yani aynı yetenek düzeyindeki bireyler farklı alt gruplardan geldikleri için maddeyi doğru

yanıtlama olasılıkları $P(\theta)$ farklılaşmıştır. Ancak bu defa MKE’de eğriler kesişmiştir. Madde farklı yetenek seviyesindeki öğrenciler için aynı avantaja sahip değildir. Ortalama seviyenin altındaki öğrenciler için madde 1. grup lehine işlerken, ortalama seviye üzerindeki bireyler için 2. grup lehine işlemiştir. Başka bir deyişle ortalama altında yetenek seviyesine sahip bireyler söz konusu olduğunda birinci grup, ortalama üstünde yetenek seviyesine sahip bireyler söz konusu olduğunda ikinci grup avantajlıdır. Yani grup üyelikleri ve θ yetenek düzeyleri arasında bir etkileşim söz konusudur. Bu yüzden bu maddede tek biçimli olmayan DMF varlığından söz edilebilir.

DMF gösteren maddeler bazı kriterlere göre sınıflandırılır. Bu sınıflandırma Educational Testing Service (ETS) tarafından üç kategori şeklinde yapılmıştır. Buna göre en düşük seviye olan A seviyesinde DMF gösteren maddeler ihmal edilip testte kalabilirken, orta derece olan B ve önemli düzey sayılan C seviyesinde DMF gösteren maddeler alan uzmanları tarafından incelenerek yanlışlık içerip içermediklerine göre testte kalmalarına veya testten çıkarılmalarına karar verilir (Ramsey’den aktaran Atalay, 2010).

2.2. Değişen Madde Fonksiyonu Belirlemede Kullanılan Yöntemler

Alan yazında birçok DMF belirleme yöntemi vardır ancak bu yöntemlerin farklı özellikleri dikkate alındığında pratikte çok azı kullanılmaktadır (Clauser & Mazor, 1998). Bu yöntemler eşleştirme için farklı kriterler kullandıklarından tam bir uyum içinde değildir bu yüzden yöntemlerin birbirlerine göre avantaj ve dezavantajları göz önünde bulundurularak birden fazla yöntemin kullanılması önerilmektedir (Holland & Wainer, 1993). DMF belirleme yöntemleri verilerin parametrik olma durumlarına, eşleme değişkeninin gözlenen veya örtük olmasına, madde puanlarının iki kategorili veya çok kategorili olmasına, tek biçimli olan ve tek biçimli olmayan DMF’yi tespit etme özelliklerine göre farklı farklı sınıflandırılabilir (Wiberg, 2007). Bu sınıflandırmalardan en yaygın olarak kullanılan yöntemin KTK ya da MTK’ya dayalı olmasıdır.

İki kategorili puanlama yapılan maddelerde, madde performansının belirlenmesine (parametrik ve parametrik olmayan) ve eşleme değişkenine (KTK ve MTK) göre Potenza & Dorans (1995) tarafından oluşturulmuş sınıflama tablosu Tablo 1’de verilmiştir.

Tablo 1

KTK ve MTK’ya Dayalı DMF Belirleme Yöntemleri

		Eşleme Değişkenine Göre	
		Klasik Test Kuramı	Madde Tepki Kuramı
Madde Performansının Belirlenmesine Göre	Parametrik	• Log-Doğrusal Modeller • Lojistik Regresyon	• MTK- Olabilirlik Oranı • Sınırlandırılmış Bilgi • MTK-LR • Lord'un Kikare testi • b parametresi indeksi
	Parametrik Olmayan	• Mantel-Haenszel • Standartlaştırma • Ki-Kare Yöntemi	• SIBTEST

Klasik Test Kuramı’na dayalı yöntemler olan Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) yöntemleri iki kategoriyle puanlama yapılan maddeleri belirlemek için güçlü yaklaşımlar olarak görülür (Erdem Keklik, 2012). Ayrıca Wiberg (2007) tarafından önerilen sınıflandırma tablosuna göre MH yöntemi TB DMF analizleri için kullanılan yöntem olarak görülürken, LR ve SIBTEST yöntemleri hem TB DMF hem de TBO DMF analizleri için kullanılabilecek yöntemler olarak sunulmuştur.

Bu çalışmada karşılaştırılmak üzere KTK’ya dayalı MH ve LR yöntemleri ile MTK’ya dayalı SIBTEST yöntemleri kullanılmıştır. Bu yüzden yalnızca bu üç yönteme ait detaylı bilgilere yer verilmiştir.

2.2.1. Mantel-Haenszel Yöntemi

1959 yılında William G. Cochran, Nathan Mantel, William Haenszel tarafından iki ‘sağ kalım’ dağılımını kıyaslamak için önerilmiş bir ki-kare testidir (Yıldız & Okyay, 2017). Sağ kalım analizi; araştırılan bir olayın, ortaya çıkma sürecine kadar üretilmiş verilerin analiz edilmesidir. Ölümcül bir hastalığa yakalanmış bir hastanın, hastalık süresince aldığı tedavinin, hastalığa verdiği tepkinin, hastanın genel durumunun her biri bir veri kabul edildiğinde bu verilerle yapılan analize sağ kalım analizi denilmektedir. Bu olay sadece hastanın ölümü ile sonuçlanmayabilir, örneğin hastanın iyileşme sürecinde ürettiği veriler de bu analize örnek olabilir (Tuncay, 2005). Neden-sonuç ilişkisinin incelendiği bu tür çalışmalarda, çalışma gruplarının uygun biçimde seçilmemesinden dolayı karıştırıcı etkenlerin neden-sonuç ilişkisini etkilemesi ile sonuçlar olduğundan farklı (daha fazla veya daha az) hesaplanabilmektedir. Söz konusu “karıştırıcı etkenler”; yaş, cinsiyet, dil, din, ırk gibi ‘yanlılık’ olarak tanımladığımız pek çok değişken olabilir. Sigaranın karaciğer kanserine olan etkisinin incelendiği bir araştırmada sigaranın karaciğer kanserini artırdığına yönelik bir sonucuna ulaşılabilir. Ancak sigara kullanan insanların çoğunluğu alkol de kullanıyorsa burada alkol aslında karıştırıcı etken kabul edilir. Kanseri tetikleyen sigara mı alkol mü olduğu sonucuna kesin olarak varılamaz (Yıldız & Okyay, 2017). Bu araştırmada alkol karıştırıcı etken yani bu araştırmanın yanlılığı olarak düşünülür.

MH yöntemi ile eşleştirilmiş gruplarda yepyeni bir süreç geliştirildiği kabul edilir. Bu süreci Holland 1985’te ve birkaç yıl sonra 1998’de de Holland ve Thayer birlikte değişen madde fonksiyonunu değerlendirmede kullanmak üzere revize etmişlerdir (Erdem Keklik, 2012).

Mantel-Haenszel, χ^2 istatistiğine dayalı bir yöntemdir. Yöntemde odak ve referans gruptaki bireyler performans gösterme durumlarına göre denkleştirilerek performansları olasılık oranlarına göre karşılaştırılır. Denkleştirme kriteri olarak toplam test puanları kullanılır.

MH tekniğinde, yetenek düzeyleri aynı olan fakat bir yönüyle gruplaşan bireylerin bulundukları grup aidiyeti ile toplam test puanları arasındaki ilişki test edilir (Ackerman &

Evans, 1992). Bu yüzden odak ve referans grupların toplam test puanları karşılaştırılır. Test edilen maddeler için yetenek düzeylerinde matris oluşturulur (Roussos ve Stout, 1996).

Tablo 2

Referans ve Odak Grubun Cevap Dağılımı

Grup	1	0	Toplam
Referans	A _j	B _j	M _{Rj}
Odak	C _j	D _j	M _{Oj}
Toplam	M _{Dj}	M _{Yj}	T _j

Tablo 2’de verilmiş değerlerle αMH değeri hesaplanır. Buna göre;

$$\alpha MH = \frac{\frac{\sum A_j D_j}{T_j}}{\frac{\sum B_j C_j}{T_j}} \text{ eşitliği elde edilir.}$$

Hesaplama sonucu bulunan α değeri bir olasılık oranıdır yani toplam test puanlarına göre denkleştirilmiş odak ve referans gruptaki bireylerin maddeye doğru yanıt verme olasılığına ait bir orandır. Bu hesaplamaya göre α değeri her iki grup için farklı değer aralığında bulunur (odak grup için 0-1 aralığı, referans grup için 1- ∞ aralığı) (Clauser & Mazor, 1998). Ortaya çıkan bu sonucun yorumlanması oldukça zor olduğu için Holland ve Thayer (1998) logaritmik bir dönüşüm ile ΔMH istatistiğini aşağıdaki formüle göre hesaplamıştır:

$$\Delta MH = (-2,35) \cdot \ln(\alpha MH)$$

ΔMH değeri bulguların yorumlanmasını kolaylaştırır (Dorans & Holland, 1992). Elde edilen değer pozitifse ‘madde odak gruba avantaj sağlamıştır’ şeklinde yorumlanırken, negatifse ‘madde referans gruba avantaj sağlamıştır’ şeklinde yorumlanır. Eğer değer 1’e eşitse madde herhangi bir gruba avantaj sağlamamıştır yani her iki grubun da maddeyi doğru yanıtlama olasılıkları birbirine eşittir (Holland & Thayer, 1986).

Hesaplanan ΔMH değerine göre eğer maddenin farklı fonksiyonlaştığı tespit edilirse teste karışan bu farklılaşmanın daha iyi anlaşılabilmesi adına DMF istatistiğinin büyüklüğünün yani etki büyüklüğünün hesaplanması gerekir. Educational Testing Service (ETS) için oluşturulmuş ve önerilen ölçütler aşağıda verilmiştir.

Tablo 3

MH Yönteminde Etki Büyüklüğünün Yorumlama Ölçüleri

DMF Düzeyi	Değer Ölçütü	Açıklaması
A	$ \Delta MH < 1$	İhmal edilebilir düzey DMF
B	$1 \leq \Delta MH < 1,5$	Orta düzey DMF
C	$ \Delta MH \geq 1,5$	Yüksek düzey DMF

Tabloda görüldüğü gibi, DMF istatistiği için etki büyüklüğü Zieky (1993) tarafından üç gruba ayrılmıştır. A düzeyi bu gruplamanın en düşük düzeyidir. Bu düzeydeki maddelerin DMF düzeyleri ihmal edilebilir ve maddeler testte kalabilir. Orta düzey olan B düzeyi ve yüksek düzey olan C düzeyindeki maddeler ise testin geçerliğini etkileyeceğinden bu maddelerin düzeltilmesi veya testten çıkarılması önerilir (Kan, 2007).

Mantel-Haenszel yöntemi örneklem sayısının çok olmadığı durumlarda da DMF belirlemeye imkan sağladığı için avantajlı bir yöntemdir. Ayrıca verilen formüllerle elle hesaplamak dışında farklı programlarla da (SAS, DIFAS, EZDIF, R, SPSS) hesaplanabilmektedir (Ayala, 2009). MH yöntemi ile etki büyüklüğünün elde edilebiliyor olması ve bununla maddeler hakkında yorumlar yapılabilmesi de bu yöntemi avantajlı kılmaktadır (Clauser & Mazor, 1998). ETS tarafından da en çok tercih gören yöntem bu yöntemdir. MH yönteminin en büyük problemi TBO DMF'yi belirleme konusundaki sınırlılığıdır (Bertrand & Boiteau, 2003; Caluser & Mazor, 1998). Bu sınırlılığa alternatif olarak Swamithan & Rogers (1990) tarafından Lojistik Regresyon yöntemi geliştirilmiştir.

2.2.2. Lojistik Regresyon Yöntemi

Lojistik regresyon, bağımlı değişkenin kategorik bir değişken olduğu regresyon durumudur. Sürekli olmayan bağımlı değişkeni sürekli bağımsız değişkenlerle açıklamaya çalışan istatistiksel yöntemdir. Lojistik regresyonun amacı, iki yönlü karakteristik (bağımlı değişken = yanıt veya sonuç değişkeni) ile ilgili bir dizi bağımsız (öngörücü veya açıklayıcı)

değişken arasındaki ilişkiyi tanımlamak için en uygun modeli bulmaktır. Lojistik regresyon ile örneğin; ‘Akciğer kanseri olma ihtimali (evet ya da hayır), kilo ve günde içilen her paket sigara için nasıl değişir?’ probleminin analizini gerçekleştirebiliriz (Swaminathan & Rogers, 1990 ; Zumbo, 1999). Bu yöntem yardımı ile bağımlı değişkenin varyansının ne kadarlık bir oranının bağımsız değişkenler yardımı ile açıklanabildiğini, bağımlı değişken üzerindeki etki büyüklüğünü ve bağımsız değişkenler arasında ilişki olup olmadığını açıklayabiliriz. Lojistik Regresyon yönteminde katsayılar maksimum olabilirlik yöntemi yardımıyla kestirilebilir (Bakan Kalaycıoğlu, 2008). Bu yöntem, maddeye verilen tepki ve toplam puan arasındaki ilişkiye dayalı olarak DMF hakkında ipucu verir (Swaminathan & Rogers, 1990 ; Zumbo, 1999).

Lojistik Regresyon aynı yetenek düzeyinde bulunan bireylerin farklı grup üyelikleri olması halinde doğru cevap verme olasılıklarının istatistiksel bakımdan karşılaştırılmasını temel almaktadır (Zumbo, 1999). Bu yöntemi DMF belirleme çalışmalarında ilk olarak Swaminathan ve Rogers (1990)’ın kullandığı bilinmektedir (Bakan Kalaycıoğlu, 2008). Lojistik regresyon yönteminde maddeye verilen tepki bağımlı değişken; toplam ölçek puanı eşleştirme değişkeni, toplam ölçek puanı ile grup etkileşimi bağımsız değişkenler olarak alınır (Zumbo’dan aktaran Ayva Yörü, 2019).

Bireyin; maddeye verdiği yanıt μ , i katılımcısının toplam puanı yani gözlenen yeteneği X_i , θ , grup üyeliği g (1 ise referans, 0 ise odak grup) ve maddeyi doğru yanıtlama olasılığında katılımcılar arasındaki yetenek farklılıkları τ_1 , madde performansında grup farklılıkları τ_2 ve yetenekle grup arasındaki etkileşim durumu τ_3 olmak üzere lojistik regresyon modeli formülü aşağıdaki gibidir:

$$P(u=1)=\frac{e^z}{[1+e^z]} \quad \text{ve} \quad z = \tau_0 + \tau_1 X_i + \tau_2 g + \tau_3 (\theta g)$$

Modelde TB DMF varlığında $\tau_2 \neq 0$ ve $\tau_3 = 0$ olurken, TBO DMF varlığında $\tau_3 \neq 0$ olmaktadır. Odd oranı olan z ’nin doğal logaritması alınırsa formül,

$z = \ln \left[\frac{p_i}{1-p_i} \right] = \tau_0 + \tau_1 X_i + \tau_2 g + \tau_3 (\theta g)$ haline dönüşür (Swaminathan & Rogers, 1990). Lojistik Regresyon’da DMF belirlenirken değişkenler modele hiyerarşik olarak bir sıra ile dahil edilir (Yıldırım, 2015). Model 1’de kriter değişkeni yani toplam puan, Model 2’de grup değişkeni yani odak ve referans gruplar, Model 3’te ise etkileşim değişkeni yani toplam puan ve grup etkileşimi modele alınır (Bakan Kalaycıoğlu, 2008). Lojistik Regresyon yöntemini DMF analizlerinde kullanmak için Model 3 kullanılır. Bu model sonucunda elde edilen X_i^2 değeri ile Model 1 sonunda elde edilen X_i^2 değeri karşılaştırılarak DMF varlığına karar verilir. Elde edilen sonucun anlamlı olması DMF varlığını gösterirken, Model 1 ve Model 3’ten elde edilen ΔR^2 değeri ile de var olan DMF’nin büyüklüğü hakkında bilgi elde edilir (Zumbo’dan aktaran Yıldırım, 2015). Zumbo(1999)’a göre Lojistik Regresyon yönteminin diğer yöntemlere göre üstün yanları kriter değişkenini kategorik hale getirmeye gerek olmaması ve TB DMF ve TBO DMF modellemelerine izin vermesidir (Bakan Kalaycıoğlu, 2008).

Mantel-Haenszel yönteminde olduğu gibi Lojistik Regresyon yönteminde de maddeler içerdikleri DMF büyüklüklerine göre farklı kategorilere ayrılırlar. Bu ayrım LR tekniğinde sıklıkla büyüklük ölçüsü olarak kullanılan Nagelkerke R^2 indeksi yorumlanması ile yapılır. ΔR^2 değeri yorumlanması için alan yazında kullanılan farklı sınıflamalar Tablo 4’te verilmiştir.

Tablo 4
 ΔR^2 Değeri Yorumlanması için Kullanılabilecek Sınıflamalar

Zumbo ve Thomas (1996)	Jodoin ve Gierl (2001)	Bakan Kalaycıoğlu ve Kelecioğlu (2008)	Açıklama	Düzey
$\Delta R^2 < 0,13$	$\Delta R^2 < 0,035$	$\Delta R^2 < 0,010$	İhmal edilebilir düzey	A
$0,13 \leq \Delta R^2 < 0,26$	$0,035 \leq \Delta R^2 < 0,070$	$0,010 \leq \Delta R^2 < 0,020$	Orta düzey	B
$0,26 \leq \Delta R^2$	$0,070 \leq \Delta R^2$	$0,020 \leq \Delta R^2$	Yüksek düzey	C

Jodoin ve Gierl (2001) yaptıkları çalışmada Lojistik Regresyonda kullanılan sınıflandırmaların yeterince iyi çalışmadığını ve iyi araştırılmadığını öne sürmüşlerdir. Zumbo ve Thomas (1996)'ın kullandığı kategorilendirme ile çalışmalarındaki maddelerin yalnızca %6,8'i B kategorisinde çıkarken, kendi oluşturdukları kategorilendirme ile maddelerinin %68,2'si B kategorisinde çıkmıştır. 2006 yılında Pitoniak, Cook, Cline ve Cahalan-Laitusis'in Mantel-Haenszel ve Lojistik Regresyon yöntemlerini karşılaştırmak için yaptıkları çalışmada Lojistik Regresyon için kullanılan sınıflandırmaların DMF miktarını belirlemede yeterince hassas olmadıklarını ifade etmiştir.

Berberoğlu ve Bakan Kalaycıoğlu 2008'de yaptıkları çalışmada LR yöntemi için belirlenmiş sınıflandırmalara bağlı kalmadan sınırlandırılmış faktör analizi yöntemlerini kullanmış ve yöntemlerin tutarlılıklarını göz önünde bulundurarak kendi sınıflandırmalarını oluşturmuştur. DMF gösterdiği halde belirlenen kriterin altında kaldığı için gözden kaçırılacak maddeler olduğunu göz önünde bulundurarak daha hassas bir sınıflandırma yöntemine gitmişler ve böylece bu maddelerin de incelenmesini amaçlamışlardır. Bu çalışmada da yöntem karşılaştırması yapılacağı için daha hassas bir sınıflandırma olması adına Berberoğlu ve Bakan Kalaycıoğlu (2008)'nin sınıflandırması baz alınmıştır.

2.2.3.Simultaneous Item Bias Test (SIBTEST)

Simultaneous Item Bias Test (SIBTEST) yöntemi, temelleri Dorans ve Kulick (1986)'in Standardizasyon yaklaşımına dayanan ve Shealy ve Stout (1993) tarafından geliştirilen bir yöntemdir. Örtük puanlar yerine bireylerin toplam puanları vasıtası ile gerçek puan kestirimi yapan bir DMF yöntemidir (Abbott, 2007). Bu özelliği ise yöntemin I. tip hatayı kontrol edebilmesine imkan tanır ve dolayısıyla da diğer yöntemlere göre daha avantajlı olmasını sağlar (Nandakumar'dan aktaran Arslan, 2020). TB DMF belirlemede örtük puanlar kullanan MH yöntemi ile paralel sonuçlar ortaya çıkarmaktadır (Clauser & Mazor, 1998). Ancak yüksek ayırt ediciliğe sahip maddelerin analizinde SIBTEST yöntemi MH yöntemine göre I. tip hatanın kontrolünde daha güvenilir sonuçlar üretir ve daha küçük örneklemelerde

de daha etkilidir. Referans ve odak grubun örneklem büyüklükleri arasındaki farkın fazla olduğu durumlarda ise MH yöntemi SIBTEST yöntemine göre daha avantajlıdır (Roussos & Stout, 1996).

Yöntemde DMF analizine bakılacak olan madde veya maddeler bir grup, kalan madde veya maddeler ise diğer bir grup olarak belirlenir. İki grup haline getirilen maddelerde g eşleme değişkeni için gözlenen puanlara dayalı olarak gerçek puanlar doğrusal regresyon aracılığı ile kestirilir (Gierl, 2005). Diğer bir deyişle gruplar ölçülmesi planlanan örtük özellik temelinde eşleştirilir ve maddeler üzerindeki performans farkları araştırılır (Roussos & Stout, 1996). SIBTEST yöntemi MTK'ya dayalı bir yöntem olmasına rağmen madde eşleştirme değişkeninin gizil olduğu bir yöntemdir (Atalay, Gök, Kelecioğlu & Arslan, 2012).

SIBTEST yöntemi bir hipotez testi ile anlamlandırılmaktadır.

$$H_0 = \beta(T) = P_R(T) - P_0(T) = 0$$

$$H_1 = \beta(T) = P_R(T) - P_0(T) \neq 0$$

Belirtilen hipotez testinde $P_R(T)$ ifadesi referans gruptaki bireylere ait T gerçek puanı yani referans gruptaki bireylerin maddeyi doğru yanıtlama olasılığı iken $P_0(T)$ ise odak gruptaki bireylere ait T gerçek puanı yani odak gruptaki bireylerin maddeyi doğru yanıtlama olasılığıdır. $P_R(T)$ ve $P_0(T)$ ifadeleri arasındaki farka $\beta(T)$ adı verilir. Bu indeksin 0'a eşit olduğu yani H_0 hipotezinin gerçekleşmiş olduğu durumda T gerçek puanları bakımından eşleştirilmiş odak ve referans gruplardaki bireylerin maddeyi doğru yanıtlama olasılıkları farklılaşmamıştır, DMF yoktur şeklinde yorumlanırken, aksi durumda maddede DMF varlığından söz edilir (Abbott, 2007). SIBTEST yöntemi ile her bir yetenek düzeyinde ayrı ayrı bir değer hesaplanarak sonrasında bu değerlerin toplamı ile toplam indeks değeri olan β elde edilir (Cappaert, 2014). DMF büyüklüğü elde edilen bu β değeri yorumlanarak belirlenebilir. DMF büyüklüğü yorumlaması için Roussos ve Stout (1996) tarafından önerilen sınıflama ölçütü Tablo 5'te verilmiştir.

Tablo 5

SIBTEST Yönteminin Değerlendirilmesinde Önerilen Sınıflama Ölçütleri

DMF Düzeyi	Açıklama	Değer Ölçütü
A	İhmal edilebilir düzey	$\beta < 0,059$
B	Orta düzey	$0,059 \leq \beta < 0,088$
C	Yüksek düzey	$\beta \geq 0,088$

Hipotez testinde elde edilen β değerinin pozitif olması maddenin referans grubu lehine, negatif olması ise odak grup lehine çalıştığının göstergesidir (Abbott, 2007).

SIBTEST'in geliştiricileri olarak bilinen Li ve Stout (1996), çalışmalarında TBO DMF belirlerken SIBTEST yönteminden daha iyi olduğunu iddia ettikleri Crossing-SIBTEST (CSIBTEST) yöntemini geliştirmişlerdir. SIBTEST ve CSIBTEST analizlerinin birlikte hesaplandığı durumlarda normalde elde edilen toplam indeks değeri β , karşımıza β_{uni} ve β_{cro} olarak çıkar. β_{uni} değeri normal toplam indeks değeri β 'ye karşılık gelirken β_{cro} , TBO DMF'yi belirlemek için daha etkili olduğu düşünülen CSIBTEST için indeks değerine karşılık gelir. Eğer bir madde TB DMF'ye sahipse β_{cro} ve β_{uni} eşit olarak karşımıza çıkar ancak maddede TBO DMF varsa genelde $|\beta_{cro}| \geq |\beta_{uni}|$ durumu söz konusudur (Chalmers, 2017).

2.3. Madde Yanlılığı

DMF çalışmaları yapılırken amaç test geliştiricilerin yanlılık içermeyen maddeler üretmesine yol göstermektir. DMF analizi yapmak madde yanlılığı için kanıt toplamaya başlama sürecidir ancak tek başına yeterli değildir (Uzun & Gelbal, 2017). Osterlind ve Everson (2009)'a göre istatistiksel süreçler sonucunda elde edilen DMF analizlerinin sonuçlarının yorumlanması en az analiz süreci kadar önem taşımakta ve iyi bir karar verme yeteneği hatta bilgelik gerektirmektedir. Çünkü bir maddede tespit edilen DMF, o maddenin her zaman bir grubun lehine çalıştığının göstergesi olmayabilir. Ironson (1982)'ın verdiği

örnekte bir kek pişirmek için gerekli olan sıcaklık sorulmaktadır. Madde istatistiksel olarak kızların lehine DMF göstermiştir ancak aşçı seçmek için kullanılacak olan bir sınav söz konusu ise bu maddede yanlılıktan bahsetmek doğru değildir (Holland & Wainer'dan aktaran Ayan, 2011).

Madde yanlılığı belirleme çalışmalarının amacı gruplar arasında ortaya çıkmış farklılıkların (DMF) gerçek farklılıktan mı (alınan puanlar doğru performans farklarını gösterir) gerçek olmayan farklılıktan mı (puanlarda ölçmeden kaynaklı hatalar söz konudur) kaynaklandığını ortaya koymaktır (Camilli & Shepard, 1994). Bir maddenin DMF analizleri madde puanlarına ve grup göstergelerine ilişkin kanıtları sağlar ancak bu maddelerin yanlılık içerip içermediğinin incelenmesi için ampirik değerlendirme ve içerik analizlerine başvurulur (Zumbo, 1990). DMF tespit edilen maddeler uzman görüşlerine sunulur maddelerin yanlılık içerip içermediğine karar verilir. Uzman görüşüne sunulan maddeler hakkında iki karar ortaya çıkabilir: madde yanlılığı ve madde etkisi (Camilli & Shepard, 1994). Eğer madde gerçekten de bireylerin bulundukları gruptan dolayı bireylere doğru yanıtlama olasılıklarında şans sağlıyorsa bu maddenin yanlı olduğu söylenir. Madde etkisinin bulunduğu bir durumda ise madde istatistiksel olarak DMF gösterir ancak maddenin içeriğinde yanlılık bulunduracak bir durum söz konusu değildir ve madde ile ölçülen özellikte gruplar arasında gerçekten de bir farklılık vardır (Zumbo, 1999). Bu farklılık genelde testin uygulandığı bireylerin bilgi birikimlerinden kaynaklanır ve madde yanlılığı ile alakası yoktur.

My uncle's *field* is computer programming.

■ Look at the sentences below. In which sentence does the word *field* mean the same as it does in the boxed sentence above?

- A. The softball pitcher knew how to *field* her position.
- B. They prepared the *field* by spraying and plowing it.
- C. I know the *field* I plan to enter when I finish college.
- D. The doctor used a wall chart to examine my *field* of vision.

Şekil 3. Yanlı madde örneği.

Popham (2017)'in "What Teachers Need To Know" kitabında yer alan yanlış madde örneği Şekil 3'te gösterilmiştir. Maddede örnek cümlede verilen 'alan' kelimesinin cümleye kattığı anlamla aynı anlamda bulunan cümlelerin bulunması istenilmiştir. Popham (2017), maddenin sosyoekonomik düzeyi yüksek ailelerin çocukları tarafından daha kolaylıkla çözüleceğini ifade etmiştir. Yüksek sosyoekonomik düzeye sahip çocukların ebeveynlerinin biri veya her ikisi bir mesleğe sahip olduğu için çocukların 'field' (alan) kelimesine daha aşina olacağı ve soruyu kolayca çözebileceğini, bu durumun da öğrencilerin performansları ile alakası olmadan başarı göstergelerini etkileyeceğini belirtmiştir. Ölçülmek istenilen durum kelimenin cümleye kattığı anlam bilgisi iken madde bazı öğrencilerin yalnızca kelimeye olan aşinalıklarını ölçmekle sınırlı kalmıştır. Madde sosyoekonomik düzeyi yüksek ailelerin çocukları lehine yanlış yanlı bir maddedir.

2.4. DMF Kaynakları

DMF, bireylerin sahip olduğu bilgi ve yetenek farkından ortaya çıkabileceği gibi (madde etkisi), istenmeyen nedenlerden de (madde yanlışlığı) orta çıkabilir. Bu nedenle madde yanlışlığı sürecinin bir parçası olarak DMF kaynaklarının incelenmesi geçerli testle oluşturulması bakımından önem arz etmektedir (Camilli & Shepard, 1994). Bir maddede yanlışlık dolayısıyla da DMF çıkmasının olası bazı nedenleri vardır. Alan yazın incelendiğinde geniş ölçekli sınavlar üzerinde yapılan madde yanlışlığı çalışmaları esnasında olası DMF kaynaklarını belirleyen çalışmalar görülmektedir (Kalaycıoğlu & Berberoğlu, 2010; Kelecioğlu vd., 2014; Yıldırım & Berberoğlu, 2009). Bu çalışmalar incelendiğinde ortaya çıkan olası DMF kaynakları şunlardır:

- Yaşanılan bölge (Örneğin; Tarım ve hayvancılıkla alakalı maddelerin kırsal kesimde yaşayan öğrenciler tarafından daha kolay yanıtlanması.)

- Öğretim esnasında kullanılan metotlar (Örneğin; ders esnasında deney yaptıran bir öğretmenin öğrencilerinin deney düzeneği ile ilgili bir maddeyi yanıtlama olasılıklarının daha fazla olması.)
- Öğrencilerin kavramlara aşinalıkları (Örneğin; ilinde metro bulunan bir öğrencinin metro ile alakalı bir maddeyi daha önce metro görmemiş bir öğrenciye göre daha kolay yanıtlaması.)
- Sosyo-ekonomik düzey (Örneğin; doktorlukla ilgili bir maddenin ebeveynleri doktor olan bireyler tarafından cevaplanma olasılığının artması.)
- Cinsiyet (Örneğin; futbol sorularını erkek öğrencilerin cevaplama olasılığının artması.)
- Madde içeriği ve formatı (Örneğin; görsel uzamsal zeka gerektiren maddelerin bu zeka türüne sahip öğrenciler tarafından cevaplanma olasılığının daha fazla olması.)

Zenisky, Hambleton ve Robin (2003)'te yaptıkları çalışmada test geliştirme süreçlerinde bulunan bireyler için kullanılmak üzere olası DMF kaynaklarını şu şekilde belirtmişlerdir:

- Günlük yaşam durumlarını içeren madde içerikleri: Özellikle uzay, dünya, bilim ve teknoloji içerikli maddelerin yarısından fazlasının DMF içerdiği belirtilmiştir.
- Görsel-uzamsal beceri gerektiren madde içerikleri: Şeklin hazır verildiği çoktan seçmeli madde türlerinde erkek öğrenciler daha başarılı iken şeklin çizilmesi veya oluşturulması istenilen açık uçlu maddelerde kız öğrencilerin daha başarılı olduğu tespit edilmiştir.
- Madde türü: Açık uçlu maddelerin genellikle kız öğrencilerin, çoktan seçmeli maddelerin ise genellikle erkek öğrencilerin lehine işledikleri belirtilmiştir.

2.5. DMF Çalışmalarının Gelişimi

DMF çalışmalarının gelişimi üç kuşakta ele alınmaktadır (Zumbo, 2007):

Birinci kuşak DMF analizlerinde, madde yanlılığı kavramının etkin olduğu görülmektedir. Başarı, sertifika, lisans vb testlerdeki tarafsızlığın sınav sonuçlarında oluşturacağı adaletsizlik endişesinden ortaya çıkmıştır. Yanlılık çalışmalarının özellikle cinsiyet ve etnik gruplar üzerinde olduğu görülmektedir.

İkinci kuşak DMF analizlerinde, madde yanlılığı kavramının yerini DMF kavramı almış ve dolayısıyla da yanlılık ve madde etkisi ayrımı yapılmıştır. Birinci kuşak çalışmalardan farklı olarak DMF belirlenen maddeler uzman görüşüne sunulmuş ve yanlı olup olmadıklarına bu şekilde karar verilerek madde etkisi gösteren maddelerin ayrımı sağlanmıştır. Ayrıca bu ayrımın ortaya çıkmasıyla farklı DMF belirleme yöntemlerine başvurulmuş ve bu yöntemler kullanıcıların erişebilirliği ve kullanım kolaylığı açısından kıyaslanmıştır. Ancak DMF varlığının nedeni ile ilgili çalışmalar bu dönemde yetersiz kalmıştır.

Üçüncü kuşak DMF analizleri, ikinci kuşakta ortaya çıkan boşluğu doldurmak adına gerçekleştirilmiştir. Bu dönem genelde DMF'nin meydana geliş sebebi üzerine odaklanılmış ve karmaşık veri yapıları üzerinde çalışılmış ve birçok değişken DMF kaynağı olarak belirlenmiştir (Zumbo'dan aktaran Ayva Yörü, 2019).

2.6. İlgili Araştırmalar

Alanyazında DMF ve madde yanlılığı ile alakalı çok sayıda çalışma bulunmaktadır. Bu çalışmalar içerisinde çeşitli DMF belirleme yöntemlerinin karşılaştırılması yapılan çalışmalar incelenmiş ve özellikle de bu çalışmada kullanılacak SIBTEST, LR ve MH yöntemlerinin karşılaştırılmasını esas alan çalışmalara aşağıda yer verilmiştir.

Narayanan ve Swaminathan (1996), simülatif veri üzerinden gerçekleştirdikleri çalışmada MH, LR ve SIBTEST yöntemlerini karşılaştırmıştır. Çalışma sonucunda TBO DMF tespitinde LR ve SIBTEST yöntemlerinin daha tutarlı sonuçlar verdiği tespit edilmiştir.

Atalay, Gök, Kelecioğlu ve Arsan (2012), DMF belirleme yöntemlerini simülasyon çalışması üzerinden karşılaştırmıştır. Çalışmada MH, LR, MTK-OO ve SIBTEST yöntemleri kullanılmıştır. Çalışmanın sonucunda ise örtük puan yöntemlerinin gözlenen puan yöntemlerine göre DMF'li maddeleri daha duyarlı belirlediği tespit edilmiştir. Ayrıca gözlenen puan yöntemleri içerisinde de MH yönteminin LR yöntemine oranla daha duyarlı olduğu sonucuna varılmıştır. Çalışma sonucunda küçük örneklerde MH, LR ve MTK-OO teknikleri, büyük örneklerde ise SIBTEST tekniği önerilmiştir.

Kan, Sünbül ve Ömür (2013), Seviye Belirleme Sınavı'na giren öğrencilerin alt testlerinin cinsiyete göre DMF gösterip göstermediklerini incelemiştir. Çalışmanın amacı Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı DMF yöntemlerini karşılaştırmaktır. Bu kapsamda Mantel-Haenszel, Lojistik Regresyon, Dönüştürülmüş Madde Güçlüğü ve Raju'nun alan ölçüsü yöntemlerini kullanmışlardır. Çalışmanın sonucunda Klasik Test Kuramı'na ait yöntemler kendi içlerinde, Madde Tepki Kuramı'na ait yöntemler ise kendi içlerinde tutarlı sonuçlar vermiştir. Ancak MTK'ya ait yöntemler KTK'ya ait yöntemlere göre daha fazla DMF'li madde belirlemiştir.

Demir ve Köse (2014), PISA 2009 Matematik okuryazarlığı alt testi ile yaptıkları çalışmada MH, SIBTEST ve LR yöntemlerini kullanmıştır. Çalışma sonucunda kullanılan üç yöntemin de belirledikleri DMF içeren madde sayısı ve DMF düzeyleri farklılık göstermiştir

Atalay-Kabasakal, Arsan, Gök ve Kelecioğlu (2014) simülatif veri ile yaptıkları çalışmalarında MTKOO, SIBTEST ve MH yöntemlerinin karşılaştırmıştır. Çalışma sonucunda TB DMF belirlenmesinde gücün en yüksek olduğu yöntem MH olarak belirtilmiştir.

Erdem Keklik (2014), 18 simülasyon koşulu kullanarak ürettiği veriler üzerinden TB DMF'nin belirlenmesinde MH ve LR yöntemlerini karşılaştırmıştır. Çalışma sonucunda MH ve LR yöntemlerinin istatistiksel güçleri karşılaştırılmış ve yetenek dağılımlarının benzer olduğu durumda yöntemlerin DMF belirleme güçlerinin yeterli olmadığı görülmüştür. Yetenek dağılımlarının farklı olduğu durumlarda MH ve LR yöntemlerinin DMF belirleme

güçlerinin çok yüksek olduğu ifade edilmiştir. Örneklem büyüklüğü arttığında ise bu iki yöntemin DMF belirleme güçlerinin azaldığı keşfedilmiştir.

Uyar ve Kaya Uyanık (2016), PISA 2012 ile ilgili yaptıkları çalışmada MH ve LR yöntemlerini karşılaştırmıştır. Çalışma sonucunda MH tekniğinin daha fazla DMF'li madde tespit ettiği görülmüştür. Ayrıca diğer çalışmalarla paralel bir biçimde DMF tespit edilen maddelerin DMF etki büyüklükleri incelendiğinde iki teknik arasında farklar olduğu görülmüştür.

Doğan, Hambleton, Yurtçu ve Yavuz (2018), yaptıkları çalışmada farklı istatistiksel tekniklerden elde edilen tahminleri karşılaştırmak için öncelikle alan uzmanlarından maddelerin DMF düzeylerini cinsiyet değişkenine göre tahmin etmelerini istemiş ve uzmanların görüş birliklerini incelemişler, sonrasında ise Lojistik Regresyon ve Mantel-Haenszel yöntemlerini kullanılarak DMF seviyelerini hesaplamışlardır. Son aşamada ise uzmanların tahminleri ile istatistiksel analiz sonuçları karşılaştırmış ve uzmanların ve istatistiksel tekniklerin kendi aralarında uyum içinde oldukları, Fen Bilimleri testinde kısmen farklı, Sosyal Bilgiler testinde ise eşit oldukları görülmüştür.

Alan yazında yer alan bu çalışmalar incelendiğinde MH yönteminin LR ve SIBTEST yöntemlerine göre daha fazla DMF'li madde tespit ettiği ve üç yöntemin DMF tespit ettikleri maddelerin ise DMF düzeylerinde farklar olduğu görülmüştür.

Çalışmalara genel olarak bakıldığında yurtdışı literatürde yöntem karşılaştırmasına yönelik DMF çalışmalarının genelde simülasyon veriler üzerinde yapıldığı yöntemlerin hangi koşullarda daha doğru sonuçlar verdiğini test etmeye yönelik olduğu görülmüştür. Türkiye'de simülasyon çalışmalarına eğilimin çok fazla olmadığı görülmektedir .

Çalışmaların büyük bir çoğunluğunda DMF belirlemeye yönelik analizler yapılmış devamında madde yanlılığını belirlemek için yalnızca uzman kanısı alınarak yanlılık olduğu belirlenen maddeler not edilmiştir. Ayrıca uzman kanısı alınan çalışmaların birçoğunda genel olarak DMF'nin olası kaynaklarına değinilmiş fakat DMF gösteren maddelerin yanlı olup olmadığına ilişkin kesin açıklamalar yapılamamıştır. Sınavların genelde cinsiyete göre

yanlılık gösterip göstermediği üzerine çalışmalar yapılmış ancak okul türüne göre yanlılık gösterip göstermediğine yönelik az sayıda çalışmaya rastlanmıştır.

Türkiye’deki araştırmalar genel olarak değerlendirildiğinde bu araştırmaların ulusal geniş ölçekte yapılan testler ve PISA üzerinde odaklandığı dikkati çekmektedir. 2005 yılından itibaren yapılan çalışmalarda özellikle Mantel- Haenszel ve Lojistik Regresyon tekniklerinin sıklıkla kullanıldığı dikkati çekmektedir.



BÖLÜM III

YÖNTEM

Bu bölümde araştırmanın modeli, çalışma grubu, veri toplama araçları, verinin üretilmesi, ve verinin analizi ile ilgili açıklamalara yer verilmiştir.

3.1. Araştırmanın Modeli

Araştırmada başarı testi geliştirilip bu testten elde edilen veriler kullanılarak Mantel-Haenszel, Lojistik Regresyon ve SIBTEST yöntemlerinin DMF belirleme performansları incelenmiştir. Araştırma, söz konusu tekniklerin hangisinin daha etkili olduğunu incelemektedir. Araştırma değişkenler arasındaki ilişkiyi belirleyerek muhtemel sonuçları tahmin etmeye çalıştığı için ilişkisel araştırma türündedir. Ayrıca araştırmanın önceden tasarlanmış ve bazı durumların manipüle edilerek belirli bir koşulun incelendiği bir yapısı olduğu için deneysel araştırma özellikleri de bulunmaktadır (Ary, Jacobs, Sorensen & Razavieh, 2010).

3.2. Çalışma Grubu

Araştırmaya konu olan veriler 2021-2022 Eğitim-Öğretim yılında Sivas'ta öğrenim gören 5.sınıf öğrencilerinden elde edilmiştir. Araştırma, DMF belirleme yöntemlerini yerleşim bölgeleri üzerinden (kırsal-kent merkezi) incelediği için öğrenciler kırsal ve kent merkezinde bulunan öğrenciler içinden rastgele seçilmiştir. Araştırmada başarı testi geliştirme süreci de bulunduğu için testin betimsel istatistiklerinin elde edilmesi ve geçerlik ve güvenirlik çalışmalarının yapılabilmesi amacıyla taslak testin uygulandığı ayrı bir çalışma grubu belirlenmiş ve bu geçerlik ve güvenirlik çalışmaları bu grup üzerinden yürütülmüştür.

Taslak başarı testi ve analizde kullanılan veri setini cinsiyet ve yerleşim bölgelerine göre dağılımları Tablo 6 ve Tablo 7'de verilmiştir:

Tablo 6

Cinsiyet ve Yerleşim Bölgelerine Göre Dağılım (Taslak Test)

Grup	N	%
Kız	211	53,68
Erkek	182	46,32
Kent Merkezi	208	52,92
Kırsal	185	47,08

Başarı testinin geçerlik ve güvenirlik çalışmalarını gerçekleştirmek için kullanılan pilot grubun cinsiyet ve yerleşim bölgelerine göre dağılımı Tablo 6'da verilmiştir. Toplamda 393 öğrenciden veri elde edilmiştir. Bu öğrencilerin 211'i (%53,68) kız, 182'si (%46,32) erkektir. Yerleşim bölgesine göre bakıldığında ise öğrencilerin 208'ini (%52,92) kent merkezinde bulunan öğrenciler oluştururken, 185'ini ise (%47,08) kırsalda yaşayan öğrenciler oluşturmaktadır. Grupların dağılımlarının birbirine yakın olduğu görülmektedir.

Tablo 7

Cinsiyet ve Yerleşim Bölgelerine Göre Dağılım (Asıl Uygulama)

Grup	N	%
Kız	840	48,80
Erkek	882	51,20
Kent Merkezi	1167	67,8
Kırsal	555	32,2

Yöntem karşılaştırılması yapılacak verilerin elde edilmesi amacıyla uygulanan başarı testi grubunun cinsiyet ve yerleşim bölgelerine göre dağılımı Tablo 7’de verilmiştir. Toplamda 1722 öğrenciden veri elde edilmiştir. Bu öğrencilerin 840’ı (%48,8) kız, 881’i (%51,20) erkektir. Yerleşim bölgesine göre bakıldığında ise öğrencilerin 1167’sini (%67,8) kent merkezinde bulunan öğrenciler oluştururken, 554’ünü ise (%32,2) kırsalda yaşayan öğrenciler oluşturmaktadır. Swaminathan ve Rogers (1990)’a göre, yapılan DMF çalışmalarında kestirimi etkileyebilecek nedenlerden biri odak ve referans grupların örneklem büyüklükleri ve örneklem büyüklüğü oranlarıdır. Alanyazın incelendiğinde KTK’ya dayalı tekniklerin grup başına 200-250 birey bulunduğu istenilen güce ulaştığı ve parametrik olmayan yöntemlerin ise gruplarda bulunan birey sayıları eşit olmadığında maximum güce ulaştığı görülmüştür (Narayanan & Swaminathan,1996; Swaminathan & Rogers, 1993). SIBTEST yönteminin ise örneklem büyüklüğü 1000 olduğunda DMF belirlemede istenilen güce ulaştığı belirtilmiştir (Gierl vd, 2004).

Gerçek test uygulamalarında genelde odak grubun örneklem genişliğinin daha az olduğu görülür. Bu nedenle, veriler bu doğrultuda toplanmıştır.

3.3. Veri Toplama Araçları

Çalışmanın kapsamından dolayı veri toplama araçları ve süreci iki kısımdan oluşmaktadır.

3.3.1. Başarı Testi

İlk aşamada içerisinde yerleşim bölgesine göre yanlı maddeler içeren bir başarı testi geliştirilmiş, geliştirilen başarı testi madde yanlılığını belirlemek için uzman görüşüne sunulmuştur. Başarı testi geliştirilirken kullanılan aşamalar aşağıda belirtilmiştir:

3.3.1.1. Testin İçeriği ve Sınıf Düzeyi

Testin içeriğinde yerleşim bölgesine göre yanlılık içeren maddeler olması planlanmıştır. Müfredat kapsamındaki derslere ait kazanımlar incelendiğinde ve uzmanlardan alınan görüşler neticesinde testin Sosyal Bilgiler dersinden oluşturulmasına karar verilmiştir. Sosyal Bilgiler dersinin sınıf düzeyindeki kazanımları incelendiğinde yanlı madde yazmaya en uygun sınıf düzeyinin 5, 6 ve 7. sınıflar olduğu görülmüştür. Öğrencilerin vereceği cevapların analiz sonucunu etkileyeceği göz önünde bulundurularak düşük sınıf düzeyi tercih edilmiştir. 5. sınıf Sosyal Bilgiler testi geliştirilmiş, test 2. dönem uygulanacağı için testin içeriğinde 1. dönem konularının kazanım ağırlıklarına göre maddeler yazılmıştır. Oluşturulan maddeler Sosyal Bilgiler öğretmenleri, Sosyal Bilgiler alan uzmanları ve Ölçme Değerlendirme uzmanlarına incelenmiş ve geri dönütler neticesinde gerekli düzenlemeler yapılarak test son halini almıştır.

3.3.1.2. Testte Bulunan Madde Sayısı

Eşleme değişkeninin gözlenen puana dayalı olduğu DMF belirleme yöntemlerinde test uzunluğunun etkisi vardır (Swaminathan & Rogers, 1990). Ancak sabit bir madde sayısı belirtilmemiş ve DMF kapsamında yürütülen simülasyon çalışmalarında genellikle 20 ile 80 aralığında madde kullanılmıştır (Clauser vd., 1993). Türkiye’de uygulanan ulusal geniş ölçekli sınavlardaki alt testler incelendiğinde madde sayılarının genellikle 12 ile 25 madde arasında olduğu, öğretmen yapımı başarı testlerinde ise bu sayının ders süresi de göz önünde

bulundurularak 20 madde olduđu görölmüştür. Bu nedenle bu çalışmada madde sayısı 20 madde olarak belirlenmiştir.

3.3.1.3. Yanlılık İçeren Madde Sayısı

Alan yazındaki çalışmalar incelendiğinde çalışmalar genelde uygulanmış bir sınavdaki DMF ve yanılılık içeren madde sayısını belirlemeye yöneliktir ve bu sınavlar analize sokulmadan önce maddelerin yanılı olup olmadığı bilinemez. Bu yüzden DMF ve yanılılık içeren madde oranında standart bir yaklaşım bulunmamaktadır. Ancak simülasyon çalışmaları incelendiğinde DMF’li madde oranlarının genellikle %0, %10 ve %20 civarlarında tutulduğu görölmüştür (Ayva Yörü, 2019). Testteki DMF’li madde oranının fazla olmasının testin geçerliğini düşürdüğüne ve yöntemlerin de DMF belirleme güçlerini azalttığına yönelik çalışmalar da bulunmaktadır (Jodoin ve Gierl, 2001). Alan yazındaki çalışmalar dikkate alındığında bu araştırmada maddeler yazılırken maddelerin %20’si (20 soruluk testte 4 madde) yerleşim bölgesine göre yanılı olacak şekilde yazılmıştır. Test geliştirilirken ilgili kazanımlar göz önünde bulundurularak testin 2, 6, 12 ve 19. Maddeleri yanılı maddeler olarak yazılmıştır. Kalan maddeler analizler esnasında herhangi bir şüpheye yer bırakmak adına, bilgi içeren, ders kitabında kullanılan ifadeler yer veren maddeler olarak yazılarak, maddelerde herhangi bir olası yanılılık bulunmamasına dikkat edilmiştir.

3.3.2. Yanlılık Anketi

Madde yanılılığı belirlemek için yerleşim bölgesine göre oluşturulan anket kullanılmıştır. Çalışmada maddelerde yanılılık olup olmadığını kontrol etmek amacıyla kullanılan anket Bakan Kalaycıoğlu ve Kelecioğlu tarafından 2011 yılında geliştirilen anketten uyarlanmıştır. Bakan Kalaycıoğlu ve Kelecioğlu (2011) yaptıkları çalışmada Öğrenci Seçme Sınavı’nı madde yanılılığı açısından incelemiş ve yanılılık için uzman görüşüne sunmak üzere cinsiyete

göre DMF içeren maddelerle oluşturulan anket kullanmıştır. Ankette cinsiyete göre olan durumlar yerleşim bölgesine göre güncellenip uyarlanarak çalışmada kullanılmıştır.

3.4. Verilerin Elde Edilmesi ve Analizi

Çalışma için gerekli veriler geliştirilen başarı testinden elde edilmiştir. Testin geçerlik ve güvenirlik çalışmalarının yapılabilmesi için öncelikle taslak test öğrencilere uygulanmış ve analizler yapılmış, sonrasında da testin oluşturulan ve düzeltilen son formunun uygulaması yapılmıştır. Bu nedenle verilerin elde edilmesi ve analizi taslak test formu ve nihai test formu başlığı 2 kısımda incelenecektir:

3.4.1. Taslak Testin Verilerinin Elde Edilmesi

Uzman görüşlerine göre düzenlenen taslak testin maddelerinin dizgisi yapılarak kitapçık formuna dönüştürülmüştür. Öğrencilerin cevaplarını toplayabilmek adına bir optik form (cevap kağıdı) tasarlanmış ve öğrencilerin cevaplarını bu forma işaretlemeleri sağlanmıştır. Optik formda öğrencilerden kişisel bilgilerini (Ad, Soyad, Sınıf, Okul) elle doldurmaları beklenmiş ve bu veriler herhangi bir şekilde depolanmamış, yalnızca öğrencilerin sınavı ciddiyetle çözmeleri adına istenilmiştir. Öğrencilerin cinsiyeti betimsel istatistikler için, yerleşim bölgeleri ise DMF istatistikleri için optik formda işaretlenerek saklanmıştır. Oluşturulan cevap kâğıdı ve kitapçıktan 500 adet çoğaltılarak kırsal ve kent merkezindeki okullara 250 adet olmak üzere dağıtılmıştır. Öğrencilerin ders saatleri esnasında sınıf düzenlerinde 40 dakikalık süre içerisinde testi yanıtlamaları istenmiştir. Test 1. dönem konularını kapsadığı için öğrencilere ‘kazanım değerlendirme testi’ olduğu belirtilerek yanıtlatılmış ve öğrencilerden sorularla alakalı gelen sorulara cevap verilmemiştir. Özellikle kırsal kesimde yaşayan öğrencilerin yanlış maddelerle alakalı soruları olmuş ancak öğrencilere herhangi bir yardımda bulunulmamıştır. Tüm öğrenciler maddeleri ciddiyetle ve duyarlılıkla çözerek çalışmaya katkı sağlamıştır. 500 adet olarak dağıtılan formun kent

merkezinden 208, kırsaldan 185 tanesi olmak üzere toplamda 393 tanesi dönmüştür. Elde edilen cevap kâğıdı sayısının geçerlik güvenirlik analizleri yapabilmek adına yeterli olduğu görülmüştür. Toplanan cevap kâğıtları optik okuyucu yardımıyla okunarak öğrencilerin cevap örüntüleri elde edilmiştir.

3.4.2. Taslak Testin Analizi

Taslak testin veri analizinin adımları aşağıda verilmiştir.

1. Elde edilen öğrenci cevap örüntüleri cevap anahtarına göre Microsoft Excel 2016’da puanlanarak doğru cevaplara 1, yanlış ve boş cevaplara 0 olacak şekilde 1-0 verilerine dönüştürülmüştür.
2. Cevap kâğıtlarındaki öğrenci cinsiyetleri kız öğrencilere 1, erkek öğrencilere 2 verilecek şekilde, yerleşim bölgeleri ise kent merkezi 1, ilçe merkezine 2, köye 3 olacak şekilde dönüştürülmüştür. Oluşturulan yanlı maddeler kent merkezindeki öğrencilerin çözebileceği şekilde yazıldığı için daha sonra 2 ve 3 olarak kodlanan yerleşim türleri 2 olarak tek tür olarak belirlenmiştir.
3. Testin geçerlik güvenirlik çalışmalarının yapılması ve madde analizlerinin elde edilmesi için 1-0 verisine dönüştürülen test Test Analysis Program (TAP) ile analiz edilmiştir. Analiz sonucunda elde edilen madde analizleri tabloda verilmiştir:

Tablo 8

Taslak Testin Madde Analizleri

Maddeler	p	d	Düzeltilmiş nokta çift serili korelasyon katsayısı
Madde 1	0,95	0,08	0,19
Madde 2	0,69	0,47	0,38
Madde 3	0,86	0,3	0,41
Madde 4	0,62	0,64	0,48
Madde 5	0,7	0,61	0,49
Madde 6	0,68	0,34	0,24
Madde 7	0,72	0,49	0,39
Madde 8	0,79	0,43	0,42
Madde 9	0,52	0,57	0,41
Madde 10	0,68	0,65	0,53
Madde 11	0,67	0,6	0,49
Madde 12	0,67	0,5	0,42
Madde 13	0,73	0,56	0,52
Madde 14	0,59	0,59	0,44
Madde 15	0,78	0,58	0,57
Madde 16	0,71	0,57	0,50
Madde 17	0,47	0,39	0,24
Madde 18	0,56	0,69	0,51
Madde 19	0,3	0,39	0,28
Madde 20	0,66	0,66	0,53

Taslak testin madde analizleri incelendiğinde 1. maddenin madde güçlüğü (p) 0,95, ayırt ediciği (d) 0,08 ve madde toplam korelasyonu (pbis) 0,23 olarak hesaplanmıştır. Büyüköztürk (2017)'ün ve Tavşancıl (2002)'in madde toplam korelasyonu için 0.30 ve üstünün yeterli olacağı ve bu maddelerin de iyi maddeler olarak nitelendirileceğini ifade ettikleri göz önüne alındığında 1.maddenin kabul edilebilir düzeyden düşük olduğu görülmüştür.

Madde incelendiğinde ve alan uzmanlarına başvurulduğunda ortak kanı maddenin çözülebilmesi için çok fazla bilgi verilmesi olmuştur. Doğal afet özelliklerinin verilerek bu özelliklerden 'deprem' ifadesinin bulunmasının beklendiği maddede 'sarsıntı' kelimesinin madde kökünde bulunmasının sorunun çözümünü çok kolaylaştırdığı ve alt gruptaki öğrenciler tarafından da maddenin kolaylıkla yanıtladığı görülmüştür. Maddenin üst grupta çözüme oranı %100, alt grupta çözülme oranı %92'dir. Maddenin seçenek analizlerine

bakıldığında çeldiricilerin hiç çalışmadığı görülmüştür. Bu durumların maddenin madde toplam korelasyonuna etkisi olduğu düşünülmüş ve kazanımlar ve müfredat çerçevesinde uzman önerilerine göre maddede düzeltmeye gidilmiştir.

Tablo 9

Taslak Testin Betimsel İstatistikleri

İstatistikler	n
Madde Sayısı	20
Öğrenci Sayısı	393
Ortalama	13,37
Standart Sapma	4,509
Güvenirlilik (Cronbach α)	0,84
Ortalama Güçlük	0,66
Ortalama Ayırt Edicilik	0,50

Testin betimsel istatistikleri incelendiğinde ortalamasının 13,37, standart sapmasının 4,50 olduğu görülmektedir. Güvenirlilik katsayısı (Cronbach α) değeri 0,84 olarak hesaplanmıştır. Akbaş(2019)'a göre güvenirlilik katsayısının 1'e yakın olması (0,70) güvenirliliğin yüksek olduğunu gösterir. Ortalama güçlük, test maddelerinin güçlük indeksleri madde sayısına bölünerek bulunur ve 0.00'a yaklaştıkça test zorlaşırken 1,00'e yaklaştıkça test kolaylaşmaktadır (Baykul, 2010). Testin ortalama güçlüğü 0,66 çıkmıştır. Değerin 0,50'ye yakın olması testin ortalama güçlükte olduğunu göstermektedir. Test ayırt ediciliği, testin bilenle bilmeyeni ayırt etme derecesidir (Baykul, 2010). Ortalama ayırt edicilik ise 0,50 bulunmuştur. Ayırt edicilik değerinin 0,40'tan yüksek olması sebebiyle testin ayırt edici bir test olduğu söylenebilir.

4. Elde edilen madde analizleri sonucunda testin yalnızca 1. maddesinde gerekli düzenlemeler yapılarak testin nihai formu elde edilmiştir.

3.4.3. Nihai Testin Verilerinin Elde Edilmesi

Nihai testin verilerinin elde edilmesinde taslak testin verileriyle aynı yol izlenmiş, kitapçıklar ve cevap kâğıtları hazırlanmış ve okullara aynı düzende uygulama yaptırılmıştır. Yaklaşık 2000 tane kitapçık ve cevap kâğıdı basılarak kırsalda ve kent merkezinde bulunan okullara dağıtımı yapılmıştır. Okulların bulunduğu bölgeler, gönderilen ve toplanan toplam evrak sayısı tabloda verilmiştir.

Tablo 10

Nihai Test Uygulanan Okullar ve Elde Edilen Veri Sayısı

İlçe	Okul	n
Zara	Hacı Bayram Elmas YBO	43
	Cumhuriyet Ortaokulu	44
	İsmet Yılmaz İHO	34
Altınyayla	Reşit Akif Paşa	36
	Yusuf Yurt İHO	54
	75.yıl	26
Kangal	Şehit Üzeyir Gençdoğan	32
Gürün	80.Yıl YBO	64
Suşehri	Atatürk Ortaokulu	114
	Cumhuriyet Ortaokulu	69
İmranlı	Asım Özden YBO	39
	Gazipaşa Ortaokulu	142
	Selçuk Ortaokulu	221
Merkez	Süleyman Demirel OO	165
	Ziya Gökalp	170
	Cüneyt Erkan	158
	Ülkü İHO	124
	Mevlana	187

Kent merkezindeki okullara 1305 tane optik ve cevap kağıdı gönderilmiş ve bu evrakların 1165 tanesi dönüş yapmıştır. Kırsalda ise 635 tane optik ve cevap kağıdı gönderilirken bu evrakların 557 tanesi dönüş yapmıştır. Toplamda 1722 tane cevap kağıdı geri dönüş yapmış, 7 farklı ilçeden veri elde edilmiştir. Eksik gelen evraklar öğrencilerin devamsızlıkları sebeplidir. Elde edilen cevap kâğıdı sayısının DMF analizlerini yapabilmek adına fazlasıyla yeterli olduğu görülmüştür. Toplanan cevap kâğıtları optik okuyucu yardımıyla okunarak öğrencilerin cevap örüntüleri elde edilmiştir.

3.4.4. Uzmanlarca Yanlı Olduğu Düşünülen Maddelerin Belirlenmesi

Çalışmada yanlılık süreci ters yönde işletileceği için nihai testin analizleri gerçekleştirilmeden önce uzmanların yanlı olduğunu düşündükleri maddeler belirlenmiştir. Bunun için önce uzmanlara madde yanlılığı konusunda bilgi verilmiş ve yanlı madde örnekleri gösterilmiştir. Sonrasında ise geliştirilen ve yanlı maddeler içeren testin tüm maddeleri yanlılık anketi ile birlikte uzmanlara incelenilerek uzmanlardan yanlı olduğu düşünülen maddeler ve bu maddelerde bulunan yanlılık nedenleri elde edilmiştir. Aşağıdaki tabloda uzmanların yanlı olduğunu düşündüğü maddeler ve kaç uzmanın bu maddelerin yanlı olduğunu düşündüğü verilmiştir.

Tablo 11

Uzmanlarca Yanlı Olduğu Düşünülen Maddeler

Madde	Kişi Sayısı
Madde 2	5
Madde 3	1
Madde 6	4
Madde 7	2
Madde 10	1
Madde 12	5
Madde 15	1
Madde 16	2
Madde 19	5

Madde 2 test geliştirme sürecinde yerleşim bölgesine göre yanlı olarak yazılmış ve uzmanların tamamınca da yanlı madde olarak düşünülmüştür. Madde içerisinde yazılan senaryoda bireyin annesi iş teklifi almakta ve aile bu nedenle şehir değiştirmektedir. Madde ‘Yerleşim bölgelerinin daha yakın olduğu konuların soru kapsamında içerilmesi nedeni ile’ yanlılık içermektedir. Genel görüşe bakıldığında ise uzmanlar, ‘bir iş teklifi alma bu bu tekliften dolayı şehir değiştirme’ durumunun il merkezindeki öğrencilerin ebeveynlerinin çalışma durumlarından dolayı daha rahat algılayabileceği, bu yüzden maddenin kent merkezindeki öğrencilerin lehine çalışacağını düşünmüştür.

Madde 6 test geliştirme sürecinde yerleşim bölgesine göre yanlı olarak yazılmış ve uzmanların dördü tarafından da yanlı olarak düşünülmüştür. Madde içerisinde Sivas Şehir Merkezinde bulunan tarihi eserlerle alakalı bilgi verilmiş ve bu maddeyi kent merkezindeki öğrenciler rahat yanıtlarken, kırsaldaki öğrencilerin maddeyi yanıtlama olasılıklarının daha düşük olacağı düşünülmüştür. Madde uygulama esnasında öğrenciler tarafından test uygulayıcısına en çok danışılan madde olmuş, öğrenciler madde hakkında çok fazla soru sormuştur. Madde bir önceki maddede olduğu gibi ‘Yerleşim bölgelerinin daha yakın olduğu konuların soru kapsamında içerilmesi nedeni ile’ yanlılık içermektedir. Genel görüşe bakıldığında ise uzmanlar, yerleşim bölgesine bağlı olarak kent merkezindeki grubun maddenin içeriğine aşina olması nedeni ile maddenin kent merkezindeki öğrencilerin lehine çalışacağını düşünmüştür.

Madde 12 test geliştirme sürecinde yerleşim bölgesine göre yanlı olarak yazılmış ve uzmanların tamamınca da yanlı madde olarak düşünülmüştür. Madde içeriğinde internet alışverişi, internette yemek siparişi durumları söz konusudur ve kırsal kesimde internette yemek siparişi kavramı öğrencilerin hayatında bulunan bir kavram değildir. Madde ‘Yerleşim bölgesine bağlı olarak bir grubun maddenin içeriğine aşina olması nedeni ile’ yanlılık içermektedir. Genel görüşe bakıldığında ise uzmanlar, ‘internette alışveriş yapma’ durumunu il merkezindeki öğrencilerin daha rahat algılayabileceği, bu yüzden maddenin kent merkezindeki öğrencilerin lehine çalışacağını düşünmüştür.

Madde 19 test geliştirme sürecinde yerleşim bölgesine göre yanlı olarak yazılmış ve uzmanların tamamınca da yanlı madde olarak düşünülmüştür. Maddenin içeriğinde bir AVM kat planından yola çıkarak seçenekleri yanıtlama durumu söz konusudur. Madde ‘Yerleşim bölgesine bağlı olarak bir grubun maddenin içeriğine aşina olması nedeni ile’ yanlılık içermektedir. Genel görüşe bakıldığında ise uzmanlar, ‘bir AVM’nin kat planına bakarak mağaza bulma’ durumunu il merkezindeki öğrencilerin daha rahat algılayabileceği, bu yüzden maddenin kent merkezindeki öğrencilerin lehine çalışacağını düşünmüştür.

Testin üç maddesi uzmanların tamamınca, bir maddesi ise uzmanların büyük çoğunluğunca yanlı olarak kabul edilmiştir. Maddelerin tamamında maddenin yanlı yazılma amacı ile uzmanların genel görüşleri incelendiğinde tutarlılık olduğu görülmüştür. Test geliştirme sürecinde yerleşim bölgesine göre yanlı olacağı düşünülerek yazılan ve uzman görüşüne sunulduğunda da uzmanlar tarafından yanlı olarak görülen bu 4 madde (2, 6, 12, 19) analizler gerçekleştirilirken yanlı maddeler olarak kabul edilmiş ve DMF analizlerinde özellikle bu maddelere odaklanılmıştır.

Test geliştirme sürecinde yanlı madde olarak yazılmamış ancak bazı uzmanlar tarafından yanlı olduğu düşünülen (Madde 3, Madde7, Madde 10, Madde 15, Madde 16) maddeler analiz sonuçlarının güveni açısından ölçme uzmanı ile birlikte incelenmiş ancak maddelerin tamamen bilgi içerikli maddeler olduğu, maddelerde herhangi bir yanlılık içerecek öğe olmadığı tespit edilmiştir.

3.4.5. Nihai Testin Analizi

Nihai testin veri analizinin adımları sırasıyla aşağıda verilmiştir.

1. Elde edilen öğrenci cevap örüntüleri cevap anahtarına göre Microsoft Excel 2016’da puanlanarak doğru cevaplara 1, yanlış ve boş cevaplara 0 olacak şekilde 1-0 verilerine dönüştürülmüştür.
2. Cevap kâğıtlarındaki öğrenci cinsiyetleri kız öğrencilere 1, erkek öğrencilere 2 verilecek şekilde, yerleşim bölgeleri ise kent merkezi 1, ilçe merkezine 2, köye 3 olacak şekilde dönüştürülmüştür. Oluşturulan yanlı maddeler kent merkezindeki öğrencilerin çözebileceği şekilde yazıldığı için daha sonra 2 ve 3 olarak kodlanan yerleşim türleri 2 olarak tek tür olarak belirlenmiştir.
3. Testin geçerlik güvenirlik çalışmalarının yapılması ve madde analizlerinin elde edilmesi için 1-0 verisine dönüştürülen test Test Analysis Program (TAP) ile analiz edilmiştir. Analiz sonucunda elde edilen madde analizleri tabloda verilmiştir;

Tablo 12

Nihai Testin Madde Analizleri

Maddeler	p	d	Düzeltilmiş nokta çift serili korelasyon katsayısı
Madde 1	0,95	0,13	0,28
Madde 2	0,77	0,47	0,43
Madde 3	0,91	0,25	0,44
Madde 4	0,71	0,52	0,44
Madde 5	0,74	0,57	0,50
Madde 6	0,68	0,35	0,23
Madde 7	0,73	0,54	0,44
Madde 8	0,86	0,35	0,42
Madde 9	0,55	0,49	0,29
Madde 10	0,73	0,59	0,50
Madde 11	0,77	0,53	0,48
Madde 12	0,77	0,4	0,37
Madde 13	0,81	0,47	0,52
Madde 14	0,63	0,49	0,30
Madde 15	0,89	0,32	0,50
Madde 16	0,78	0,5	0,47
Madde 17	0,53	0,47	0,26
Madde 18	0,55	0,58	0,37
Madde 19	0,73	0,56	0,48
Madde 20	0,7	0,52	0,40

Nihai testin madde analizleri incelendiğinde düzeltilen 1. Maddenin madde güçlüğü(p) 0,95, ayırt ediciği (d) 0,13 ve madde toplam korelasyonu (pbis) 0,33 olarak hesaplanmıştır. Büyüköztürk (2017)'ün ve Tavşancıl (2002)'in madde toplam korelasyonu için 0.30 ve üstünün yeterli olacağı ve bu maddelerin de iyi maddeler olarak nitelendirileceğini ifade ettikleri göz önüne alındığında 1.maddenin kabul edilebilir düzeye ulaştığı görülmektedir. Maddenin ayırt ediciliğinin düşük olmasının nedeni maddenin öğrencilere aşırı kolay

gelmesi olarak değerlendirilmiştir. Ancak madde toplam korelasyonu belirtilen aralıkta olduğu için madde bu şekliyle kabul edilmiştir.

Tablo 13
Nihai Testin Betimsel İstatistikleri

İstatistikler	n
Madde Sayısı	20
Öğrenci Sayısı	1722
Ortalama	14,77
Standart Sapma	4,096
Güvenirlilik (Cronbach α)	0,82
Ortalama Güçlük	0,73
Ortalama Ayırt Edicilik	0,45

Testin betimsel istatistikleri incelendiğinde ortalamasının 14,77, standart sapmasının 4,09 olduğu görülmektedir. Güvenirlilik katsayısı (Cronbach α) değeri 0,82 olarak hesaplanmıştır. Akbaş(2019)'a göre güvenirlilik katsayısının 1'e yakın olması (0,70) güvenirliliğin yüksek olduğunu gösterir. Ortalama güçlük, test maddelerinin güçlük indeksleri madde sayısına bölünerek bulunur ve 0.00'a yaklaştıkça test zorlaşırken 1,00'e yaklaştıkça test kolaylaşmaktadır (Baykul,2010). Testin ortalama güçlüğü 0,73 çıkmıştır. Değerin 0,50'ye yakın olması testin ortalama güçlükte olduğunu göstermektedir. Test ayırt ediciliği, testin bilenle bilmeyeni ayırt etme derecesidir (Baykul,2010). Ortalama ayırt edicilik ise 0,45 bulunmuştur. Ayırt edicilik değerinin 0,40'tan yüksek olması sebebiyle testin ayırt edici bir test olduğu söylenebilir.

Tablo 14

Nihai Testin Betimsel İstatistikleri (Kırsal)

İstatistikler	n
Madde Sayısı	20
Öğrenci Sayısı	557
Ortalama	13,54
Standart Sapma	4,32
Güvenirlilik (Cronbach α)	0,82
Ortalama Güçlük	0,67
Ortalama Ayırt Edicilik	0,51

Alt grupların betimsel istatistiklerine ayrı ayrı incelendiğinde; Kırsal kesimde yaşayan öğrencilerin ortalamasının 13,54, standart sapmasının 4,32 olduğu görülmektedir. Güvenirlilik katsayısı (Cronbach α) değeri 0,82 olarak hesaplanmıştır. Akbaş (2019)'a göre güvenirlilik katsayısının 1'e yakın olması (0,70) güvenirliliğin yüksek olduğunu gösterir. Testin ortalama güçlüğü 0,67 çıkmıştır. Değerin 0,50'ye yakın olması testin ortalama güçlükte olduğunu göstermektedir. Ortalama ayırt edicilik ise 0,51 bulunmuştur. Ayırt edicilik değerinin 0,40'tan yüksek olması sebebiyle testin ayırt edici bir test olduğu söylenebilir.

Tablo 15

Nihai Testin Betimsel İstatistikleri (Kent Merkezi)

İstatistikler	n
Madde Sayısı	20
Öğrenci Sayısı	1165
Ortalama	15,35
Standart Sapma	3,84
Güvenirlilik (Cronbach α)	0,81
Ortalama Güçlük	0,76
Ortalama Ayırt Edicilik	0,41

Kent merkezinde yaşayan öğrencilerin ortalamasının 15,35, standart sapmasının 3,84 olduğu görülmektedir. Güvenirlik katsayısı (Cronbach α) değeri 0,81 olarak hesaplanmıştır. Akbaş (2019)'a göre güvenirlik katsayısının 1'e yakın olması (0,70) güvenirliliğin yüksek olduğunu gösterir. Testin ortalama gücü 0,76 çıkmıştır. Değerin 0,50'ye yakın olması testin ortalama güçlükte olduğunu göstermektedir. Ortalama ayırt edicilik ise 0,41 bulunmuştur. Ayırt edicilik değerinin 0,40'tan yüksek olması sebebiyle testin ayırt edici bir test olduğu söylenebilir.

4. DMF analizlerinden önce eşleme değişkeninin geçerlik kanıtlarının sağlanması gerekir (Dorans ve Holland. 1993). Bu nedenle DMF analizlerinin temel ihtiyacı olarak tek boyutluluk varsayımının kontrolü gerekmektedir. Araştırmada verilerin tek boyutluluk varsayımını kontrol etmek amacıyla öncelikle IBM SPSS Statistics 23 programında Açıklayıcı Faktör Analizi (AFA) yapılmış daha sonra belirlenen faktör yapısının doğruluğunu test etmek amacıyla LISRELL 8.80 programı yardımıyla tek boyutlu Doğrulamalı Faktör Analizi (DFA) yapılmıştır.

AFA uygulamasından önce, örneklem büyüklüğünün uygunluğunu test etmek amacıyla Kaiser- Meyer-Olkin (KMO) testi uygulanmıştır. Analiz sonucunda elde edilen KMO değeri ve Bartlett küresellik testi sonuçları tabloda gösterilmiştir.

Tablo 16

Kaiser- Meyer-Olkin (KMO) Testi Sonucu

KMO ve Bartlett Küresellik Testi		
	KMO	,936
	χ^2	5438,954
Bartlett Küresellik Testi	sd	190
	p.	,000

KMO değerinin ,936 olduğu belirlenmiştir. Ayrıca Bartlett küresellik testi sonuçları incelendiğinde, elde edilen ki-kare değerinin ($\chi^2= 5438,954$) anlamlı olduğu görülmüştür (p= .000). Yapılan analizde kayıp değer ve uç değerlerin olmadığı belirlenmiştir.

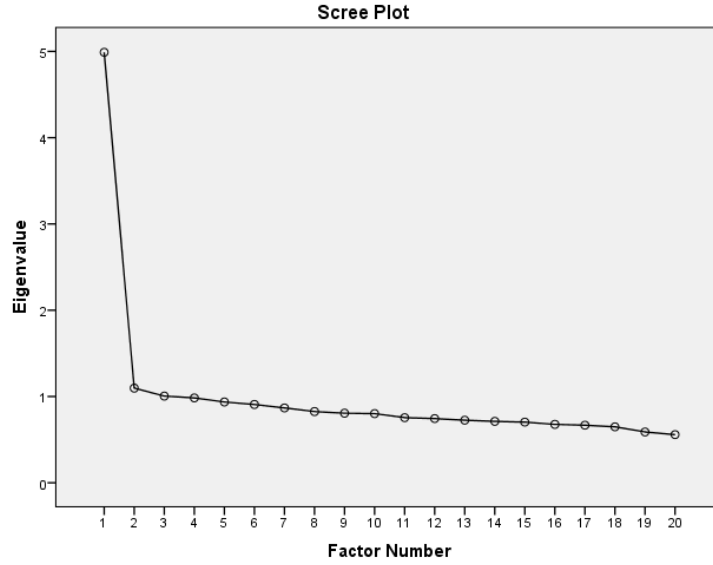
Faktör yapısını ortaya koymak amacıyla en çok olabilirlik (Maximum Likelihood) yöntemi ile döndürülmemiş faktör analizi uygulanmıştır. 1'den büyük öz değerler bir faktör olarak alınmıştır.

Tablo 17

Maddelerin Faktörlerdeki Yük Değerleri

Madde	Faktör		
	1	2	3
Madde 13	0,586	0,067	-0,062
Madde 15	0,571	0,128	-0,19
Madde 7	0,552	-0,52	-0,04
Madde 5	0,542	0,07	0,111
Madde 10	0,542	0,08	0,111
Madde 19	0,526	0,068	0,029
Madde 11	0,525	0,116	0,017
Madde 16	0,518	0,056	-0,044
Madde 3	0,499	0,119	-0,194
Madde 4	0,485	-0,011	0,027
Madde 2	0,478	-0,063	0,007
Madde 8	0,472	-0,004	-0,06
Madde 20	0,429	0,085	0,139
Madde 12	0,411	0,131	-0,11
Madde 18	0,396	0,057	0,245
Madde 14	0,32	-0,004	0,153
Madde 1	0,317	0,051	-0,145
Madde 9	0,309	-0,051	0,179
Madde 17	0,28	-0,048	0,093
Madde 6	0,238	0,164	0,076

Tablo 12 incelendiğinde maddelerin 1.faktör altında toplandığı görülmektedir. Ayrıca öz değer tablosu incelendiğinde ilk faktörün tek başına toplam varyansın yaklaşık %25'ini açıkladığı görülmüştür. Bu oran diğer faktörlere göre oldukça yüksek bir orandır (Pallant, 2007). Birinci faktörün açıkladığı toplam varyans ikinci faktörün 4,55 katıdır. İlk iki değer arasındaki oranın 3,5 olması tek faktörlülük için yeterli kabul edilir (Baykul, 2015).



Şekil 4. Yamaç birikinti grafiği.

Yamaç birikinti grafiği (Scree Plot) incelendiğinde de grafiğin birinci faktörüyle ikinci faktörü arasında büyük bir kopma olduğu görünmektedir.

AFA sonucunda tek faktörlü çıkan yapının doğruluğunu test etmek için yapıya DFA uygulanmıştır. DFA'da model veri uyumunu incelemek için dikkate alınması gereken istatistikler vardır (Çokluk, Şekercioğlu ve Büyüköztürk, 2002). Tek faktörlü olarak test edilen yapıya ait elde edilen istatistikler tabloda verilmiştir. Tabloda verilen kriter değerler Çelik ve Yılmaz (2013)'ten elde edilmiştir.

Tablo 18
DFA Model Veri Uyumu İstatistikleri

Uyum Ölçüleri	İyi Uyum	Kabul Edilebilir Uyum	Test İstatistikleri
<i>RMSEA</i>	$0 < RMSEA < 0,05$	$0,05 \leq RMSEA \leq 0,08$	0,024
<i>CFI</i>	$0,97 \leq CFI \leq 1$	$0,95 \leq CFI < 0,97$	0,99
<i>GFI</i>	$0,95 \leq GFI \leq 1$	$0,90 \leq CFI < 0,95$	0,98
<i>AGFI</i>	$0,90 \leq GFI \leq 1$	$0,85 \leq CFI < 0,90$	0,98

Testin tek boyutluluğuna ilişkin model veri uyumu incelendiğinde RMSEA, CFI, GFI, AGFI değerlerine bakılmıştır. Testin kriter değerlere göre yeterli düzeyde model veri uyumu sağladığı gözlenmiştir. Bu nedenle testin tek boyutluluk varsayımını karşıladığı görülmüştür. Verilerin betimsel istatistikleri ve faktör analizleri incelendikten sonra DMF analizlerine geçilmiştir.

5. Cinsiyet verisi betimsel istatistik verisi olarak kullanıldığı ve DMF verilerinde kullanılmayacağı için bu sütun Excel tablosundan çıkarılmıştır. Daha sonra Excel dosyası CSV formatında kaydedilerek R’da analiz edilmeye hazır hale getirilmiştir.

6. Verilerin analizi R programında ve IBM SPSS Statistics 23 programında gerçekleştirilmiştir. R programında analizler gerçekleştirilirken programda yer alan ve DMF belirleme tekniklerinin neredeyse tamamının analizini gerçekleştirebilen; bu tekniklerin hem tek başlarına hem de karşılaştırmalı olarak analiz edilmesine olanak sağlayan “difR” paketi kullanılmıştır (Magis, Beland & Raiche, 2018).

BÖLÜM IV

BULGULAR VE YORUM

Bu bölümde her bir probleme ilişkin bulgulara ve bu alanda yapılmış çalışmalarla olan benzerlik ve farklılıklara ilişkin yorumlamalara yer verilmiştir. Bulgu ve yorumlar sunulurken araştırma soruları sırası ile takip edilmiştir.

4.1. Araştırma Problemi 1'e Yönelik Bulgular

Yerleşim bölgesine göre uzman görüşü alınarak yanlış olduğu belirlenmiş maddeler ile Mantel-Haenszel yöntemine göre DMF gösteren maddeler tutarlı mıdır? Probleminin çözümü için R'da bulunan “difR” paketi ile MH yöntemine yönelik analizler gerçekleştirilmiş ve elde edilen bulgular tablolarda verilmiştir.

Tablo 19

MH Yöntemi İle DMF Analizi Sonuçları

Madde	αMH	ΔMH	p	Düzye
Madde 1	0,61	1,151	0,049*	B
Madde 2	0,68	0,921	0,01*	A
Madde 3	0,73	0,746	0,14	A
Madde 4	0,43	1,969	0,00*	C
Madde 5	0,67	0,956	0,00*	A
Madde 6	0,45	1,859	0,00*	C
Madde 7	0,81	0,488	0,14	A
Madde 8	0,86	0,368	0,41	A
Madde 9	0,78	0,592	0,03*	A
Madde 10	0,86	0,360	0,30	A
Madde 11	0,97	0,063	0,91	A
Madde 12	0,65	0,983	0,00*	A
Madde 13	0,69	0,860	0,03*	A
Madde 14	0,96	0,084	0,81	A
Madde 15	0,58	1,263	0,01*	B
Madde 16	0,62	1,125	0,00*	B
Madde 17	0,67	0,925	0,00*	A
Madde 18	0,78	0,595	0,03*	A
Madde 19	0,53	1,475	0,00*	B
Madde 20	1,12	-0,269	0,42	A

Not: * $p < 0,05$ ve $|\Delta MH| < 1$ ise A, $1 \leq |\Delta MH| < 1,5$ ise B, $|\Delta MH| \geq 1,5$ ise C

Tablo 17'ye göre yapılan analizde αMH istatistiği 13 maddede (Madde 1, 2, 4, 5, 6, 9, 12, 13, 15, 16, 17, 18, 19) manidar çıkmıştır ($p < 0,05$). ETS'nin ΔMH değeri için belirlediği kategori aralıklarına göre αMH istatistiği manidar çıkmış 13 maddenin 7 tanesi (Madde 2, 5, 9, 12, 13, 17, 18) A düzeyinde (ihmal edilebilir düzey), 4 tanesi (Madde 1, 15, 16, 19) B düzeyinde (orta düzey), 2 tanesi (Madde 4 ve 6) C düzeyinde (yüksek düzey) DMF göstermiştir. αMH istatistiği manidar çıkmayan ($p > 0,05$) 7 maddede yalnızca A düzeyi (ihmal edilebilir) DMF çıkmıştır. Uzmanlarca yanlı olduğu düşünülen maddeler tablosu (Tablo 11) ve MH yöntemine göre bulunan düzeyler birlikte incelendiğinde; 5 uzmanın yanlı olduğunu belirttiği 2. madde A düzeyinde, 4 uzmanın yanlı olduğunu belirttiği 6. madde C düzeyinde, 5 uzmanın yanlı olduğunu belirttiği 12. madde A düzeyinde, 5 uzmanın yanlı

olduğunu belirttiği 19. Madde B düzeyinde DMF göstermiştir. Ayrıca MH yöntemine göre uzmanlarca yanlış olduğu düşünülmeyen 1. maddede B düzeyinde, 4. maddede C düzeyinde DMF tespit edilirken 2 uzmanın yanlış olduğunu düşündüğü 16. maddede B, 1 uzman tarafından yanlış olduğu düşünülen 15. maddede ise B düzeyinde DMF tespit edilmiştir.

4.2. Araştırma Problemi 2'ye Yönelik Bulgular

Yerleşim bölgesine göre uzman görüşü alınarak yanlış olduğu belirlenmiş maddeler ile Lojistik Regresyon yöntemine göre TB ve TBO DMF gösteren maddeler tutarlı mıdır? Probleminin çözümü için SPSS 23 programı ile LR yöntemine yönelik analizler gerçekleştirilmiş ve elde edilen bulgular tablolarda verilmiştir.

Tablo 20

LR Analizi β Değerleri ve DMF Türü

Madde	Grup		Grup*Toplam		DMF Türü
	β	p	β	p	
Madde 1	-,123	,597	-,034	,512	-
Madde 2	-0,65	,646	-,036	,326	-
Madde 3	,081	,697	-,094	,076	
Madde 4	-,513	,000*	-,016	,661	TB DMF
Madde 5	-,002	,987	-,059	,136	-
Madde 6	-,542	,000*	-,039	,167	TB DMF
Madde 7	,122	,377	-,071	,047*	TBO DMF
Madde 8	,249	,155	-,031	,461	-
Madde 9	,032	,783	-,084	,006*	TBO DMF
Madde 10	,231	,108	-,062	,122	
Madde 11	,349	,021*	-,048	,232	TB DMF
Madde 12	-,132	,331	-,030	,808	-
Madde 13	,013	,936	,027	,564	-
Madde 14	,237	,048*	-,044	,150	TB DMF
Madde 15	-,118	,554	,004	,945	-
Madde 16	-,177	,224	-,014	,715	-
Madde 17	-,120	,290	-,074	,013*	TBO DMF
Madde 18	,065	,591	-,058	,096	-
Madde 19	-,234	,089	,012	,759	-
Madde 20	,406	,002*	,019	,589	TB DMF

Not: * $p < 0,05$

Tablo 19’da TB DMF ve TBO DMF istatistikleri anlamlı çıkan maddelerin DMF düzeyleri verilmiştir. Düzeyler belirlenirken hassas ölçümler yapabilmek adına Kelecioğlu ve Bakan Kalaycıoğlu (2008)’nun sınıflandırmaları esas alınmıştır.

Tablo 21

LR Analiz Sonuçları ve DMF Düzeyleri

Madde No	χ^2	p	ΔR^2	DMF Düzeyi
Madde 4	15,525	,000*	0,010	A
Madde 6	21,991	,000*	0,017	B
Madde 9	7,36	,007*	0,005	A
Madde 11	5,46	,019*	0,004	A
Madde 14	3,95	,047*	0,004	A
Madde 17	6,01	,014*	0,005	A
Madde 20	9,52	,002*	0,006	A

Not: * $p < 0,05$ ve $\Delta R^2 < 0,010$ ise A, $0,010 \leq \Delta R^2 < 0,020$ ise B, $0,020 \leq \Delta R^2$ ise C, Kelecioğlu, Bakan Kalaycıoğlu (2008)

Tablo 18’e göre yapılan analizde elde edilen β değerlerinin anlamlılık düzeylerine göre toplamda 7 maddenin (madde 4, 6, 9, 11, 14, 17 ve 20) DMF istatistikleri anlamlı çıkmıştır ($p < ,05$). Bu maddelerin 4 tanesi (Madde 4, 6, 11, 14 ve 20) TB DMF gösterirken, 2 tanesinde (Madde 9 ve 17) TBO DMF mevcuttur. Kelecioğlu ve Bakan Kalaycıoğlu(2008)’nun ΔR^2 değeri için belirlediği kategori aralıklarına göre istatistiği manidar çıkmış 7 maddenin 6 tanesi (Madde 4, 9, 11, 14, 17 20) A düzeyinde (ihmal edilebilir düzey), 1 tanesi ise (Madde 6) B düzeyinde (orta düzey) DMF göstermiştir. Uzmanlarca yanlı olduğu düşünülen maddeler tablosu (Tablo 11) ve LR yöntemine göre bulunan düzeyler birlikte incelendiğinde; 5 uzmanın yanlı olduğunu belirttiği 2, 12 ve 19. maddelerin LR yöntemine göre DMF istatistiği anlamlı düzeyde çıkmamış, 4 uzmanın yanlı olduğunu belirttiği 6. maddede B düzeyinde DMF tespit edilmiştir. LR yönteminin bulduğu DMF düzeylerinin genellikle A düzeyinde olduğu görülmüştür.

4.3. Araştırma Problemi 3'e Yönelik Bulgular

Yerleşim bölgesine göre uzman görüşü alınarak yanlış olduğu belirlenmiş maddeler ile SIBTEST yöntemine göre TB ve TBO DMF gösteren maddeler tutarlı mıdır? Probleminin çözümü için R’da bulunan “difR” paketi ile SIBTEST yöntemine yönelik hem TB DMF hem TBO DMF analizleri gerçekleştirilmiş ve elde edilen β_{cro} ve β_{uni} değerleri tabloda verilmiştir.

Tablo 22

SIBTEST Yöntemi İle DMF Analiz Sonuçları

Madde	β_{uni}	p	β_{cro}	p
Madde 1	0,023	0,07	0,023	0,07
Madde 2	0,057	0,01*	0,057	0,01*
Madde 3	0,024	0,14	0,024	0,14
Madde 4	0,151	0,00*	0,151	0,00*
Madde 5	0,067	0,01*	0,067	0,01*
Madde 6	0,177	0,00*	0,177	0,00*
Madde 7	0,047	0,03*	0,056	0,03*
Madde 8	0,017	0,35	0,017	0,35
Madde 9	0,064	0,02*	0,064	0,02*
Madde 10	0,027	0,19	0,027	0,19
Madde 11	0,001	0,96	0,007	0,93
Madde 12	0,068	0,00*	0,068	0,00*
Madde 13	0,050	0,01*	0,050	0,01*
Madde 14	0,017	0,52	0,003	0,53
Madde 15	0,048	0,01*	0,048	0,01*
Madde 16	0,072	0,00*	0,072	0,00*
Madde 17	0,093	0,00*	0,093	0,00*
Madde 18	0,057	0,03*	0,042	0,06
Madde 19	0,094	0,00*	0,094	0,00*
Madde 20	-0,013	0,57	0,013	0,57

Not: *p<,05

Chalmers (2017)’a göre eğer bir madde TB DMF’ye sahipse β_{cro} ve β_{uni} eşit olarak karşımıza çıkar ancak maddede TBO DMF varsa genelde $|\beta_{cro}| \geq |\beta_{uni}|$ durumu söz konusudur. Tablo 20 incelendiğinde istatistiği manidar çıkmış maddelerin tamamında

$\beta_{cro}=\beta_{uni}$ olduğu görülmektedir. Yani istatistiği manidar çıkmış maddelerin tamamında TB DMF mevcuttur.

Tablo 20'ye göre, yapılan SIBTEST DMF analizinde 13 maddede (Madde 2, 4, 5, 6, 7, 9 12, 13, 15, 16, 17, 18, 19) TB DMF manidar çıkmıştır ($p<0,05$).

Tablo 23
SIBTEST Yöntemi DMF Düzeyleri

Madde	β_{uni}	p	Düzye
Madde 2	0,057	0,01*	A
Madde 4	0,151	0,00*	C
Madde 5	0,067	0,01*	B
Madde 6	0,177	0,00*	C
Madde 7	0,047	0,03*	A
Madde 9	0,064	0,02*	B
Madde 12	0,068	0,00*	B
Madde 13	0,050	0,01*	A
Madde 15	0,048	0,01*	A
Madde 16	0,072	0,00*	B
Madde 17	0,093	0,00*	C
Madde 18	0,057	0,03*	A
Madde 19	0,094	0,00*	C

Not: * $p<0,05$ ve $\beta<0,059$ ise A, $0,059\leq\beta<0,088$ ise B, $\beta\geq 0,088$ ise C, Roussos ve Stout (1996)

DMF büyüklüğü yorumlaması için Roussos ve Stout (1996) tarafından önerilen sınıflama aralıklarına göre istatistiği manidar çıkmış 13 maddenin 5 tanesi (Madde 2, 7, 13, 15, 18) A düzeyinde (ihmal edilebilir düzey), 4 tanesi (Madde 5, 9, 12, 16) B düzeyinde (orta düzey), 4 tanesi (Madde 4, 6, 17 ve 19) C düzeyinde (yüksek düzey) DMF göstermiştir. İstatistiği manidar çıkmayan ($p>0,05$) 7 maddede yalnızca A düzeyi (ihmal edilebilir) DMF çıkmıştır. Uzmanlarca yanlış olduğu düşünülen maddeler tablosu (Tablo 11) ve SIBTEST yöntemine göre DMF bulunan maddelerin düzeyleri birlikte incelendiğinde; 5 uzmanın yanlış olduğunu

belirttiği 2. madde A düzeyinde, 4 uzmanın yanlı olduğunu belirttiği 6. madde C düzeyinde, 5 uzmanın yanlı olduğunu belirttiği 12. madde B düzeyinde, 5 uzmanın yanlı olduğunu belirttiği 19. Madde C düzeyinde DMF göstermiştir. Ayrıca SIBTEST yöntemine göre; uzmanlarca yanlı olduğu düşünülmeyen 4.maddede C düzeyinde ve 9.madde B düzeyinde, 17. maddede C düzeyinde ve 5.maddede ise B düzeyinde DMF tespit edilmiştir. 2 uzmanın yanlı olduğunu düşündüğü 16. maddede ise B düzeyinde DMF tespit edilmiştir.

4.4. Araştırma Problemi 4'e Yönelik Bulgular

Yerleşim bölgesine göre uzman görüşü alınarak yanlı olduğu belirlenmiş maddeler ile MH, LR ve SIBTEST yöntemleri karşılaştırıldığında hangi yöntem daha tutarlı sonuç vermektedir? Problemi için uzmanlarca yanlı olduğu düşünülen maddeler ve MH, LR ve SIBTEST yöntemlerine göre DMF göstermiş maddelerin düzeyleri tabloda birlikte verilmiştir.

Tablo 24

Uzman Görüşleri ve Yöntemlerce Bulunan DMF Düzeyleri

Madde	Yanlı Olduğunu Düşünen Uzman Sayısı	MH Yöntemi	LR Yöntemi	SIBTEST Yöntemi
Madde 1	-	B	-	-
Madde 2	5	A	-	A
Madde 3	1	-	-	-
Madde 4	-	C	A	C
Madde 5	-	A	-	B
Madde 6	4	C	B	C
Madde 7	2	-	-	A
Madde 8	-	-	-	-
Madde 9	-	A	A	B
Madde 10	1	-	-	-
Madde 11	-	-	A	-
Madde 12	5	A	-	B
Madde 13	-	A	-	A
Madde 14	-	-	A	-
Madde 15	1	B	-	A
Madde 16	2	B	-	B
Madde 17	-	A	A	C
Madde 18	-	A	-	A
Madde 19	5	B	-	C
Madde 20	-	-	A	-

Tablo 22’de uzman görüşlerine göre yanlı olduğu düşünülen maddeler ve MH, LR ve SIBTEST yöntemleri birlikte incelendiğinde uzmanların büyük çoğunluğu tarafından testin başlangıcında yanlı olarak yazılan 4 maddede yanlılık olduğunu tespit ettikleri (Madde 2, Madde 6, Madde 12 ve Madde 19) gözlenmiştir.

Bu maddeler yöntemlerle karşılaştırıldığında;

- 5 uzmanın yanlı olduğunu belirttiği 2. maddede LR yöntemine göre DMF manidar çıkmazken, MH ve SIBTEST yöntemlerine göre ise A düzeyi DMF tespit edilmiştir. Yani bu maddedeki DMF yöntemler tarafından tespit edilememiştir.
- 4 uzmanın yanlı olduğunu belirttiği 6. maddede MH ve SIBTEST yöntemlerine göre C düzeyi, LR yöntemine göre ise B düzeyi DMF tespit edilmiştir.

- 5 uzmanın yanlı olduğunu belirttiđi 12. maddede LR yöntemine göre DMF manidar çıkmazken, MH yöntemine göre A, SIBTEST yöntemine göre ise B düzeyinde DMF tespit edilmiştir.
- 5 uzmanın yanlı olduğunu belirttiđi 19. maddede LR yöntemine göre DMF manidar çıkmazken, MH yöntemine göre B, SIBTEST yöntemine göre ise C düzeyinde DMF tespit edilmiştir.
- Test geliştirme sürecinde yanlı olarak yazılmamış ve uzmanların tamamı tarafından da yanlı olarak görülmeyen 4.maddede MH ve SIBTEST yöntemleri C düzeyinde, LR yöntemi ise B düzeyinde DMF tespit etmiştir.

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

Bu bölümde, araştırmadan elde edilen bulgulara dayalı sonuçlar ve tartışmalara ayrıca bu sonuçlara dayalı önerilere yer verilmiştir.

5.1. Sonuç ve Tartışma

Yapılan çalışmalarda DMF'yi belirlemede etkili tek bir yöntem önerilememiş ve DMF belirlemede birden fazla yöntemin bir arada kullanılması tavsiye edilmiştir. Bu nedenle yöntemlerin etkinliğini karşılaştırmak için çok sayıda çalışma yapılmıştır. Yapılan araştırmalar sonucunda ise tek bir nihai sonuca ulaşılamamış ve farklı DMF belirleme yöntemlerinin farklı sonuçlar ortaya çıkardığı ve yöntemlerin arasındaki uyumun düşük ya da orta düzeyde olduğu tespit edilmiştir.

Bu araştırmada yerleşim bölgesine göre yanlı olarak yazılmış ve uzmanlarca da yanlılığı onaylanmış maddeleri bulunan bir testte, DMF belirleme yöntemlerinden Mantel-Haenszel (MH), Lojistik Regresyon (LR) ve Simultaneous İtem Bias Test (SIBTEST) yöntemlerinin belirtilen maddeleri tespit etme tutarlılıklarını belirlemek amaçlanmıştır. Araştırmanın amacı doğrultusunda 4 tane yanlı maddesi bulunan 5. Sınıf Sosyal Bilgiler testi geliştirilmiştir. Maddeler yazılırken yerleşim bölgelerine göre yanlı olmaları amaçlanarak

yazılmış ve uygulama kırsal ve kent merkezine yapılmıştır. Geçerlik ve güvenilirlik çalışmaları yapılan test uzman görüşüne sunulmuş ve uzmanlarca da bu 4 maddenin yanlı olduğu onaylanmıştır. Testin üç maddesi uzmanların tamamınca, bir maddesi ise uzmanların büyük çoğunluğunca yanlı olarak kabul edilmiştir. Maddelerin tamamında maddenin yanlı yazılma amacı ile uzmanların genel görüşleri incelendiğinde tutarlılık olduğu görülmüştür. Test geliştirme sürecinde yerleşim bölgesine göre yanlı olacağı düşünülerek yazılan ve uzman görüşüne sunulduğunda da uzmanlar tarafından yanlı olarak görülen bu 4 madde (2, 6, 12, 19) analizler gerçekleştirilirken yanlı maddeler olarak kabul edilmiş ve DMF analizlerinde özellikle bu maddelere odaklanılmıştır. Sonrasında maddelerin MH, LR ve SIBTEST yöntemleri ile analizleri gerçekleştirilerek DMF içerme durumları incelenmiştir. Bu araştırmanın öne çıkan sonuçları aşağıda maddeler halinde özetlenmiştir.

MH yöntemi ile elde edilen bulgular incelendiğinde uzmanların yanlı olarak kabul ettiği dört maddenin tamamında DMF tespit edilmiştir. DMF tespit edilen bu maddelerin iki tanesinde A düzeyinde, diğer maddelerin bir tanesinde B, diğerinde ise C düzeyinde DMF tespit edilmiştir. Ayrıca uzmanlarca yanlı olarak kabul edilmeyen bir maddede C düzeyinde, bir maddede B düzeyinde ve çok az uzman tarafından yanlı olarak kabul edilen iki maddede ise B düzeyinde DMF tespit edilmiştir. Yöntemin bulduğu DMF düzeyleri ve uzmanlarca onaylanan yanlı maddeler karşılaştırıldığında yöntemin yanlı olan 4 maddenin yalnızca ikisini tespit edebildiği, ayrıca yanlı olarak görülmeyen 4 maddede de B ve C düzeylerinde DMF tespit ettiği gözlemlenmiştir (A düzeyinde DMF gösteren maddeler ihmal edilmiştir). LR yöntemi ile elde edilen bulgular incelendiğinde uzmanların yanlı olarak kabul ettiği dört maddenin yalnızca bir tanesinde B düzeyinde DMF gözlenmemiştir. LR yöntemi genel olarak incelendiğinde, en hassas aralık düzeyi (Bakan Kalaycıoğlu & Kelecioğlu, 2011) kullanıldığında dahi DMF tespit edilen maddelerin neredeyse tamamında A düzeyinde DMF bulunmuştur. Yöntemin bulduğu DMF düzeyleri ve uzmanlarca onaylanan yanlı maddeler karşılaştırıldığında yöntemin yanlı olan dört maddenin yalnızca bir tanesini tespit edebildiği ortaya çıkmıştır. Alan yazındaki çalışmalar incelendiğinde LR yönteminin genelde DMF

belirlemede düşük düzeylerde kaldığı ifade edilmiştir. Çalışmada da alan yazındaki diğer çalışmalarla paralel sonuçlar ortaya çıkmıştır (Acar, 2008; Arslan, 2020; Kan vd., 2013).

SIBTEST yöntemi ile elde edilen bulgular incelendiğinde uzmanların yanlı olarak kabul ettiği dört maddenin tamamında DMF tespit etmiş, ve bu maddelerin bir tanesinde A, bir tanesinde B ve iki tanesinde de C düzeyinde DMF olduğu görülmüştür. Ayrıca uzmanlarca yanlı olarak kabul edilmeyen iki maddede C, iki maddede B, iki uzman tarafından yanlı kabul edilen bir maddede ise B düzeyinde DMF tespit edilmiştir (A düzeyinde DMF gösteren maddeler ihmal edilmiştir). Yöntemin bulduğu DMF düzeyleri ve uzmanlarca onaylanan yanlı maddeler karşılaştırıldığında yöntemin yanlı olan dört maddenin tamamını tespit edebildiği ve bunlarda da bir tanesi hariç diğerlerinde bulunan DMF'leri B ve C düzeylerinde tespit edebildiği görülmüştür. Ancak yöntemin yanlılık bulunmayan maddelerde de yüksek düzeyde DMF bulduğu tespit edilmiştir.

Yöntemlerin tamamı ve yanlı olduğu onaylanmış dört madde birlikte düşünüldüğünde LR yöntemi hassas aralıklar kullanıldığında dahi yanlı maddelerle tutarlı sonuçlar verememişken, MH yöntemi iki maddeyi (B ve C), SIBTEST yöntemi ise 3 maddeyi (C, B ve C) tespit edebilmiştir. Kullanılan yöntemler arasında bu çalışmada yanlı maddeleri en tutarlı tespit eden yöntem SIBTEST yöntemi olmuştur.

Araştırmanın önemi bölümünde bahsedilen ve bu araştırmanın yapılması gerekliliğini ortaya çıkaran ilk olumsuzluk yanlı olduğu halde DMF göstermediği için uzmanların görüşüne hiç sunulmayan ve testin geçerliğini gizlice tehdit etmeye devam eden maddelerin olmasıdır. Standart madde yanlılığı prosedürü işletildiğinde LR yöntemine göre bir madde, MH yöntemine göre iki madde, SIBTEST yöntemine göre ise üç madde uzman görüşüne sunulacaktı ve en iyi ihtimalle SIBTEST yöntemi kullanılsaydı dahi 1 maddenin yanlılığı tamamen göz ardı edilebilirdi.

Uzmanlar tarafından yanlı görülmediği halde DMF gösteren maddeler incelendiğinde ise; SIBTEST yöntemi uzmanların hiçbirisi tarafından yanlı olarak görülmeyen iki maddede C düzeyinde, iki maddede B düzeyinde, sadece iki uzman tarafından da yanlı kabul edilen bir

maddede ise B düzeyinde DMF tespit etmiştir. MH yöntemi ise hiçbir uzman tarafından yanlış görülmeeyen bir maddede C düzeyinde, iki maddede ise B düzeyinde DMF tespit etmiştir. Burada ortaya araştırmanın önemi kısmında bahsedilen ikinci olumsuzluk çıkmaktadır. Alandaki çalışmalar incelendiğinde uzmanların DMF göstermiş maddelere kesin olarak yanlış demeseler dahi olası yanlışlık kaynakları ile cevap verdikleri göz önünde bulundurulduğunda, standart madde yanlışlığı prosedürü işletildiği zamanlarda bu maddelerde DMF bulunmasının uzmanların yanlışlık görüşünü etkileyeceği ve maddeler yanlış olmasa bile olası yanlışlık kaynakları üzerinden maddelerin yanlış olarak ifade edilebileceği düşünülmektedir. SIBTEST yöntemine göre fazladan bulunmuş dört, MH yöntemine göre ise fazladan bulunmuş üç maddenin uzmanlar tarafından bir nedenle yanlış kabul edilip, yanlış olmadıkları halde testten çıkarılma ihtimalleri göz önünde bulundurulmalıdır.

Alan yazında bu üç yöntemin karşılaştırılması ile alakalı çalışmalar incelendiğinde, LR yöntemi ile MH ve SIBTEST yöntemine göre daha fazla DMF tespit edilmesinin beklenildiği ifade edilmektedir (Gierl, Khalig & Boughton, 1999; Hidalgo & Lopez-Pina, 2004). Bu çalışmada yöntemlerin DMF tespit etme durumları göz önünde bulundurulduğunda SIBTEST ve MH yöntemlerine göre 13 maddede, LR yöntemine göre ise 7 maddede DMF tespit edilmiştir. Alandaki beklentiye zıt yönde bir durum ortaya çıkmıştır. MH, LR ve SIBTEST analiz yöntemleri ile geniş ölçekli sınavlar üzerinde yapılan DMF araştırmalarında genellikle B düzeyinde DMF belirlenmiştir (Narayanan & Swaminathan, 1994; Rogers & Swaminathan, 1993). Çalışma incelendiğinde ise SIBTEST ve MH yöntemlerinin beklenen doğrultuda (B düzeyinde daha fazla) DMF gösterdiği ancak LR yönteminin genellikle A düzeyinden ileri gidemediği, yalnızca bir maddede B düzeyi DMF tespit edebildiği görülmüştür.

Son olarak bu çalışmayla Suna (2012) ve Demir (2013), çalışmalarında bu çalışmada olduğundan farklı olarak SIBTEST yönteminin MH ve LR yöntemlerine göre daha tutarsız sonuçlar verdiğini öne sürmüşlerdir.

Alan yazında bulunan ve bu çalışmanın sonucuyla uyumlu ve çelişkili olan farklı çalışmalar bulunması DMF yöntemlerinden hangisinin en doğru sonucu verdiğine dair genel bir durumun olmadığına bir kanıt niteliğindedir (Gierl, Khalig & Boughton, 1999). Ancak çalışma sonucu LR yönteminin DMF belirleme oldukça zayıf ve güvenilirliği düşük bir yöntem olduğuna işaret etmektedir. Ayrıca SIBTEST ve MH yöntemlerinin ise olması gerekenden fazla DMF belirlediğine yönelik bulgular belirlenmiştir.

5.2. Öneriler

Araştırmanın önerileri iki başlık altında incelenmiştir.

5.2.3. Araştırma Sonuçlarıyla İlgili Öneriler

- Araştırma sonucuna göre yanlış olduğu uzmanlarca onaylanmış maddeleri en iyi tespit eden yöntem SIBTEST yöntemi olduğu için araştırmacılara yanlış maddeleri belirlerken bu üç yöntem içerisinde SIBTEST yöntemini kullanmaları önerilebilir.
- LR yöntemleri alan yazında bulunan çalışmalar incelendiğinde ve bu çalışma sonuçları da gösterdiği üzere diğer yöntemlere göre daha zayıf DMF tespit etme gücüne sahiptir. Bu yüzden araştırmacıların tek yöntem kullanacakları durumlarda bu yöntemin LR yöntemi olmamasına dikkat etmeleri önerilebilir.

5.2.3. Araştırmacılara Yönelik Öneriler

Bu çalışmada yerleşim bölgesine göre yanlış olarak yazılmış ve uzman görüşü ile yanlışlığı doğrulanmış maddeleri bulunan testteki yanlış maddeleri yöntemlerin tespit etme tutarlılıkları incelenmiştir.

- Bu çalışmada alan yazında en çok tercih edilen KTK'ya dayalı yöntemlerden MH ve LR ile, MTK'ya dayalı yöntemlerden SIBTEST yöntemleri karşılaştırılmıştır. Diğer DMF

belirleme yöntemleri de eklenerek yöntemlerin karşılaştırılmasında daha geniş yol haritası sunacak bir çalışma yapılabilir.

- DMF analizlerine yönelik çok fazla çalışma bulunmasına karşı uzman görüşleri ile alakalı kısımlara çok fazla rastlanmamıştır. Bu çalışmada da yanlışlık konusuna hakim uzman bulmakta zorlanılmış ve alan uzmanlarına kısaca konu hakkında bilgi verilmiştir. Ancak uzman görüşlerindeki etkililiğin arttırılması adına uzmanlara yönelik yanlışlık çalışmalarına yer verilebilir.

- Yöntemlerin farklı örneklem büyüklüklerinde I. Tip hata oranlarının farklılaştığı göz önünde bulundurularak farklı örneklem büyüklükleri ile çalışma yapılabilir.

- Testte bulunan DMF'li madde sayısının yöntemlerin gücünü etkilediği göz önünde bulundurularak farklı yanlış madde sayısı içeren testler geliştirilip karşılaştırmalar tekrar yapılabilir.

KAYNAKLAR

- Abbott, M. L. (2007). A confirmatory approach to differential item functioning on an ESL reading assessment, *Language Testing* 24(1), 7-36.
<https://doi.org/10.1177%2F0265532207071510>
- Acar, T. (2008). *Maddenin farklı fonksiyonlaşmasını belirlemede kullanılan genelleştirilmiş aşamalı doğrusal modelleme, lojistik regresyon ve olabilirlik oranı tekniklerinin karşılaştırılması*. Doktora Tezi, Hacettepe Üniversitesi, Ankara.
- Ackerman, T. A. (1991). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, 29, 67-91. Retrieved from <https://eric.ed.gov/?id=ED333014>
- Ackerman, T. A., & Evans, J. A. (1992). *An investigation of the relationship between reliability, power, and the type I 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, San Francisco, CA, April 1992. Retrieved from <http://files.eric.ed.gov/fulltext/ED344937.pdf>
- Akbaş, U. (2019). Eğitimde Ölçme ve Değerlendirme. B. Çetin (Ed.), *Ölçme araç ve sonuçlarında bulunması gereken nitelikler* (s. 45-78). Ankara: Anı.
- Akcan, R. (2018). *2016 lisans yerleştirme sınavı ingilizce testinin madde yanlılığı açısından incelenmesi*. Yüksek Lisans Tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

- Angoff, W. H. (1993). Perspectives on differential item functioning methodology. P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 3–23). New Jersey: Lawrence Erlbaum.
- Ary, D., Jacobs, L. C., Sorensen, C. & Razavieh, A. (2010). *Introduction to research in education*. Belmont: Cengage
- Atalay, K. (2010). *PISA 2006 öğrenci anketinde yer alan tutum maddelerinin değişen madde fonksiyonu açısından incelenmesi*. Yüksek Lisans Tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Ateşok Deveci, N. (2008). *Üniversitelerarası kurul yabancı dil sınavının madde yanlılığı bakımından incelenmesi*. Doktora Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Atılgan, H. (Ed.). (2014). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı.
- Ayala, R. J. (2009). *Methodology in the social sciences: The theory and practice of item response theory*. New York: The Guilford.
- Ayan, C. (2011). *PISA 2009 fen okuryazarlığı alt testinin değişen madde fonksiyonu açısından incelenmesi*. Yüksek Lisans Tezi, Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Ayva Yörü, F. G. (2019). *Değişen madde fonksiyonu belirleme yöntemlerinin madde ve yetenek parametre kestirimine etkisi*. Yüksek Lisans Tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Bakan Kalaycıoğlu, D. (2008). *Öğrenci seçme sınavı'nın madde yanlılığı açısından incelenmesi*. Yüksek Lisans Tezi, Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: Pegem Akademi.

- 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>
- Büyüköztürk, Ş., Kılıç Çakmak. E., Akgün. Ö.E., Karadeniz. Ş. & Demirel. F. (2011). *Bilimsel araştırma yöntemleri*. Ankara: Pegem.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. California:Sage.
- Cappaert, K. (2014). *Dissecting the impact of DIF/DBF on ability estimation and person fit*. Doctoral Dissertation, The University of Wisconsin-Milwaukee in Educational Psychology.
- Chalmers, R. P. (2017). Improving the crossing-SIBTEST statistic for detecting non-uniform DIF. *Psychometrika*. 83(6). <http://dx.doi.org/10.1007/s11336-017-9583-8>
- Clauser, B., Mazor, K., & Hambleton, R. K. (1993). The effects of purification of the matching criterion on the identification of DIF using the Mantel-Haenszel procedure. *Applied Measurement in Education*, 6(4), 269-279. Retrieved from <https://eric.ed.gov/?id=EJ478666>
- Clauser, B. E. & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement Issues and Practice*, 17(1), 31-44. Retrieved from <https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1745-3992.1998.tb00619.x>
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Orlando: Holt, Rinehart and Winston.
- Çelik, H. E. & Yılmaz, V. (2013). *LISREL 9.1 ile yapısal eşitlik modellemesi: Temel kavramlar ve uygulamalar*. Ankara: Anı.
- 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.

- Çetin, B. (Ed.), (2019). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Anı
- Çokluk, O., Şekercioğlu, G. & Büyüköztürk, S. (2010). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve Lisrel uygulamaları*. Ankara: Pegem.
- Demir, S. (2013). *PISA 2009 matematik okuryazarlığı alt testinde bulunan maddelerinin mantel-haenszel, sıbtest ve lojistik regresyon yöntemleri ile değişen madde fonksiyonunun incelenmesi*, Yüksek Lisans Tezi, Abant İzzet Baysal Üniversitesi, Eğitim Bilimleri Enstitüsü, Bolu.
- Demir, S. & 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.
- Dogan, N., Hambleton, R. K., Yurtcu, M., & Yavuz, S. (2018). The comparison of differential item functioning predicted through experts and statistical techniques. *Cypriot Journal of Educational Sciences*, 13(2), 137–148. <https://doi.org/10.18844/cjes.v13i2.2427>
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenzel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35-66). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc
- 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. Retrieved from https://www.jstor.org/stable/1434554?seq=3#metadata_info_tab_contents
- Douglas, J. A., Roussos, L. A. & Stout, W. (1996). Item-bundle DIF hypothesis testing: Identifying suspect bundles and assessing their differential functioning. *Journal of Educational Measurement*, 33(4), 465-484. Retrieved from <https://www.jstor.org/stable/1435335?seq=1>

- 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.
- French, B. F., & Finch, W. H. (2013). Extensions of Mantel-Haenszel for multilevel DIF detection. *Educational and Psychological Measurement*, 73, 648-671. <https://doi.org/10.1177/0013164412472341>
- Gierl, M.J., Rogers, W.T. & Klinger, D.A. (1999). Using statistical and judgmental reviews to identify and interpret translation differential item functioning. *The Alberta Journal of Educational Research*. 45(4), 353-376. <https://doi.org/10.11575/ajer.v45i4.54709>
- Gierl, M. J., Gotzmann, A., & Boughton, K. A. (2004). Performance of SIBTEST when the percent of DIF items is large. *Applied Measurement in Education*, 17(3), 241-264. https://doi.org/10.1207/s15324818ame1703_2
- Gierl, M. J. (2005). Using dimensionality-based DIF analyses to identify and interpret constructs that elicit group differences. *Educational Measurement: Issues and Practice*, 24(1), 3-14. <https://doi.org/10.1111/j.1745-3992.2005.00002.x>
- Holland, P. W. & Wainer, H.(Eds.) (1983). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum.
- Holland, P. W. & Thayer, D. T. (1986). Differential item functioning and the Mantel-Haenszel procedure. *ETS Research Report Series*. 2, 1-24. Retrieved from <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2330-8516.1986.tb00186.x>
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the mantelhaenszel procedure. In H. Holland & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ: Erlbaum. Retrieved from <http://files.eric.ed.gov/fulltext/ED272577.pdf>
- Jodoin M. G. & Gierl, M. J. (2001).Evaluating Type I Error and Power Using an Effect Size Measure with the Logistic Regression Procedure for DIF Detection. *Applied Measurement in Education*, 14(4), 329-349. Retrieved from

https://www.tandfonline.com/doi/pdf/10.1207/S15324818AME1404_2?needAccess=true

- Kalaycıoğlu, D. B. & Berberoğlu, G. (2010). Differential item functioning analysis of the science and mathematics items in the University Entrance Examinations in Turkey. *Journal of Psychoeducational Assessment*, 29(5), 467-478.
- Kan, A. (2007). Test yansızlığı: Yabancı dil muafiyet sınavının cinsiyete ve bölümlere göre DMF analizi. *Eurasian Journal of Educational Research*, 29, 45-58.
- Kan, A., Sünbül, Ö. & Ömür, S. (2013). 6.-8. sınıf seviye belirleme sınavları alt testlerinin çeşitli yöntemlere göre değişen madde fonksiyonlarının incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 9(2), 207-222.
- Karami, H. (2012). An introduction to differential item functioning. *The International Journal of Educational and Psychological Assessment*, 11(2), 59-76. Retrieved from https://www.researchgate.net/publication/310153908_An_Introduction_to_Differential_Item_Functioning/link/5829c4e808aeff66e480f774/download
- Karakaya, İ. & Kutlu, Ö. (2012). Seviye Belirleme Sınavındaki Türkçe alt testlerinin madde yanlılığının incelenmesi. *Eğitim ve Bilim Dergisi*, 37(165).
- Kelecioğlu, H., Karabay, B., & Karabay, E. (2014). Seviye Belirleme Sınavı'nın madde yanlılığı açısından incelenmesi. *İlköğretim Online*, 13(3), 934-953.
- Meadows, D. H., (2008). *Thinking in Systems: A Primer*. White River Junction, VT: Chelsea Green. Retrieved from <https://wtf.tw/ref/meadows.pdf>
- Li, Z. (2009). *Impact of differential item functioning on statistical conclusions*. Doctoral Dissertation, University of British Columbia, Canada.
- Li, Z., & Zumbo, B. D. (2009). Impact of differential item functioning on subsequent statistical conclusions based on observed test score data. *Psicologica*: 30, 343-370. Retrieved from <https://www.redalyc.org/pdf/169/16911991011.pdf>

- Magis, D., Beland, B. Tuerlinckx, F., & De Boeck P. (2018). A General Framework and An R Package for The Detection of Dichotomous Differential Item Functioning. *Behavior Research Methods*, 42 (3), 847- 862.
- Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Routledge.
- Millsap, R. E. & Meredith, W. (1992). Inferential conditions in statistical detection of measurement bias. *Applied Psychological Measurement*, 16, 389-402.
<https://doi.org/10.1177/014662169201600411>
- Narayanan, P. & Swaminathan, H. (1994). Performance of The Mantel-Haenszel and Simultaneous Item Bias Procedures for Detecting Differential. *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.
- Osterlind, J. S. (1983). *Test item bias*. London: Sage Publications. Retrieved from <https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=0&sid=0e08925f-17b7-43fa-82d3-f3f84b76e726%40redis>
- Osterlind J. S. & Everson T. H. (2009). *Differential Item Functioning*, London: Sage Publications, Inc.
- Özer Özkan, Y. (2019). Eğitimde ölçme ve değerlendirme. B. Çetin (Ed.), *Ölçme ve değerlendirme ile ilgili temel kavramlar* (s. 1-22). Ankara: Anı.
- Pallant, J. (2007). *SPSS Survival manual: A step by step guide to data analysis using SPSS for windows (3. Baskı)*. Berkshire, England: Open University.
- Popham, W.J. (2014). *Classroom assessment: What teachers need to know*. Boston: Pearson.
- 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.

- Roussos, L. A. & Stout, W. F. (1996). 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.
- Ryan, C. H. (2008). *Using hierarchical generalized linear modeling for detection of differential item functioning in a polytomous item response theory framework: An evaluation and comparison with generalized Mantel-Haenszel*. Doctoral Dissertation Georgia State University, College of Education, Atlanta.
- Roth, W.-M., Oliveri, M. E., Sandilands, D., Lyons-Thomas, J., & Ercikan, K. (2013). Investigating sources of differential item functioning using expert think-aloud protocols. *International Journal of Science Education*, 35, 546–576. Retrieved from http://web.uvic.ca/~mroth/PREPRINTS/TAPs_301r.pdf
- Sireci, S. G., & Rios, J. A. (2013). Decisions that make a difference in detecting differential item functioning. *Educational Research and Evaluation*. 19(2-3), 170-187. Retrieved from <https://www.tandfonline.com/doi/pdf/10.1080/13803611.2013.767621>
- Suna, H.E. (2012). *TIMSS 2007 fen bilimleri testindeki maddelerin dil ve cinsiyet yanlılığı açısından incelenmesi*. Yüksek Lisans Tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Swaminathan, H. & Rogers, H. J. (1990). Detecting Differential Item Functioning Using Logistic Regression Procedures. *Journal of Educational Measurement*, 27(4), 361-370.
- Tan, Ş . (2020). *Öğretimde Ölçme ve Değerlendirme*. Ankara: Pegem
- Tavşancıl, E. (2002). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel
- Tecim, V., (2004). Sistem yaklaşımı ve soft sistem düşüncesi, *D.E.Ü. İ.İ.B.F.Dergisi* 19(2), 75-100. <https://docplayer.biz.tr/7873605-Sistem-yaklasimi-ve-soft-sistem-dusuncesi-vahap-tecim.html> sayfasından erişilmiştir.

- Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of item response theory in the study of group differences in trace lines. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp.147-169). Hillside, NJ: Erlbaum.
- Uyar, Ş. ve Kaya-Uyanık, G. (2016). PISA 2012 bilişsel maddelerinin kültüre göre değişen madde fonksiyonu bakımından incelenmesi. *Eğitim ve Öğretim Araştırmaları Dergisi*, 5(3), 230-240.
- Uzun, N. & Gelbal, S. 2017. PISA Fen başarı testinin madde yanlılığının kültür ve dil açısından incelenmesi. *Kastamonu Eğitim Dergisi*, 25(6), 2427-2446.
- Walker, C. M. (2011). What's the DIF? Why differential item functioning analyses are an important part of instrument development and validation. *Journal of Psychoeducational Assessment*,29(4),364-376. Retrieved from <https://journals.sagepub.com/doi/pdf/10.1177/0734282911406666>
- Wiberg, M. (2007). *Measuring and detecting differential item functioning in criterion-referenced licensing test:A theoretic comparison of methods*. Umea University. Retrieved from http://coshima.davidrfikis.com/EPRS9360/Articles/Wiberg_07.pdf
- Yıldırım, H. (2015). *2012 yılı Seviye Belirleme Sınavı'nın madde yanlılığı açısından incelenmesi*. Yüksek Lisans Tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara
- Yıldırım, H. & Berberoğlu, G. (2009). Judgmental and statistical DIF analyses of the PISA-2003 mathematics literacy items. *International Journal of Testing*, 9, 108–121.
- Yıldız, F., & Okyay, P. (2017). Karıştırıcı etkenlerin kontrolünde mantel haenszel testi. *Sağlık ve Toplum Dergisi*, 1, 53- 60.
- Yurdugül, H. (2003). *Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı'nın madde yanlılığı açısından incelenmesi*. Doktora tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

- Zenisky, A.L., Hambleton, R.K., & Robin, F. (2003). DIF detection and interpretation in large-scale science assessments: Informing item writing practices. *Educational Assessment*, 9(1-2), 61-78.
- Zumbo, B. D., & Gelin, M. N. (2005). A matter of test bias in educational policy research: Bringing the context into picture by investigating sociological/community moderated (or mediated) test and item bias. *Journal of Educational Research and Policy Studies*, 5(1), 1-23. Retrieved from <https://files.eric.ed.gov/fulltext/EJ846827.pdf>
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF)*. Ottawa: National Defense Headquarters.
- Zumbo, B. D. & Thomas, D. R. (1996). *A measure of DIF effect size using logistic regression procedures*. Paper presented at the National Board of Medical Examiners, Philadelphia.

EKLER



Ek 1. Veri Elde Edilmesinde Kullanılan Başarı Testi

5. SINIF SOSYAL BİLGİLER KAZANIM DEĞERLENDİRME TESTİ

1. Ülkemizde en fazla can ve mal kaybına neden olan doğal afettir. Ülkemizin hemen hemen her yerinde meydana gelir. Yer kabuğundaki ani kırılmalar nedeniyle oluşur.

Yukarıdaki metinde hangi doğal afetten bahsedilmektedir?

- A) Yangın B) Sel C) Heyelan D) Deprem

2. Ayşe Hanım ve Erdinç Bey, çocukları Mehmet ve Zeynep ile Sivas'ta yaşamaktadır. Ayşe Hanım, İstanbul'dan bir iş teklifi almıştır ve ailecek İstanbul'a taşınmışlardır. Ailenin üyeleri yaşadıkları bu durum karşısında farklı şekillerde etkilenmiştir. Erdinç Bey, eşi için gelen iş teklifi için sevinirken Mehmet, arkadaşlarından ayrıldığı için üzgündür. Zeynep ise yeni bir ev, yeni bir çevre edindikleri için çok heyecanlıdır.

Buna göre ailenin yaptığı bu şehir değişikliğinin aile bireyleri üzerindeki etkisiyle ilgili verilen bilgilerden hangisi kesinlikle doğrudur?

- A) Ayşe Hanım olaydan hiç etkilenmemiştir.
B) Ailenin tamamı bu durumdan dolayı üzülmüştür.
C) Olayın birden fazla sonucu olmuştur.
D) Erdinç Bey Sivas'tan ayrıldığına üzülmemiştir.

3. Aşağıdakilerden hangisi sosyal bilgiler dersinin kişinin gelişimine sağladığı faydalardan biri değildir?

- A) Dünyadaki farklı kültürleri tanımak
B) Sorumluluk almaktan kaçınmak
C) Çevremize karşı daha duyarlı olmak
D) Hak ve sorumluluklarımızı bilmek

4. Aşağıda Yasin'in bir günüyle ilgili bilgiler verilmiştir.

09.00 - 10.00	Kız kardeşimle birlikte annemin kahvaltı hazırlamasına yardım ettim.
12.00 - 13.00	Öğretmenimin verdiği ödevleri tamamladım.
18.00 - 19.00	Babam bize arkadaşları ile olan anılarını anlattı.
20.00 - 21.00	Kuzenlerimle birlikte akşam yemeği yedik.

Buna göre Yasin bu zaman dilimlerinde hangi rolde bulunmamıştır?

- A) Kardeş B) Öğrenci
C) Arkadaş D) Akraba

5. Thales (Tales), Pythagoras (Pisagor), Hipokrates (Hipokrat), Homeros, Diogenes (Diyojen) gibi ünlü bilim insanları İyonlar Dönemi'nde yetişmiştir.

Bu bilgide İyonların aşağıdaki özelliklerinden hangisi vurgulanmaktadır?

- A) Kullandıkları alfabe
B) İnanç sistemleri
C) Ekonomik etkinlikleri
D) Bilimde ileri olmaları

6. Ülkemiz çeşitli doğal güzelliklere ve tarihî eserlere sahiptir.

Aşağıda verilen tarihi eserlerden hangisi ilimiz merkezinde bulunmamaktadır?

A)



Ulu Cami

B)



Kale Cami

C)



Kongre Binası

D)



Uç Kümbetler

7. Bir ile üniversite açılması ildeki yaşamı büyük oranda etkiler. İlde kafe, öğrenci yurdu gibi işletmelerin sayısı artarken bu işletmeler insanlar için yeni iş imkânları oluşturur. Üniversite ilin hem eğitimini hem de insanların yaşamını etkiler.

Yukarıda verilen metin aşağıdakilerden hangisi ile daha çok bağlantılıdır?

- A) Olaylar toplum içinde hazır bulunur.
B) Olayların birbiriyle ilişkisi yoktur.
C) Olaylar bir arada yaşamının sonucudur.
D) Bir olayın, birden çok sonucu olabilir.

8. Aşağıdakilerden hangisinde “eğitim hakkının ihlali” söz konusudur?

- A) Okul çağındaki bir çocuğun okula gönderilmemesi
B) Çocukların düşüncelerini özgürce ifade etmelerinin engellenmesi
C) Bir ailenin konutuna izinsiz müdahale edilmesi
D) Yeni doğan bebeğin nüfus kütüğüne kaydedilmemesi

9. Mardin’de gerçekleştirilen bir arkeolojik çalışma sırasında günümüzden 2800 yıl önce hazırlanmış bir tapu belgesi bulunmuş. Bir meyve bahçesine ait olan bu tapu belgesi, bir kil tabletin üzerine Asur yazısıyla yazılmış. Bu belgenin şimdiye kadar bulunmuş en eski tapu belgesi olduğu belirtiliyor.

Buna göre Asurlular ile ilgili aşağıdaki yargılardan hangisine ulaşılamaz?

- A) Yazıyı kullanmışlardır. B) Toplumda sınıf farklılıkları görülür.
C) Yerleşik bir yaşam sürmüşlerdir. D) Mülkiyet hakkı vardır.

10. Nüfusun dağılışını etkileyen doğal unsurlar vardır. Örneğin dağlık ve engebeli arazilerde nüfus daha azken düzlük alanlarda fazlalaşmaktadır.

Bu bilgide nüfus dağılışını etkileyen doğal unsurlardan hangisi anlatılmıştır?

- A) Bitki örtüsü B) İklim türü C) Yeryüzü şekilleri D) Toprak verimliliği

11. Türkiye fiziki haritası incelendiğinde haritanın doğusunda kahverengi tonunun daha fazla kullanılmasının nedeni aşağıdakilerden hangisidir?

- A) Yükseltinin fazla olması B) İklimin soğuk olması
C) Nüfusun seyrek olması D) Ova ve yaylaların çok olması

12. Ahmet ve ailesi internet üzerinden eve yemek siparişi vermiştir. Sipariş esnasında güvenli alışveriş yapmak için uyguladıkları yöntemlerden bazıları şunlardır:

- I. Kapıda ödeme yapmak
- II. Yanlış adres bilgisi vermek
- III. İnternet sitesinde içeriği detaylı biçimde anlatılmış yemeği sipariş vermek
- IV. İnternetteki yorumları en iyi olan firmadan sipariş vermek

Buna göre davranışlarından hangilerini uygularlarsa internet üzerinden güvenli alışveriş yapmış olurlar?

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

13. Geçmişte düğünler evlerin önünde 2-3 gün süren kır düğünü şeklinde yapılırdı. Günümüzde ise düğün salonlarında yaklaşık 3-4 saat süren düğünler yapılmaktadır.

Bu bilgilerden yola çıkarak aşağıdaki ifadelerden hangisine ulaşılabılır?

- A) Düğünlerde müzik eşliğinde türküler söylenir.
- B) Kültürel unsurlar zaman içinde değişebilmektedir.
- C) Düğünler ülkenin her yerinde aynı şekilde yapılır.
- D) Kültürel özellikler farklı bölgelerde aynı olabilir.

14. Aşağıdaki görsellerden hangisi tarihi - kültürel bir öge değildir?

A)



Görün Gölü

B)



İzmir Saat Kulesi

C)



Çini Tabak

D)



Ayasofya Camii

15. Aşağıdakilerden hangisi teknolojinin hayatımızdaki olumsuz etkilerine örnek verilebilir?

- A) Nurcan'ın dünyadaki haberleri internetten takip etmesi
- B) Ahmet'in yurt dışında yaşayan arkadaşlarıyla görüntülü konuşması
- C) Aytaç'ın sosyal medyada çok vakit geçirerek ailesini ihmal etmesi
- D) Zeynep'in iş seyahatlerini hızlı bir biçimde uçakla yapması

16. Mehmet'in babası çocukları için 'Bilim Çocuk' dergisine abone olmuştur.

Buna göre eve gelen dergiler sonucunda Mehmet ve kardeşlerinde aşağıdaki değişimlerden hangisi beklenmez?

- A) Araştırma yapmaya istekli olurlar.
- B) Bilimi sevmeye başlarlar.
- C) Bilgileri sorgulamadan kabul etmeye başlarlar.
- D) İyi birer gözlemci olurlar.



17. Aşağıda bazı doğal afetlerle ilgili haberler yer almaktadır.



Aşağıda verilenlerden hangisi bu doğal afetlerden birine önlem olamaz?

- A) Ormanlık alanlarda ateş yakmamalıyız.
- B) Ormanlık alanlardaki ağaçların kesilmesini önlemeliyiz.
- C) Doğal bitki örtüsünün korunmasını sağlamalıyız.
- D) Yerleşim yerlerini sağlam zeminler üzerine inşa etmeliyiz.

18. Ahmet Bey bir ev yaptırmak istiyor. Konuştuğu firma Ahmet Bey'e yaptıracığı evle alakalı şu önerilerde bulunuyor:



Ahşabı kolay temin edebildiğimiz için evinizi ahşaptan yapalım. Ayrıca ilimiz her mevsim yağış aldığı için evinizin çatısını geniş turalım.

Buna göre Ahmet Bey'in yaşadığı il aşağıdakilerden hangisi olabilir?

- A) Sivas
- B) İstanbul
- C) Sinop
- D) Diyarbakır

19. Sivas'ta bulunan PrimeMall Alış Veriş Merkezi'ne (AVM) ait kat planı ile ilgili aşağıda bazı bilgiler verilmiştir.

- Birinci katta yalnızca giyim mağazaları vardır.
- İkinci katta yalnızca giyim mağazaları ve ayakkabı mağazası vardır.
- Üçüncü katta yemek mekanları ve kat planında belirtilmemiş ancak dolu olduğu bilinen farklı türde dükkanlar bulunmaktadır.

Buna göre verilenlerden hangisi doğrudur?

- A) Çocuk oyun alanına gitmek isteyen bir kişi üçüncü kata çıkabilir.
- B) Yemek yiyecek bir kişi ikinci kata çıkabilir.
- C) Bilgisayar almak isteyen bir kişi birinci katta alışveriş yapabilir.
- D) Tabak almak isteyen bir kişi ikinci kata çıkabilir.

20. Doğu Anadolu Bölgesi engebeli ve yüksek bir arazi yapısına sahiptir. Bu nedenle Doğu Anadolu Bölgesi'nde ulaşım zordur. Kışlar uzun ve soğuk, yazlar kısadır. Tarım arazileri dar ve küçüktür. Ancak çayır ve otlaklar geniştir ve süt-entegre tesisleri yaygındır.

Bu metne göre Doğu Anadolu Bölgesi'nin temel geçim kaynağı aşağıdakilerden hangisi olabilir?

- A) Tarım
- B) Hayvancılık
- C) Madencilik
- D) Sanayi

Ek 2. Madde Yanlılığı Anketi

Değerli katılımcı;

Yanlılık, bir maddenin farklı gruplardan gelen ancak benzer yetenek düzeyinde bulunan bireyler tarafından doğru yanıtlanma olasılığının farklılık göstermesidir. Yapılan analizlere göre, farklı yerleşim bölgelerinde yaşayan (kırsal/ kent merkezi) benzer yetenek seviyesindeki öğrencilerin maddeleri doğru yanıtlama oranları arasında farklılıklar bulunabilir.

Bu durum maddenin yanlış olabileceğine dair bir kanıttır. Hazırlanmış olduğum başarı testi maddelerini göz önünde bulundurarak, yerleşim bölgesine göre yanlışlık oluşturabileceğini düşündüğünüz maddeleri ve hangi grup lehine yanlışlık oluşturabileceğini tabloda işaretleyiniz. Yanlılık oluşturabileceğini düşündüğünüz maddelerde yanlışlığın hangi nedenden kaynaklandığını ise ikinci kısımdaki formda doldururuz. Yanlılığın listelenen nedenler dışında bir nedenden kaynaklandığını düşünüyorsanız “diğer” kısmına ekleyebilirsiniz.

Katılımlarınız için teşekkür ederim.

Nurcan CEYLAN

Katılımcı Bilgileri

Adı, Soyadı:

Görev Yeri:

Unvanı:

Alanı:

Testte yanlışlık olduğunu düşündüğünüz maddeleri hangi grup lehine yanlışlık oluşturduğuna göre ‘X’ ile işaretleyiniz. (Benzer yetenek seviyelerinde bulunan hangi grup bulunduğu yerleşim yerinden dolayı maddeyi diğer gruba göre daha kolay yanıtlar? Eğer hiçbirinde olmadığını düşünüyorsanız maddeyi boş bırakınız)

Madde No/ Grup	Kırsal	Kent Merkezi
Madde 1		
Madde 2		
Madde 3		
Madde 4		
Madde 5		
Madde 6		
Madde 7		
Madde 8		
Madde 9		
Madde 10		
Madde 11		
Madde 12		
Madde 13		
Madde 14		
Madde 15		
Madde 16		
Madde 17		
Madde 18		
Madde 19		
Madde 20		

İşaretlediğiniz her madde ile alakalı aşağıdaki formdan maddeler için ayrı ayrı doldururuz.

Madde No:

(X ile işaretleyiniz)	Kent Merkezi	Kırsal
Sizce madde hangi yerleşim bölgesi lehine çalışmaktadır?		

Yanılığın nedeni aşığıdakilerden hangisi olabilir? (Birden fazla işaretleyebilirsiniz.)		
	Evet	Sözcük veya ifadeyi yazınız
Maddede kullanılan sözcük ya da ifade nedeni ile		
	Evet	İçeriğı yazınız
Yerleşim bölgesine bağılı olarak bir grubun maddenin içeriğine aşına olması nedeni ile		
	Evet	Konuyu yazınız
Yerleşim bölgelerinin daha yakın olduğı konuların soru kapsamında içerilmesi nedeni ile		
	Evet	Tarzı yazınız
Maddede kullanılan tarzın yerleşim bölgesi özelliklerine göre farklı olması nedeni ile (sözel anlatımın uzun olması, grafik içermesi vs. gibi)		
	Evet	Açıklayınız
Diğer:		

Ek 3. Gazi Üniversitesi Etik Komisyonu Görüşü ve Onayı

Evrak Tarih ve Sayısı: 11.04.2022-E.336916



**T.C.
GAZİ ÜNİVERSİTESİ
Öğrenci İşleri Daire Başkanlığı**

Sayı : E-17311665-044-336916
Konu : Anketler (Nurcan CEYLAN)

11.04.2022

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : a) 17.03.2022 tarihli ve 80287700-044- 318297 sayılı yazı.
b) 29.03.2022 tarihli ve 46680793 sayılı yazı.

Enstitünüz Eğitim Bilimleri Ana Bilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı yüksek lisans öğrencisi Nurcan CEYLAN'ın, Prof. Dr. Bayram ÇETİN danışmanlığında yürüttüğü "Uzman Görüşüne Göre Yanlı Olduğu Belirlenen Maddelerin Değişen Madde Fonksiyonu Belirleme Yöntemleri İle Tutarlılığının İncelenmesi" konulu tez çalışması kapsamında Sivas İl Milli Eğitim Müdürlüğüne bağlı Sivas İlinde bulunan resmi ortaokullarda öğrenim gören 5. sınıf öğrencilerine uygulama yapmasına izin verilmesine ilişkin ilgi (a) yazınız Sivas Valiliğine iletilmiş olup, alınan ilgi (b) cevabı yazı ve ekleri ilişikte gönderilmiştir.

Bilgilerinizi ve gereğini rica ederim.

Ek:İlgi (b) yazı ve ekleri (8 sayfa)

Ek 4. Sivas Milli Eğitim Müdürlüğü Onayı

Evrak Tarih ve Sayısı: 11.04.2022-E.336916

Evrak Tarih ve Sayısı: 11.04.2022-E.336870



T.C.
SİVAS VALİLİĞİ
İl Millî Eğitim Müdürlüğü

Sayı : E-92255297-605.01-46680793
Konu : Araştırma İzni (Nurcan CEYLAN)

29.03.2022

DAĞITIM YERLERİNE

- İlgi : a)Gazi Üniversitesi Rektörlüğü Öğrenci İşleri Daire Başkanlığının 18/03/2022 tarihli ve E-173116 65-044-318488 sayılı yazısı.
b)Valilik Makamının 28/03/2022 tarihli ve E-92255297-605.01-46563389 sayılı onayı.
c)Millî Eğitim Bakanlığı Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 21/01/2020 tarihli ve 81576613-10.06.02-E.1563890 sayılı 2020/2 no'lu genelgesi.

Gazi Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Ana Bilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı yüksek lisans öğrencisi Nurcan CEYLAN'ın, "Uzman Görüşüne Göre Yanlı Olduğu Belirlenen Maddelerin Değişen Madde Fonksiyonu Belirleme Yöntemleri ile Tutarlılığının İncelenmesi" konulu çalışması kapsamında, onaylı bir örneği müdürlüğümüzde muhafaza edilen veri toplama araçlarının, gönüllülük esas olmak kaydıyla ilimiz genelinde bulunan ortaokullarda uygulanması valilik makamının ilgi (b) onayı ile uygun görülmüş olup onay örneği yazımız ekinde gönderilmiştir.

Söz konusu araştırma çalışmasının bitiminde, araştırma yapan kişi tarafından sonuç raporunun bir örneğinin CD ortamında müdürlüğümüze gönderilmesi hususunda;

Bilgilerinizi ve gereğini arz ve rica ederim.

Ergüven ASLAN
Millî Eğitim Müdürü

Ek:

- 1-İlgi (b) Onay Örneği (1 Sayfa)
2-Veri Toplama Araçları

Dağıtım:

Gereği:

-Ankara Gazi Üniversitesi Rektörlüğü
(Öğrenci İşleri Daire Başkanlığı)

Bilgi:

-16 İlçe Millî Eğitim Müd.
-Mrk. Ortaokul Müd.

Bu belg

Ek 5. Gönüllü Katılım Formu

GAZİ ÜNİVERSİTESİ ETİK KOMİSYONU FORM-2

Rev-2
25.01.2022



T.C. GAZİ ÜNİVERSİTESİ ETİK KOMİSYONU

KATILIMCILAR İÇİN BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU

Sizi, Gazi Üniversitesi Etik Komisyonu'ndan tarih / sayı ile izin alınan ve Prof. Dr. Bayram ÇETİN danışmanlığında Nurcan CEYLAN tarafından yürütülen "Uzman Görüşüne Göre Yanlı Olduğu Belirlenen Maddelerin Değişen Madde Fonksiyonu Belirleme Yöntemleri İle Tutarlılığının İncelenmesi" başlıklı araştırmaya davet ediyoruz. Bu çalışmaya katılmak tamamen gönüllülük esasına dayanmaktadır. Çalışmaya katılmama veya katıldıktan sonra herhangi bir anda çalışmadan çıkma hakkına sahipsiniz. Bu çalışmaya katılmanız için sizden herhangi bir ücret istenmeyecektir. Çalışmaya katıldığınız için size bir ödeme yapılmayacaktır. Çalışmadan elde edilecek bilgiler tamamen araştırma amacı ile kullanılacak olup kişisel bilgileriniz gizli tutulacaktır.

*Gazi Üniversitesi Etik Komisyon izini alındıktan sonra doldurularak kullanılacaktır.

Araştırmanın Amacı	Bazı test maddeleri aynı yetenek düzeyinde olan bireyler tarafından farklı şekillerde cevaplanabilir. Böyle maddelere 'Yanlı Madde' adı verilir. Yanlı olan maddeleri belirlerken kullanılan farklı istatistiksel yöntemler bulunmaktadır. Bizim amacımız öğrencilerin teste verdiği cevaplar sonucunda elde ettiğimiz veriler vasıtası ile bu istatistiksel yöntemleri birbirleri ile karşılaştırarak hangi yöntemin daha tutarlı sonuçlar verdiğini hesaplamaktır.
Araştırmanın Yöntemi	Çalışmada istatistiksel yöntemleri karşılaştırmak için kırsal ve kent merkezindeki öğrencilerin testteki maddelere verdikleri cevaplar kullanılacaktır. Bu cevaplar 5.sınıf Sosyal Bilgiler dersi için üretilen ve 1.dönem konularını içeren 20 soruluk bir test ile sağlanacaktır. Türkiye'de yapılan geniş ölçekli ulusal sınavlarda alt test uzunlukları genelde 12 ve 25 soru arasında değişmektedir Bu yüzden bu çalışmada 20 soru esas alınacaktır.
Araştırmanın Öngörülen Süresi (Başlama ve Bitiş Tarihi Başvurudaki Başlangıç ve Bitiş Tarihi ile Uyumlu Olmalıdır.)	01.04.2022 - 30.06.2022
Araştırmaya Katılması Beklenen Katılımcı/Gönüllü Sayısı	1500
Araştırmanın Yapılacağı Yerler	Sivas Kent Merkezi ve Kırsal Ortaokulları 5. Sınıfları
Görüntü ve/veya ses kaydı alınacak mı?	Evet ☐ Hayır ☒

Tablo katılımcıların anlayabileceği biçimde, akademik dil kullanılmadan yazılacaktır.

KATILIMCI BEYANI

Yukarıda amacı ve içeriği belirtilen bu araştırma ile ilgili bilgiler tarafıma aktarıldı. Bu bilgilerden sonra araştırmaya katılımcı olarak davet edildim. Bu çalışmaya katılmayı kabul ettiğim takdirde gerek araştırma yürütülürken gerekse yayımlandığında kimliğimin gizli tutulacağı konusunda güvence aldım. Bana ait verilerin kullanımına izin veriyorum. Araştırma sonuçlarının eğitim ve bilimsel amaçlarla kullanımı sırasında kişisel bilgilerimin dikkatle korunacağı konusunda bana yeterli güven verildi. Araştırmanın yürütülmesi sırasında herhangi bir sebep göstermeden çekilebilirim. Araştırma için yapılacak harcamalarla ilgili herhangi bir parasal sorumluluk altına girmiyorum. Bana herhangi bir ödeme yapılamayacaktır. Araştırma ile ilgili bana yapılan tüm açıklamaları ayrıntılarıyla anlamış bulunmaktayım. Bu çalışmaya hiçbir baskı altında kalmadan kendi bireysel onayım ile katılıyorum. İmzalı bu form kağıdının bir kopyası bana verilecektir.

Araştırma yürütücüsü(Tez çalışmalarında Danışman tarafından imzalanacaktır.)

GAZİ ÜNİVERSİTESİ ETİK KOMİSYONU FORM-2

Rev-2
25.01.2022

Adı ve Soyadı	Bayram ÇETİN	Tarih ve İmza
Adres ve telefonu	Katılımcı	
Adı ve Soyadı		Tarih ve İmza
Adres ve telefonu	Velayet veya Vesayet Altındaki Katılımcılar için Veli/Vasi	
Adı ve Soyadı		Tarih ve İmza
Adres ve telefonu		

Ek 6. Tez Benzerlik Raporu

UZMAN GÖRÜŞÜNE GÖRE YANLI OLDUĞU BELİRLENEN MADDELERİN DEĞİŞEN MADDE FONKSİYONU BELİRLEME YÖNTEMLERİ İLE TUTARLILIĞININ İNCELENMESİ

ORJİNALLİK RAPORU

% 6	% 6	% 1	% 1
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

BİRİNCİL KAYNAKLAR

1	www.openaccess.hacettepe.edu.tr:8080 İnternet Kaynağı	% 3
2	acikbilim.yok.gov.tr İnternet Kaynağı	% 1
3	dergipark.org.tr İnternet Kaynağı	% 1
4	9lib.net İnternet Kaynağı	% 1

Alıntıları çıkart üzerinde
Bibliyografyayı Çıkart üzerinde

Eşleşmeleri çıkar < %1



GAZİLİ OLMAK AYRICALIKTIR...