

T.C
HASAN KALYONCU ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME TEZLİ YÜKSEK LİSANS
PROGRAMI

**MADDE YANLILIĞININ UYGULAMA ÖNCESİNDE UZMAN GÖRÜŞÜNE VE
SONRASINDA KLASİK TEST KURAMINA GÖRE BELİRLENMESİ**

YÜKSEK LİSANS TEZİ

HAZIRLAYAN
BEHÇET ÖZTİMUR

GAZİANTEP - 2022

T.C
HASAN KALYONCU ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME TEZLİ YÜKSEK LİSANS
PROGRAMI

**MADDE YANLILIĞININ UYGULAMA ÖNCESİNDE UZMAN GÖRÜŞÜNE VE
SONRASINDA KLASİK TEST KURAMINA GÖRE BELİRLENMESİ**

YÜKSEK LİSANS TEZİ

HAZIRLAYAN
BEHÇET ÖZTİMUR

TEZ DANIŞMANI
DOÇ. DR. UFUK AKBAŞ

GAZİANTEP - 2022



**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ MÜDÜRLÜĞÜNE
YÜKSEK LİSANS KABUL VE ONAY FORMU**

Eğitimde Ölçme ve Değerlendirme Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Tezli Yüksek Lisans Programı öğrencisi Behçet ÖZTİMUR tarafından hazırlanan “Madde Yanlılığının Uygulama Öncesinde Uzman Görüşüne ve Sonrasında Klasik Test Kuramına Göre Belirlenmesi” başlıklı tez, 13/01/2022 tarihinde yapılan savunma sınavı sonucu başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Görevi

Unvanı, Adı ve Soyadı

İmzası:

Kurumu/Üniversitesi

Tez Danışmanı

Doç. Dr. Ufuk AKBAŞ

Jüri Başkanı

Prof. Dr. Şener BÜYÜKÖZTÜRK

Jüri Üyesi

Doç. Dr. İbrahim YILDIRIM

Bu tez Enstitü Yönetim Kurulunca belirlenen yukarıdaki jüri üyeleri tarafından uygun görülmüş ve Enstitü Yönetim Kurulu kararı ile onaylanmıştır.

Prof. Dr. İbrahim Halil GÜZELBEY
Enstitü Müdürü

TEZ ETİK VE BİLDİRİM SAYFASI

Yüksek Lisans Tezi olarak sunduğum “Madde Yanlılığının Uygulama Öncesinde Uzman Görüşüne ve Sonrasında Klasik Test Kuramına Dayalı Olarak İncelenmesi” başlıklı çalışmanın tarafımca, bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu ve bunlara atıf yapılarak yararlanmış olduğumu belirtir ve onurumla doğrularım. 26/12/2021

Behçet ÖZTİMUR



ÖNSÖZ

Her konuda beni dinleyen ve hiçbir zaman yardım ve desteğini esirgemeyen hem öğretmenim hem de danışmanım olan Sn. Doç.Dr. Ufuk AKBAŞ hocama,

Her zaman bizi dikkatle ve ilgiyle dinleyen, kapısının bizlere her daim açık olduğunu bildiğimiz Sn. Prof. Dr. Şener BÜYÜKÖZTÜRK hocama,

En zor anlarda dahi zaman ayıran, her durumda desteğini esirgemeyen Sn. Merve YILDIRIM SEHERYELİ hocama,

Bu günlere gelmemde büyük emekleri olan Sn. Doç. Dr. Yeşim ÖZER ÖZKAN, Sn. Dr. Ersoy KARABAY, Sn. Doç. Dr. Burcu GÜRKAN, Sn. Doç. Dr. Halil İbrahim SARI hocalarıma,

Çalışmamda her konuda yardım eden Şehit Samet Kaymakçı İmam Hatip Ortaokulu idaresine ve öğretmenlerine,

Çalışmam için gerekli izinleri sağlayan Gaziantep İl Milli Eğitim Müdürlüğü'ne,

Çalışmamda yardımcı olan okullar ve okulların yetkili amirlerine,

Süreç boyunca her durumda yardımlarını esirgemeyen sevgili arkadaşlarım Tayfun ÖZTÜRK, Ahmet Gürkan KARATAŞ ve Serhat TAN'a,

Uzman olarak destek sağlayan, her anlamda destek olan Sn. Dr. Öğr. Üyesi Safiye BİLİCAN DEMİR'e, Sn. Doç. Dr. Özen YILDIRIM'a, Sn. Dr. Öğr. Üyesi Fatih KEZER'e ve aynı zamanda hocam olan Sn. Dr. Öğr. Üyesi Ersoy KARABAY'a

Beni her konuda destekleyen, dünyalar güzeli biricik eşim Fatma BİLGİN ÖZTİMUR'a,

Her zaman yanımda olduğunu hissettiren annem Hatice ÖZTİMUR'a, babam Murat ÖZTİMUR'a, kardeşlerim Fethi ÖZTİMUR ve Mustafa ÖZTİMUR'a teşekkürlerimi sunarım.

Gaziantep, 2022

Behçet ÖZTİMUR

ÖZET

Çalışmanın amacı sınavlardan sonra analiz edilen madde yanlılığının sınav uygulanmadan tespit edilip/ edilemeyeceği üzerine bir araştırma gerçekleştirmektir. Çalışma Güneydoğu bölgesinde bulunan bir ilde yer alan 8.sınıf düzeyindeki öğrencilerle yapılmıştır. Araştırma için önceden yapılan araştırmalar incelenerek yanlı madde konuları belirlenmiş ve bu madde içerikleri ile ilgili kazanımlara uygun maddeler hazırlanmıştır. Uzman görüşleri alındıktan sonra gerekli izinler alınarak 245 kişilik deneme uygulaması yapılmıştır. Deneme uygulamasından sonra veriler AFA ile analiz edilmiş ve faktör yükü .30 altında olan maddeler testten çıkarılmıştır. Nihai test ile 881 kişilik bir öğrenci grubuna ana uygulama yapılmıştır. Yapılan uygulamadan sonra hangi maddelerin yanlı ve hangi düzeyde olduğu ile alakalı analizler yapılmadan uzman görüşleri alınmıştır. Uzman görüşleri alındıktan sonra Klasik Test Kuramı yanlılık belirleme yöntemlerinden olan Mantel – Haenszel ve Lojistik Regresyon yöntemi ile madde yanlılığı incelenmiştir. Çıkan analiz sonuçlarına göre beş maddede yanlılık olduğu saptanmıştır. Uzman grubu cinsiyete göre yanlılık beklenen on maddeden beş maddeyi ve düzeylerini yüksek oranda doğru tahmin etmiştir. Düzey olarak ise uyumsuzluklar olsa da yüksek bir düzeyde ortak görüş bildirilmiştir.

Anahtar Kelimeler: Madde Yanlılığı, Klasik Test Kuramı, Lojistik Regresyon, Mantel – Haenszel, Değişen Madde Fonksiyonu

ABSTRACT

The aim of the study is to carry out a research on finding the item bias analyzed after the exams without taking the exam. The study was conducted with students at 8th grade in South East region. For the research, the previous researches were examined and the biased item topics were determined and the items related to these topics were prepared. After getting expert opinions, necessary permissions were obtained and a trial application was made for 245 people. After the trial application, the data were analyzed by EFA and items with a factor load below .30 were excluded from the test. With the final test, the main application was made to a group of 881 students. After the application, expert opinions were taken without analyzing which items were biased and at which level. After receiving expert opinions, item bias was examined with Mantel – Haenszel and Logistic Regression method, which is one of the Classical Test Theory bias determination methods. According to the results of the analysis, it was determined that there was a bias in five items. According to the results of the analysis, it was determined that there was a bias in five items. Experts have correctly estimated the gender element in the wrong items at the rate of 80%. As for the level, it can be said that it has been estimated at a high level, although there are disagreements.

Keywords: Item Biased, Classical Test Theory, Logistic Regression, Mantel Haenszel, Differential Item Functioning

İÇİNDEKİLER

Sayfa No.

ÖNSÖZ.....	i
ÖZET	ii
ABSTRACT	iii
TABLolar LİSTESİ.....	vi
ŞEKİLLER LİSTESİ.....	vii
KISALTMALAR LİSTESİ	viii

BİRİNCİ BÖLÜM

GİRİŞ.....	1
1.1. Problem Durumu	1
1.2. Problem Cümlesi	2
1.3. Araştırmanın Önemi	2
1.4. Araştırmanın Amacı	3
1.5. Sayıtlar.....	3
1.6. Araştırmanın Sınırlılıkları.....	3

İKİNCİ BÖLÜM

KURAMSAL ÇERÇEVE	3
2.1. Ölçme ve Değerlendirme.....	4
2.2. Hata Türleri.....	4
2.3. Güvenirlik	6
2.4. Geçerlik	7
2.5. Yanlılık	8
2.6. Değişen Madde Fonksiyonu (DMF).....	9
2.7. Klasik Test Kuramı.....	10
2.7.1 Mantel Haenszel Yöntemi.....	11
2.7.2. Lojistik Regresyon	12
2.8. İlgili Araştırmalar	13
2.8.1. Yurtiçinde yapılan yanlılık çalışmaları	13
2.8.2. Yurtdışında yapılan yanlılık çalışmaları	17

ÜÇÜNCÜ BÖLÜM

YÖNTEM.....	20
3.1. Araştırmanın Modeli.....	20
3.2. Çalışma Grubu.....	21
3.3. Veri Toplama Araçları.....	22

3.4. Veri Toplama Süreci.....	23
3.5. Verilerin Analizi	23
DÖRDÜNCÜ BÖLÜM	
BULGULAR	24
4.1. Betimsel İstatistik Analiz Sonuçları	24
4.2. Açımlayıcı Faktör Analizi Sonuçları	24
4.3. Madde istatistikleri	30
4.4. Mantel – Haenzsel Analizi Sonuçları	32
4.5. Lojistik Regresyon Analiz Sonuçları.....	34
4.6. Mantel – Haenzsel ve Lojistik Regresyon DMF Karşılaştırma.....	37
4.7. Uzman Görüş Tablosu	38
BEŞİNCİ BÖLÜM	
YORUM	42
ALTINCI BÖLÜM	
SONUÇ VE ÖNERİLER.....	45
KAYNAKÇA	47
EKLER.....	55
Ek 1. Pilot uygulama Soruları.....	55
Ek.2. AYSE Başvuru Formu	63
Ek.3. Sosyal Bilimler Enstitü Müdürlüğü Dilekçe	64
Ek.3. Gaziantep Valiliği İl Milli Eğitim Müdürlüğü İzin Yazısı.....	65
Ek.4. Araştırma İzni Taahhütnamesi	66
Ek.5. Örneklem Toplanacak Okullar Listesi	67
Ek. 6. İl Milli Eğitim'in Gaziantep Valiliğine Yazdığı İzin Yazısı	68
Ek.7. Sınav Yönergesi	69

TABLÖLAR LİSTESİ

Sayfa No.

Tablo 1. Mantel Haenszel 2x2 Olasılık Tablosu	12
Tablo 2. Örneklem Tablosu	21
Tablo 3. Madde Kazanımları Tablosu	22
Tablo 4. Betimsel İstatistik Analiz Sonuçları	24
Tablo 5. Faktör Yükleri Tablosu	25
Tablo 6. Özdeğer Analiz Sonuçları	26
Tablo 7. Çoklu Doğrusallık Varsayımı Analiz Sonuçları.....	29
Tablo 8. Madde İstatistik Analiz Sonuçları	30
Tablo 9. Mantel – Haenszel Analizi Sonuçları.....	32
Tablo 10. Lojistik Regresyon Analizi Sonuçları	34
Tablo 11. Zumbo and Thomas ve Jodoin and Gierl indeksi analiz sonuçları	36
Tablo 12. MH – LR Karşılaştırma Tablosu	37
Tablo 13. Uzman Görüş Tablosu.....	39
Tablo 14. Karşılaştırmalı Analiz Sonuç Tablosu.....	42

ŞEKİLLER LİSTESİ

Sayfa No.

Şekil 1. Tek Boyutlu DMF Örneği.....	10
Şekil 2. Tek Boyutlu Olmayan DMF Örneği	10
Şekil 3. Araştırmanın birinci aşaması adımları	20
Şekil 4. Araştırmanın ikinci aşaması adımları	20
Şekil 5. Yamaç – Birikinti Grafiği (Scree Plot)	27
Şekil 6. Mantel Haenszel Analizi Sonuç Şekli	34
Şekil 7. Lojistik Regresyon Analizi Sonuç Şekli.....	35



KISALTMALAR LİSTESİ

AFA	:	Açımlayıcı Faktör Analizi
CMH	:	Cochran Mantel Haenszel
DMF	:	Değişen Madde Fonksiyonu
JG	:	Jodoin and Gierl
KMO	:	Kaiser-Meyer-Olkin
KTK	:	Klasik Test Kuramı
LR	:	Lojistik Regresyon
MH	:	Mantel – Haenszel
p	:	Madde Güçlük İndeksi
PGT	:	Psikometrik Giriş Testi
rjx	:	Madde Ayırt Edicilik Gücü İndeksi
VIF	:	Variance Inflation Factor (Varyans Büyütme Faktörü)
ZT	:	Zumbo and Thomas
$\bar{x}$	:	Aritmetik Ortalama

BİRİNCİ BÖLÜM

GİRİŞ

Bu bölümde araştırmanın problem durumuna, araştırmanın amacı ve önemine, problem cümlesine, sayıtlara, sınırlılıklara ve tanımlara ilişkin bilgiler verilmiştir.

1.1. Problem Durumu

Eğitim, insanların öğrenme kabiliyetinin oluşmaya ve gelişmeye başlaması ile ortaya çıkan ve hayatı devam ettikçe de devam eden bir süreçtir. Eğitim; bireylerin toplum yaşamında yer edinebilmeleri için gerekli bilgi, beceri ve konseptleri elde etmelerine, belirgin özelliklerini geliştirmelerine okulda veya okul dışında, doğrudan veya dolaylı olarak yardım etme, görgü olarak adlandırılır (TDK,2021). İnsan eğitilebilir bir varlıktır. Bu sebeple de önemli bir arayışı ve bilgi birikimine ihtiyaç duygusunu beraberinde getirmiştir (Aydın,2013:5) Eğitim, kişilerin hal ve hareketinde istendik farklılıklar meydana getirme sürecidir ve sistemsal bir durumdur (Turgut ve Baykul, 2015:1). Sistem; Bir sonuç elde etmeye yarayan yöntemler düzeni olarak adlandırılır (TDK, 2021). Ulusların en önemli unsurlarından olan eğitim ise sistem kelimesi ile kullanılarak ilerlemektedir. Eğitimin amacı kişinin sahip olduğu özelliklerin farkına varması, bu özellikleri geliştirmesi ve nihayet olarak kendini gerçekleştirme temeline dayanır. Eğitim sistemine giren her kişi bu sisteme girdi olarak kaydedilir. Her sistemde olduğu gibi bu sistemin de girdileri, çıktıları ve arada geçen süre olarak süreci ve en sonunda ise kontrolü vardır (Baykul, 2015:83). Eğitimde, bireylere bazı davranışlar kazandırma, kazandırılan davranışları geliştirme, kazanılmış yanlış davranışları ise düzeltme amaçlanır. Bu durumları etkileyen, bazen davranışların değişmesine yardım eden, kolaylaştırıcı, bazen değişmede direnmeye sebep olan zorlaştırıcı, bazen de etkinliklerin yapılabilmesi için gerekli olan etkenler vardır. Bunların hepsi eğitimin bir unsurunu oluşturur. Bu unsura eğitim sisteminde girdi adı verilir. Yeni davranış biçimlerinin oluşturulması, istenmeyen davranış biçimlerinin silinmesi, hatalı olduğu düşünülen davranış biçimlerinin de düzeltilmesi için yapılan etkinlikler bu modelde süreç ismini alır. Eğitimin sonunda bireyler, eğitimin başındaki durumlarına göre artık farklıdırlar; önceden sahip oldukları bazı davranış biçimleri değişmiş, bazı yeni davranış biçimleri kazanmışlardır. İşte bu değişen yeni davranışlara eğitim sisteminin çıktısı denir. Çıktılar sistemin ortaya koymuş olduğu ürünlerdir. Eğitim etkinlikleri sonunda, amaçlanan bazı davranış biçimlerinin istenen düzeyde olup olmaması durumu kontrol mekanizması ile yapılır. Eğitim sisteminin kontrolü değerlendirme unsuru tarafından yerine getirilir. Buradan hareketle değerlendirmenin rolünün, sistemin onarılmasını sağlamak olduğu sonucuna varılabilir (Turgut

ve Baykul, 2015: 65-67). Eğitim ise bir sistem olarak ölçme ve değerlendirme kavramlarından bağımsız düşünülemez. Eğitimde kazandırılmak istenen davranışların amaçları neticesinde ölçme yapılır. Ölçme sonucu ile bireyde ortaya çıkan veya kazandırılmak istenen davranışlar ile ilgili ne derecede amaca ulaşıldığı ortaya çıkarılmaya çalışılır. Eğitim bir sistem olarak düşünüldüğü zaman ölçmenin ve sonrasında değerlendirmenin bu sistemin en önemli unsurlarından birisi olduğu söylenebilir. Bu amaçla ölçme ve değerlendirme sürecinin hata türlerinden arınık olması istenen bir durumdur. Ölçme hataları güvenirlik ve geçerliği etkileyen unsurlardır. Bu hatalardan birisi ise madde yanlılığıdır. Yanlılık; belirli bir grubun üyeleri için test çıkarımlarının anlamını çarpıtan sistematik hata olarak düşünülebilir (Camilli and Shepard, 1994:8). Yanlılık bulunurken klasik test kuramına göre kestirim yöntemlerinden olan Mantel – Haenszel ve Lojistik Regresyon analiz yöntemleri duruma göre kullanılır.

1.2. Problem Cümlesi

Bir testteki maddelerin cinsiyete göre yanlılık gösterip göstermeme durumunun sınav uygulamadan önce uzman görüşlerine dayalı olarak tespit edilip edilemeyeceğinin incelenmesidir.

1.3.Araştırmanın Önemi

Türkiye gibi genç nüfusun yoğunlukta olduğu ülkelerde tanımaya ve yerleştirmeye yönelik sınav uygulamaları ile ölçme ve sonrasında ise değerlendirme yapıldığı görülmektedir. Ölçmenin yapıldıktan sonraki aşaması ise değerlendirme aşamasıdır. Değerlendirme kavramı ölçme sonuçlarını belirli bir ölçüte dayanarak ölçümü olan nitelik hakkında yargıya varma sürecidir. Değerlendirmenin güvenilir olması için ölçümü yapılan durum, geçerli bir ölçüt ve yapılan işlemlerde yanlış ve hataya yer verilmemesi yerinde bir karar vermek için uygun olacaktır (Turgut ve Baykul, 2015:3). Bu durumlarda ise uygulanan sınavlara bazı değişkenler etki edebilmektedir. Bu değişkenlerden birisi de hata türleridir. Bütün uygulanan sınav veya etkinliklerde hata olabilmektedir (Güler, 2017:30). Bu hata türlerinden birisi de yanlılık olarak bilinen kavramdır. Yanlılık; belirli bir grubun üyeleri için test çıkarımlarının anlamını çarpıtan sistematik hata olarak düşünülebilir (Camilli and Shepard, 1994:8). Ayrıca değerlendirmeyi isabetli yapabilmek için, ölçme sonuçlarının hatalarını azaltmak gerekmektedir (Turgut ve Baykul, 2015:3). Araştırmanın önemi ise yukarıda anlatılan sistematik hatayı sınav uygulanmadan önce bulmaya çalışmaktır. Yanlılık genel olarak değerlendirme uygulamasının üzerinden zaman geçtikten sonra analiz edilir. Yapılan bu işlem gecikmiş bir durum tespitidir. Önemli olan yanlılık içeren veya içermeme durumları olan maddelerin uygulanmasından önce

tespit etmek veya bu maddeleri kullanmamaktır. Bu bağlamda, bu araştırmanın problemi, bir testin uygulanmasından önce cinsiyete göre yanlılık içeren maddelerin tespit edilip edilemeyeceğinin incelenmesidir.

1.4.Araştırmanın Amacı

Araştırmanın amacı; sınavlar uygulandıktan sonra tespit edilen madde yanlılığının sınavlar uygulanmadan önce uzman görüşlerine dayalı olarak tespit edilip edilemeyeceğidir. Yapılan sınavların hatalardan arınlık derecesinin artırılmasına katkı sağlamaktır.

1.5. Sayıtlar

Bütün testlerin aynı şartlar altında uygulandığı kabul edilmiştir. Öğrencilerin motive bir şekilde bütün soruları bilgilerini açığa çıkaracak şekilde ciddiye ve özenle cevapladıkları kabul edilmiştir.

1.6.Araştırmanın Sınırlılıkları

Bu çalışma Güneydoğu Anadolu Bölgesi'nde yer alan bir ilin ilçelerinde yer alan seçili okullardaki öğrenciler ile sınırlandırılmıştır.

Türkçe “*okuduğunu anlama*” testi olarak uygulanmıştır.

Yanlılığa ilişkin incelemeler cinsiyet değişkeniyle; Klasik Test Kuramı'na göre MH ve LR yöntemleriyle gerçekleştirilmiştir.

Yalnızca belirlenen devlet okullarından veri toplanmıştır.

Yüz yüze uygulama olduğu için Covid – 19 pandemisinden etkilenilmiştir.

İKİNCİ BÖLÜM

KURAMSAL ÇERÇEVE

Araştırmanın bu bölümünde ölçme ve değerlendirme tanımı yapılarak önemine yer verilecek. Ölçme ve değerlendirmede olmaması gereken unsur olan hata türlerine, hata türlerinden sonra ise ölçme araçlarında olması istenen özellikler olan güvenirlik ve geçerlik kavramlarına yer verilecektir. Testteki geçerlik ve güvenirliği olumsuz anlamda etkileyen yanlılık kavramına ve yanlılığın oluşturduğu değişen madde fonksiyonuna değinilecektir. Ölçme ve değerlendirme kuramlarından klasik test kuramına ve klasik test kuramı ile yanlılık kestirim yöntemleri olan MH ve LR yöntemleri ile ilgili ifadeler yer verilecektir. Son olarak ise yurtiçinde ve yurtdışında yapılan ilgili araştırmalara yer verilecektir.

2.1.Ölçme ve Değerlendirme

Ölçme en geniş tanımla bir niteliği gözlemlemek ve gözlemlenen durumun sonucunu sembol veya sıfatlarla ifade etmektir (Baykul, 2015:86). Ölçme bir tasvir işlemidir. Ortada olan bir durumun veya varlığın nitelenmek amacıyla bir ölçme aracıyla ölçülmesi işlemidir denilebilir. Ölçme işleminin sonunda tespit edilen niteliğe veya niceliğe bir derece veya bir sıfatın uygun görülmesidir (Baykul, 2015:84). Ölçme işlemi ile ilgili birçok tanım yapılabilir. Ölçme, varlıkların nicelik ve nitelik yönünden betimlenmesi ve bu özelliklere sahiplik belirleme işlemidir. Bu belirleme işleminde görülmeyen ama gerçekte var olan durumların, sembollerle ifade edilmesidir (Özçelik, 2016:13). Ölçme işlemi, bir performansın niceliğini tespit etme ya da gözlemlenen başarıma rakam ile derece verme sürecidir (Kilmen, 2014:44). Ölçme durumundan hemen sonra ise değerlendirme aşamasına geçilir. Değerlendirme ölçülen niteliğin veya niceliğin bir kriter ile karşılaştırılarak sonuca varma sürecidir. Değerlendirmenin nesnel bir şekilde yapılabilmesi için bir ölçüte ihtiyaç duyulmaktadır. Elde olan ölçüt ile karşılaştırılan ölçme sonucuna göre bir karar verilir. Bu işleme ise değerlendirme denilmektedir (Güler, 2017:12). Değerlendirme kavramı ölçme sonuçlarını belirli bir ölçüte dayanarak ölçümü olan nitelik hakkında yargıya varma sürecidir. Değerlendirme diğer bir tanımla, ölçme sonuçlarının bir kıstas ile karşılaştırma yapılarak nihai bir karar verme sürecidir. Diğer bir tanıma göre ise amaçlanan davranışın müfredat ve öğretim programı yoluyla uygulama aşamasında ne derecede gerçeğe ulaştığını tespit işlemidir. Ama amaçlanan davranışlar istenen değişiklikler olması gerektiğinden, bireylerin davranış kalıplarında istenen değişiklikler olduğundan, değerlendirme aşaması bu değişimlerin görünüşte ne kadar gerçekleştiğini belirleme sürecidir (Tyler, 2014: 94-95). Değerlendirme, belirli bir eğitimle ilgili yargı verilebilmesi için bilgi edinme sürecidir (Nitko and Brookhart, 2014: 25). Değerlendirme için öncelikle ölçüt gereklidir. Bu ölçüt ile ölçme sonucu karşılaştırılarak bir sonuca varılır ve bu aşamaya değerlendirme işlemi denir (Güler, 2017:12). Değerlendirmenin güvenilir olması için ölçümü yapılan durum, geçerli bir ölçüt ve yapılan işlemlerde yanlış ve hataya yer verilmemesi yerinde bir karar vermek için uygun olacaktır. Ayrıca değerlendirmeyi isabetli yapabilmek için, ölçme sonuçlarının hatalarını azaltmak gerekmektedir (Turgut ve Baykul, 2015:68).

2.2.Hata Türleri

Ölçme ve değerlendirme unsurlarının hata ve hata türlerinden arınık olması istenir. Bu aşamada hata kavramına bakmak, hata kaynaklarını incelemek ve hata türlerine bakmak yerinde olacaktır. Ölçme sonuçlarının gerçek değerinden olan farkına ölçmede hata adı verilir (Güler, 2017:30). Klasik test teorisinde, gözlenen puan içinde, gerçek puan ile ölçmede yapılan tesadüfi

hata dan gelen bir miktar hata puanı bulunur. M evreninden tesadüfi olarak seçilen X=gözlenen puanı, T=gerçek puanı ve E=hata puanını temsil eder.

Formül şu şekildedir;

$$X = T + E \text{ (Baykul, 2015:107-108).}$$

Ölçme işlemlerinde hata olmaması söz konusu olamaz. Hatalar teorik olarak önlenabilir ama pratikte mümkün değildir. Her ölçme işleminde mutlaka hata bulunur (Güler, 2017). Bu hatalar çeşitli sebeplerle ortaya çıkmaktadır. Bu duruma hata kaynakları denir. Hata kaynakları şu şekildedir;

Ölçmeyi gerçekleştiren şahıstan ortaya çıkan hatalar (Örneğin; Öğretmenin puanlama yaparken ki dikkatsizliği, psikolojisi, fizyolojik eksikliği vb. etkenler)

Ölçmenin yapıldığı şahıstan kaynaklanan hatalar (Öğrencinin uygulama esnasındaki kaygı düzeyi, rahatsızlıkları, kopya çekmesi, şans faktörü vb. etkenler)

Ölçmenin yapıldığı ortamdan kaynaklanan hatalar (Gürültü, ışığın yetersiz olması, sınıfın sıcaklığının çok fazla ya da çok düşük olması vb. etkenler)

Ölçme aracından kaynaklanan hatalar (Yazılı yoklamalarda puanlamanın sübjektif olması, çoktan seçmeli testlerde şans başarısının olması vb. etkenler) (Güler, 2017:31).

Eğitimde yer alan ölçmelerde veya psikolojide yer alan ölçmelerde hangi özellik ölçülürse ölçülsün mutlaka ölçme sonuçlarına bir miktar hata payı olacaktır. Hatasız ölçme söz konusu değildir (Köse, 2014:87). Yapılan ölçme işlemlerinde ne kadar her şey doğru ve eksiksiz görünse de hatalar meydana gelmektedir. Bu hataları minimuma indirmek esastır. Bu bakımdan ölçmede hata denklemi aşağıdaki gibidir;

Hata= Gözlenen ölçme sonucu – Gerçek ölçme sonucu (Güler, 2017:30). Ölçme literatüründe hata türleri üç başlık altında incelenebilir;

Sabit Hata: Yapılan ölçmeler sonucunda bütün ölçümlerde eşit derecede yer alan hata türüne sabit hata denir (Turgut ve Baykul, 2015:120). Ölçümler sonucu miktarı değişmeyen hata türüne sabit hata denir. Bu hatanın genele yayılımı söz konusudur; yani ölçme sonuçlarına aynı oranda hata karışmasıdır (Köse, 2014:90). Diğer bir tanıma göre ise ölçmeler arasında miktarı ve yönü değişmeyen ve çıkış yeri belirlenebilen hata türüdür (Güler, 2017: 31).

Sistematik Hata: Ölçüm yapılan büyüklüğe, ölçme işlemine ya da ölçme koşullarına bağımlı olarak görülen hata türüne sistematik hata denir. İstatistik alanında bu türlü hatalara yanlışlık denir. (Turgut ve Baykul, 2015:121). Sistematik hata türünde ölçme işlemine karışan hatanın miktarı değişkenlik gösterir. Hatanın miktarı ölçme işlemine giren bütün unsurlardan (puanlayıcı, ölçme aracı, testle ölçülen özellik, uygulama yapılan ortam) kaynaklı olabilir. Bu hata türünde genellikle hata puanlayıcı veya testten kaynaklıdır (Köse, 2014:90). Diğer bir tanımla yapılan ölçmelerde ölçmeden ölçmeye miktarı veya yönü değişen, çıkış yeri belirlenebilen hatalardır (Güler, 2017:31).

Tesadüfi Hata: Bu hatalar çıkış yerleri pek bilinmeyen hata türü olup ölçmenin nihai sonucuna rastgele karışırlar. (Turgut ve Baykul, 2015:121). Tesadüfi hata, ölçme sonuçlarına karışan, çıkış yönü ve ölçme sonuçlarını olumsuz veya olumlu yönde etkilediği tahmin edilemeyen hatalardır. Tesadüfi hata güvenilirliği düşürerek etkilerken, sabit ve sistematik hatalar ise geçerliği düşürmektedir (Köse, 2014:90). Yapılan ölçmeler sonucunda miktarı ve yönü değişen ama çıkış yeri belirlenemeyen hata türüdür (Güler, 2017:31). Bir ölçme işlemi sonucunda bu bahsedilen hataların biri bulunabileceği gibi duruma göre iki türden hatta üç türden hata birlikte bulunabilir. Sabit ve sistematik hataların miktar ve kaynakları bellidir. Fakat tesadüfi hatalar kaynakları kesin olarak bilinmeyen ve miktarları hesaplanamayan, ancak kestirilebilen hata türüdür. Sabit ve sistematik hatalar önceden yapılacak işlemlerle teorik olarak da olsa önlenabilir, fakat tesadüfi hata için böyle bir durumdan söz edilemez. Bu sebepten dolayı tesadüfi hataların bulunması ölçme alanında önem kazanır. Tesadüfi hatalar ölçmede güvenilirlik kavramını doğurmuştur (Turgut ve Baykul, 2015:121-122).

2.3.Güvenirlik

Bir test uygulandığı zaman, testi yapan birey, aynı kişilere benzer koşullarda uygulandığında aynı sonuçları alabileceğinin güvencesini ister. Testten alınan bu puanların kararlılığı durumuna güvenilirlik denir (Crocker and Algina, 2008: 105-106). Güvenirlik bir başka deyişle ölçme sonuçlarının tesadüfi hatalardan arınmışlık derecesini belirten bir kavramdır (Turgut ve Baykul, 2015: 123). Yani tesadüfi hatalar ne kadar az ise o ölçme sonucu o oranda güvenilir. Ölçme hatası ile güvenilirlik arasında ters yönlü bir ilişki söz konusudur (Yılmaz, 2014:44-45). Diğer bir anlamda ise güvenilirlik, yapılan ölçme sonuçlarının duyarlık derecesini ifade eder (Turgut ve Baykul, 2015:123). Ölçme sonuçlarının belirli bir zaman aralığında değişmeden kalması, yani kararlılığıdır. Tekrarlı ölçümlerde aynı veya benzer sonuçların elde edilmesi durumudur. Ölçme sonuçlarının tutarlılığı, ölçme sonuçlarının hatasızlığı oranında gerçekleşecektir (Köse, 2014: 91). Ölçme sonuçları üzerinde hataya yol

açan her unsur, güvenilirliği de doğrudan etkilemektedir. Güvenirlik katsayısı üzerinde olumsuz etkiye sahip olan unsurlar da vardır. Bunlar; Puanlamanın nesnelliği, puanlayıcı güvenilirliği, test edilen grubun değişkenliği, testteki madde sayısı ve testin güçlük düzeyidir (Köse, 2014:103-105). Bir ölçme sonucu barındırdığı tesadüfi hataların azlığı oranında güvenilirirdir (Turgut ve Baykul, 2015: 123). Bir testte gerçek ve gözlenen puanlar arasındaki ilişkinin derecesini ifade eden korelasyon katsayısının karesi veya gerçek puanlar varyansının gözlenen puanlar varyansına oranı, güvenilirlik indeksi olarak bilinir (Baykul,2015:139). Ölçme aracında olması istenen özelliklerden birisi güvenilirlik kavramı iken diğer bir kavram ise geçerliktir.

2.4.Geçerlik

Bir ölçme aracının kullanım amacına hizmet etme derecesidir. Bir ölçme unsuru kullanılış amacına ne kadar uygun ise o düzeyde geçerlilik özelliğine sahiptir. (Yılmaz, 2014:49-50). Geçerlik, bir test açısından bakıldığında, testi belli bir kapsama bağlı ölçmek istenen nesneyi ne kadar iyi ölçtüğüne bağlı bir hüküm verme ve tahmin etme işlemidir (Cohen and Swerdlik 2020:108). Geçerlik, arzu edilen niteliği ölçme ve bu işi başka niteliklerin etkisini ölçülere yansıtmadan yapabilme durumunu gösterir (Özçelik, 2016:125). Bir ölçme unsurunun ölçmesi istenen niteliği ne kadar ölçebildiğinin derecesine o ölçme aracının geçerliği adı verilir. Başka bir söylemle ölçme unsurunun ölçtüğü değişkeni farklı, başka hiçbir parametrenin karışmadan ölçebilmesinin derecesidir. Geçerlik kavramı kendi içinde türlere ayrılmaktadır. Geçerlik türleri şu şekildedir;

Kapsam Geçerliği: Ölçme unsurunun ölçmesi arzu edilen kaplamı, eşit bir şekilde ne olası eksik olarak ölçmesi ne de olası fazla olarak ölçmemesi durumudur (Güler, 2017:58).

Yapı Geçerliği: Kişilerde olduğu varsayılan ama gözle görülemeyen nitelikler “yapı” olarak adlandırılır. Yapı geçerliği ölçme aracının kuramda var olduğu varsayılan, gözle görülemeyen bu yapıları ölçebilme derecesine denir (Güler, 2017:58).

Ölçüt Dayanaklı Geçerlik: Diğer bir deyişle kriter geçerliği; test puanının kişinin bir ya da birden fazla ölçüte endeksli başarımlar ihtimali gücünü gösterir. Bu sebeple kişinin testteki performansı (yordayıcı) bir ölçütle karşılaştırılır (Köse, 2014: 75). Ölçüt dayanaklı geçerlik ise kendi içinde ikiye ayrılır;

Uygunluk Geçerliği: Bir ölçme unsurundan alınan puanların o ölçme unsuruyla ya da ölçme unsurunun uygulanmasından bir zaman sonra ölçütten elde edilen puanlarla olan korelasyonudur (Köse, 2014:77).

Yordama Geçerliđi: Ölçme unsurunun kişinin gelecekteki başarımını kestirim gücü olarak ifade edilebilir (Köse, 2014:76).

Güvenirlik ve geçerlik ölçme ve değerlendirme uygulamalarında olması istenen özelliklerdir. Uygulanan bir testin güvenilir ve geçerli olması gerekmektedir. Güvenirlik ve geçerliđi etkileyen faktörler vardır bu unsurlardan birisi ise yanlılık kavramıdır. Yanlılık; testte yer alan bir maddenin bir gruba (cinsiyet, dil, etnik grup vb.) avantaj veya dezavantaj sağlamasıdır (Zumbo, 1999:12). Testten alınan puanların geçerliđi önemlidir. Bu yüzden bir sistematik hata sorunu olan yanlılık göz ardı edilmemelidir. Bu bağlamda yanlılık kavramına bakılması yerinde olacaktır.

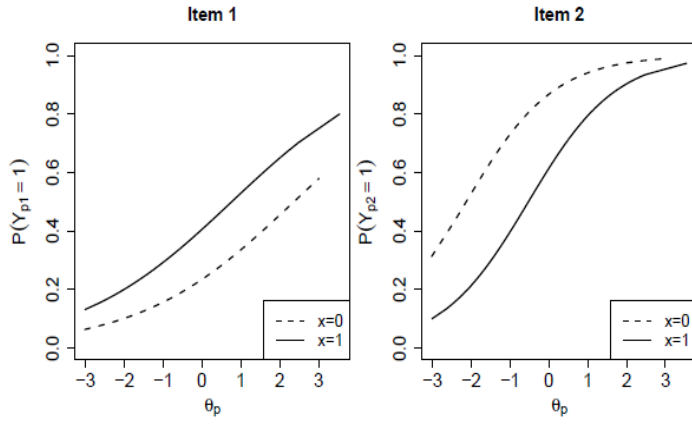
2.5.Yanlılık

Test maddelerindeki yanlılık, ölçümde sistematik bir hatanın varlıđı olarak da kabul edilir. Maddeler, belirli bir alt grup için, aynı popölasyondan oluşturulan başka bir alt grubun veya grupların performans(lar)ı ile karşılaştırıldığında nispeten daha az veya çok zor olarak değerlendirilir. Yanlılık bir puanın varlıđı değildir. Aynı yetenek düzeyinde olduğuna inandığımız iki grup arasındaki farktır. Zihinsel testleri ve diğer ölçümleri tanımlamak için kullanıldığında "yanlılık" terimi, belirli ve benzersiz bir anlama sahiptir. Yanlılık, ölçüm sürecindeki sistematik bir hata olarak tanımlanır (Osterlind,1983:10-11). Test içeriğindeki, test yönetimindeki ve/veya puanlama prosedürlerindeki sistematik hatalar, bazı test katılımcılarının gerçek yeteneklerinden daha düşük veya daha yüksek puanlar almasına neden olabilir. Yanlılığın kaynađı, testin ölçmeyi amaçladığı özellik ile alakasızdır (NCME, 2017). Hiçbir test karşıdaki bireysel yapıyı tam olarak ölçemez. Ölçüm hatası farklı grupların üyeleri için, puanları eşit olarak etkilediđi sürece yanlılıktan söz edilemez. Yanlılık etkisi, test performansında ve teste bađlı olarak kararlar alınırken orantısızlıđa neden olan grup farklılıklarında ortaya çıkar. Madde yanlılıđı analizlerinin amacı gerçek grup farklılıkları ile ölçümdeki farklılıđı ayırt etmektir. Testte yer alan maddeler arasındaki farklılıklar yanlılık kanıtı olarak kullanılamaz. Çünkü puan farklılıkları ve bilgi ve tecrübedeki grup farklılıklarının geçerli yansımaları olabilir. Testte yer alan maddeler bir grup için nispeten daha zorsa ve bu zorluğun kaynađı test ile alakasız ise bu maddenin yanlı olduğu söylenebilir (Camilli and Shepard, 1994:8). Bir testteki her test maddesi, her grupta tam olarak aynı madde yanıtlama fonksiyonuna sahip olsaydı, o zaman herhangi bir belirli yetenek veya beceri düzeyindeki kişiler, grup üyeliklerine bakılmaksızın, öđeyi dođru yanıtlama konusunda tam bir eşitliğe sahip olacaktı. Böyle bir test tamamen tarafsız olacaktır. Bir maddenin bir grup için diğerinden farklı bir madde yanıt işlevi varsa, maddenin yanlı olduğu açıktır. Belirli bir yetenek düzeyindeki

birinci gruptaki kişilerin, diğer gruptaki eşit yetenekteki kişilere göre maddeyi doğru yanıtlama şansı daha yüksektir. Madde yanlılığı, test ögesinin bazı özellikleri veya test amacı ile ilgili olmayan test durumu nedeniyle, bir grubun -sınava girenlerin- bir ögeyi doğru yanıtlama (veya bir ögeyi onaylama) olasılıkları başka bir grubun adaylarından daha düşük olduğunda ortaya çıkar (Zumbo, 1999:12). Yanlılık kavramının bir testte olması veya olmaması durumu “Değişen Madde Fonksiyonu” ile belirlenir.

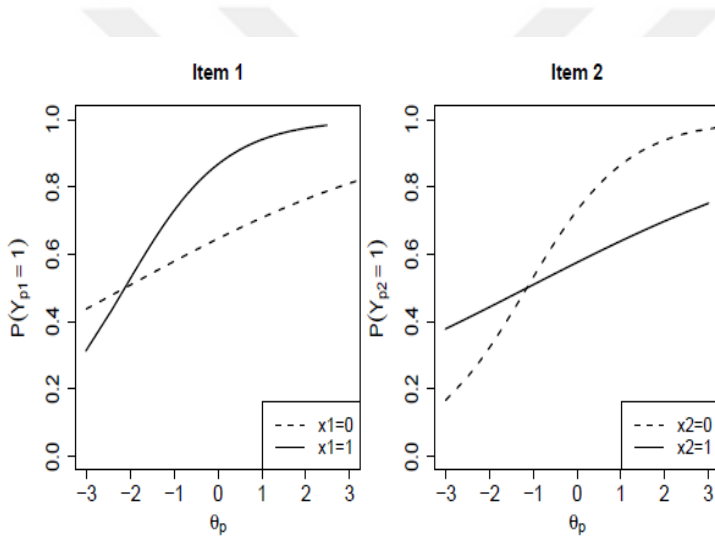
2.6.Değişen Madde Fonksiyonu (DMF)

Değişen madde fonksiyonu, gizli özellik veya daha genel olarak maddeler veya test tarafından ölçülen özellik ile eşleşen, çoğu zaman demografik olmak üzere gruplar arasındaki öğelerin işleyişindeki farklılıkları ifade eder (Camilli, 2006:220-256). Farklı gruplar için, farklı işlev gören tüm maddeleri tanımlamak için DMF ifadesi kullanılmaktadır. (Camilli and Shepard, 1994). Göreceli zorluk ve yanlılık arasındaki ayrımı korumak için ham veya yorumlanmamış göreceli zorluğu bilinen maddeye DMF içeren madde tanımı kullanılır (Holland and Thayer,1988:129-145). Mantıksal olarak farklı gruplar için farklı işlev gören tüm maddeler için DMF istatistikleri kullanılır. Bazı maddeler bir grup aleyhine nispeten daha zor olarak bulunur ve DMF unsurunun alt kümesi olarak kabul edilir ve muhtemelen testten çıkarılır (Camilli and Shepard, 1994:15-17). DMF, farklı gruplardan sınananlar, maddenin ölçmeyi amaçladığı temel yetenekle eşleştikten sonra madde üzerinde farklı başarı olasılıkları gösterildiğinde (veya onaylandığında) ortaya çıkar (Zumbo, 1999:12). DMF analizi, bir testteki maddelerin beklenmeyen davranışlarının bir göstergesini sağlar. Farklı gruplardan kişilerin belirli bir yanıt verme olasılıkları farklıysa, bir öge DMF göstermez; ancak ve ancak aynı temel gerçek yeteneğe sahip farklı gruplardan insanların belirli bir yanıt verme olasılıklarının farklı olması durumunda DMF gösterir. DMF unsurunu değerlendirmek için yaygın prosedürler klasik test teorisi tabanlı yöntemler olarak Mantel-Haenszel ve lojistik regresyondur (Zumbo, 2007: 223–233). Değişen madde fonksiyonu (DMF), bir maddenin iki grup için nasıl işlev gördüğüne ilişkin psikometrik bir farklılığı ifade eder. DMF, karşılaştırılabilir iki sınav grubu arasındaki, yani test tarafından ölçülen yapıya göre eşleşen gruplar arasındaki madde performansındaki bir farkı ifade eder. Eşleşen veya karşılaştırılabilir grupların karşılaştırılması önemlidir, çünkü madde işleyişindeki farklılıkları gruplar arasındaki farklılıklardan ayırt etmek önemlidir (Dorans and Holland, 1991:35-66). DMF biçim olarak tek biçimli (TB) ve tek biçimli olmayan (TBO) olarak iki grupta incelenebilir. Bir madde özelinde bir grup tüm yetenek düzeylerinde daha avantajlı durumdaysa TB DMF, maddenin doğru cevaplanma olasılığı farklı düzeylerde farklı grupları avantajlı duruma getiriyorsa TBO DMF olduğu söylenir (Zumbo, 1999: 12-14).



(Berger and Tutz, 2015:12)

Şekil 1. Tek Boyutlu DMF Örneği



(Berger and Tutz, 2015:21)

Şekil 2. Tek Boyutlu Olmayan DMF Örneği

DMF bulunurken Madde Tepki Kuramı ve Klasik Test Kuramı teknik ve yöntemleri kullanılır. Bu araştırmada ise klasik test kuramı ile yanlışlık kestirimi yöntemleri kullanılmıştır.

2.7. Klasik Test Kuramı

Klasik test kuramı, maddelerin zorluğu veya testi yapanların yeteneği gibi psikolojik testlerin sonuçlarını öngören ilgili psikometrik teorisinin bir bütünüdür. Bir kişinin bir testte gözlemlediği veya elde ettiği puanın gerçek bir puanın (hatasız puan) ve bir hata puanının toplamı olduğu fikrine dayanan bir test teorisidir (NCME, 2017). Klasik test kuramında, gözlenen puan içinde, bireyin elde ettiği gerçek puan ile ölçme esnasında yapılan hatadan kaynaklı olan bir miktar hata puanı bulunur. Bu durum, klasik test teorisinin temelini oluşturur.

İfade;

Gerçek Puan=Gözlenen Puan $\pm$ Hata Puanı

$X = T + E$ şeklindedir ve bu ifade klasik test teorisinin temelini oluşturur. Bu ifadede X =toplam puanı, T =gözlenen puanı, E =hata puanını temsil eder (Baykul, 2015:108).

Klasik test teorisi aşağıdaki fikirler kavramsallaştırıldıktan sonra oluşturulmuştur;

1. Ölçümlerde hataların varlığının tanınması,
2. Bu hatanın rastgele bir değişken olarak anlaşılması,
3. Korelasyon kavramı ve nasıl ifade edileceği (Traub, 1997:8-14).

KTK'nın bir özelliği de bir testin standart ölçüm hatasının tutarlı olmasıdır ve bu durum bütün bir grup için geçerlidir. Yani standart hata kişiden kişiye farklılık göstermez. Standart hata kişiden kişiye değil, bunun yerine teste giren çok sayıda kişi tarafından üretilir ve daha sonra potansiyel sınav katılımcıları popülasyonuna genelleştirilir. Ek olarak ham test puanına bakılmaksızın, standart hata her birey için aynıdır (Kline, 2005:91-94). Bütün kuramlarda olduğu gibi klasik test kuramının da varsayımları vardır. Klasik test kuramının varsayımları şunlardır (Baykul, 2000; Hambleton ve Swaminathan, 1985; Lord ve Novick, 1968):

Gözlenen puan gerçek puan ile hata puanının toplamıdır.

Hataların evrendeki dağılımının ortalaması sıfıra eşittir.

İki ayrı ölçmeden elde edilen hata puanları arasındaki korelasyon sıfıra eşittir.

Bir ölçme kümesine ait hata puanları ile aynı değişkeni ölçen başka bir ölçme kümesine ait gerçek puanlar arasındaki korelasyon sıfıra eşittir. Klasik test kuramının kendine ait yöntem ve teknikleri vardır. Araştırma yanlılık tespitine yönelik bir çalışma olduğundan dolayı KTK ile yanlılık kestirme yöntemleri kullanılacaktır. KTK'da yanlılık kestirim yöntemi olarak duruma göre "Mantel Haenszel" ve "Lojistik Regresyon" kullanılır.

2.7.1 Mantel Haenszel Yöntemi

Mantel Haenszel, adını William G. Cochran, Nathan Mantel ve William Haenszel'den almıştır (Cochran, 1954). İstatistikte, Cochran-Mantel-Haenszel testi (CMH), tabakalı veya eşleşmiş kategorik verilerin analizinde kullanılan bir testtir. Bir araştırmacının, kategorik durumu hesaplarken ikili bir yordayıcı veya doğru veya yanlış kontrol durumu gibi ikili bir sonuç arasındaki ilişkiyi test etmesine izin verir (Agresti, 2002: 231-232). Bir testin

kategorik bir cevaba veya birkaç gruba uzantıları genellikle Cochran–Mantel–Haenszel istatistikleri olarak adlandırılır (Mantel, 1963).

Tablo 1. Mantel Haenszel 2x2 Olasılık Tablosu

	Doğru	Yanlış	Toplam
Referans grubu	Doğru cevaplayan referans grubu öğrenci sayısı	Yanlış cevaplayan referans grubu öğrenci sayısı	Toplam referans grubu öğrenci sayısı
Odak grubu	Doğru cevaplayan odak grubu öğrenci sayısı	Yanlış cevaplayan odak grubu öğrenci sayısı	Toplam odak grubu öğrenci sayısı
Sütun Toplamı	Doğru cevaplayan toplam öğrenci sayısı	Yanlış cevaplayan toplam öğrenci sayısı	Toplam cevaplayan öğrenci sayısı

Tablo 1’de de görüldüğü üzere MH yönteminin hesaplanması bu şekilde olmaktadır. Öğrenciler odak ve referans grup olarak ikiye ayrılır. Doğru ve yanlış cevap veren öğrencilerin toplama olan oranına bakılarak bir değer yargısına varılır. Bunun yanında Zwick and Erickan (1989) Delta (Δ) DMF etki boyutunu değerlendirmek için aşağıdaki yorumlama yönergelerini önerdiler:

Delta Δ <1 ise A düzeyindedir yani yok denilebilir veya ihmal edilebilir düzeydedir.

1< Delta Δ < 1,5 ise B düzeyinde yani orta düzeyde DMF vardır.

1,5< Delta Δ ise C düzeyindedir yani yüksek düzeyde DMF vardır demektir (Zwick and Erickan, 1989). Örneklem durumu ikili bir sonuç değişkeni ve durum değişkeni (Doğru – Yanlış) gibi ikili bir belirleyiciyi göz önünde bulundurur. Gözlemler katmanlar halinde gruplandırılır. Gruplandırılan veriler, her değişken için bir tane olmak üzere bir dizi 2×2 olasılık tablosunda özetlenmiştir. Tablo ise aşağıda verilmiştir;

2.7.2. Lojistik Regresyon

Lojistik Regresyon (LR) yöntemi de DMF araştırmalarında yaygın olarak kullanılmaktadır (Wang, 2004:450-480). Yöntem ayrıca, tespit edildiğinde DMF' nin büyüklüğünü ölçen bir etki büyüklüğü ölçüsü sağlar. Bağımlı değişken olarak madde yanıtını (0 veya 1), gruplama değişkeni (1= referans, 2= odak olarak şeklinde kodlanmış), her konu için toplam ölçek puanı karakterize edilir ve etkileşimine göre bir grup olarak kullanılır.

İstatistikte lojistik model (veya logit model), belirli bir sınıf veya olayın geçme/başarısız olma, kazanma/kaybetme, canlı/ölü veya sağlıklı/hasta gibi var olma olasılığını modellemek için kullanılır. Testte yer alan her maddeye 0 ile 1 arasında bir olasılık atanır. Bir olumlu sıfır olumsuz olarak düşünülebilir. Lojistik regresyon, daha birçok karmaşık uzantı mevcut olmasına rağmen, temel formunda ikili bağımlı değişkeni modellemek için kullanılan istatistiksel bir modeldir. Regresyon analizinde lojistik regresyon bir lojistik modelin parametrelerini tahmin etmektir (Tolles and William, 2016: 533-534). Matematiksel olarak ifade edilecek olursa, bir ikili lojistik model, iki değer "0" ve "1" olarak etiketlendiği bir gösterge değişkeni ile temsil edilen geçti/kaldı veya doğru/yanlış gibi iki olasılığı olan değere sahip bağımlı bir değişkene sahiptir. İkili lojistik regresyon modelinde bağımlı değişkenin iki kategorik düzeyi vardır. İki'den fazla değere sahip çıktılar multinominal lojistik regresyon ile ve birden fazla kategori sıralanırsa sıralı lojistik regresyon ile modellenir (Walker and Duncan, 1967: 167-178). İkili lojistik regresyonda, sonuç genellikle "0" veya "1" olarak kodlanır, çünkü bu şekilde en basit yorumlamaya ulaşılabilir (Hosmer and Lemeshow 2000:7). İkili lojistik regresyon, bağımsız değişkenlerin değerlerine dayalı bir vaka olma ihtimalini tahmin etmek için kullanılır. 1-0 olarak kodlanma sebebi ise başarılı/başarısız veya doğru/yanlış gibi değerlendirme yapabilmek içindir. Lojistik regresyon hem tek biçimli hem de tek biçimli olmayan DMF'yi belirleme konusunda etkilidir. LR tekniği ile DMF belirlenirken, madde etkisi değişkeni anlamlı olduğunda maddede DMF olduğu, grup-madde etkileşim değişkeni anlamlı olduğunda tek biçimli olmayan değişen madde fonksiyonu olduğu sonucuna varılır (Gök, Kelecioğlu ve Doğan, 2010:7). Lojistik regresyon ile modele dahil edilen her bir terimi test ederek ilgilenilen öge üzerinde tek biçimli ve tek biçimli olmayan DMF'nin varlığını değerlendirmek için bir ki-kare testi kullanılır (Ukanda vd., 2019: 878-887). Lojistik regresyon, sürekli değişkeni işleyebildiği ve çok değişkenli analiz gerçekleştirebildiği için CMH testinden daha geneldir. CMH testi uygulanabildiğinde, koşullu lojistik regresyonun CMH testi istatistiği ve skor testi istatistiği aynıdır (Day and Byar, 1979:623-630).

2.8. İlgili Araştırmalar

2.8.1. Yurtdışında yapılan yanlılık çalışmaları

Kurnaz (2006), çalışmada, MH ve LR ile analiz yapılmış olup, cinsiyete göre MH yönteminde 16 LR yönteminde ise 14 maddede yanlılık bulunduğu rapor edilmiştir. Sosyoekonomik düzeye göre ise MH yönteminde 33, LR yönteminde ise 20 maddenin DMF gösterdiği anlaşılmıştır. İki yöntemin uyumlu olduğu sonuçlar vermediği rapor edilmiştir.

Cinsiyet değişkenine göre yanlılığa bakılan, sorularda MH yöntemine göre 95 sorudan 1 sorunun A düzeyinde 8 maddenin B düzeyinde 7 maddenin ise C düzeyinde DMF içerdiği sonucuna ulaşılmıştır. Referans grup olarak kız öğrenciler seçilmiştir. 8 madde kız öğrencilerden yana yanlı çıkmıştır. 8 madde ise erkek öğrencilerden yana yanlı çıkmıştır. LR sonuçlarına göre ise toplam 14 maddede yanlılık olduğu, 2 maddenin A düzeyinde olduğu, 2 maddenin B düzeyinde olduğu ve 10 maddenin ise C düzeyinde yanlı olduğu tespit edilmiştir.

Bakan Kalaycıoğlu (2008), çalışmasında 2005 yılı üniversite seçme sınavında yer alan maddelerin cinsiyete göre ve okul türü değişkenlerine göre DMF içerip içermediği incelenmiştir. Yapılan bu çalışmada cinsiyet değişkenine göre özellikle felsefe bölümünün maddelerinde kız öğrencilerin, fizik ile alakalı maddelerde ise erkek öğrencilerin başarılı olduğu belirtilmiştir. Aynı doğru sayısına ulaşmış öğrencilerin farklı maddeleri yanıtlayarak birbirine denk doğru cevap sayıları verdikleri gözlemlenmiştir. Tarih testine erkek öğrenciler özellikle savaş ve siyaset odaklı konularda daha çok doğru yanıt verirken, aynı doğru sayısına kız öğrenciler felsefe alt testine daha çok doğru cevap vererek ulaşmışlardır. Alt test olarak matematik testinde geometri maddelerinde erkek öğrencilerin daha başarılı olduğu görülürken, işlem yapılan maddelerde ise kızların daha başarılı olduğu sonucu ortaya çıkmıştır. Aynı şekilde fen bilimleri alt testine bakıldığında fizik konusu ile ilgili maddelerin erkekler tarafından daha çok doğru cevaplanma oranı var iken biyoloji maddelerinin kız öğrenciler tarafından daha çok cevaplanma oranı vardır. Okul türü değişkenine göre bakılacak olursa Anadolu lisesi öğrencilerinin bilgi gerektiren maddelerde daha başarılı olduğu sonucuna ulaşılmıştır. Genel lise öğrencilerinin ise daha kolay olan maddelerde başarılı olduğunu belirtmiştir. Bu aşamada ise öğrencilerden ziyade okul eğitim düzeyi ve imkanların farklı olmasının etkili olabileceği göz önüne alınmıştır.

Ateşok Deveci (2008), çalışmasında 2005 yılı mart ayında yapılan Üniversitelerarası Dil Sınavı'nın eğitim alanlarına ve cinsiyete göre DMF içerip içermediği incelenmiştir. Üniversitelerarası Dil Sınavı Sosyal Bilimler İngilizce testi maddelerinin, % 11'i Eğitim Bilimleri - Öğretmen Yetiştirme alanı ile Beşeri ve İdari Bilimler alanlarından gelen, yanıtlayıcı grupları için DMF'ye yol açmaktadır. DMF'ye yol açan madde konularının siyaset siyaset, ekonomi, bankacılık ve bilgisayar olanları Beşeri ve İdari Bilimler alanına; tarih, askeri operasyon ve arkeoloji konusu içerikli maddelerin ise Eğitim Bilimleri - Öğretmen Yetiştirme alanına avantaj sağladığını belirtmiştir. Üniversitelerarası Dil Sınavı'nın İngilizce testinde yer alan maddelerin %16'sı tüm kız öğrenci ve erkek öğrenci cevaplayıcı gruplarında DMF'na rastlanmıştır. Doğum, ekonomi, para ve sanat olan maddeler kız öğrenciler lehine; askerlik, döviz kuru, tarih ve ticaret odak konulu maddelerin ise erkek öğrenciler lehine avantaj sağladığı

belirtilmiştir. Testte kız öğrencilerin genel olarak erkeklere göre daha avantajlı olduğu rapor edilmiştir.

Sırgancı (2012), çalışmasında PISA sınavı kapsamında uygulanan öğrenci anketinin kültürel ve dilsel farklılıklara göre DMF içerip içermediğini incelemiştir. Doğrulayıcı faktör analizi sonuçlarına göre faktör yapısı bütün kültürler için aynı çıkmıştır. Bu amaçla referans ülke olarak Avustralya alınmıştır. Avustralya örnekleminde hareketle, aynı dil konuşulan ve benzer kültür olan Yeni Zelanda odak grup olarak belirlenip karşılaştırılmıştır. Aynı dil farklı kültür, odak grup olarak Amerika Birleşik Devletleri seçilerek karşılaştırma yapılmıştır. Farklı dil, farklı kültür ve odak grup olarak Türkiye örneklemini seçilmiştir. Yapılan analizler sonucunda kültür unsurları arasındaki uçurum arttıkça ve dil yapısındaki değişimlerin farklılığı arttıkça yanlılık durumunun arttığı belirtilmiştir.

Suna (2012), yaptığı çalışmada, ulusal sınavlardan olan ve 2007 yılında uygulanan TIMSS sınavının Fen bilimleri alt testinin maddelerinin dil ve cinsiyet unsuruna göre yanlılık durumuna bakılmıştır. Bu amaçla Türkiye örneklemini alınarak İngiltere örneklemini ile karşılaştırılmıştır. Karşılaştırma yapılırken LR, MH ve SIBTEST teknikleri kullanılarak aralarındaki uyuma da bakılmıştır. Türkiye lehine yedi maddenin dile dayalı olarak DMF gösterdiği, geriye kalan dört maddenin ise İngiliz öğrenciler lehine avantaj sağladığı belirtilmiştir. Bu çalışmada cinsiyet değişkenine göre ise sadece Türkiye örneklemini ele alınmıştır.

Karakaya ve Kutlu (2012), yapılan araştırmada seviye belirleme sınavının alt testi olan Türkçe testi incelenmiştir. Testte yanlı bulunan sorular ile ilgili uzman görüşüne başvurulmuştur. Uzmanlar maddenin konusunun akvaryum, balık gibi unsurlar olduğunu ve bu sebeple de bu maddenin erkekler lehine avantaj sağlayacağı yönünde fikir bildirmişlerdir. Diğer yandan Türkçe alt testinde yer alan okuduğunu anlama sorusunun ise kızlar lehine avantaj sağlayacağı yönünde fikir bildirmişlerdir.

Başusta (2013), yaptığı çalışmada 2006 yılında yapılan uluslararası sınav olan PISA sınavındaki fen bilimleri testlerini kültür dil açısından yanlılık unsuru olup olmadığına göre incelemiştir. Bu bağlamda Kanada kültürü aynı dili farklı olarak incelenmiştir. Kanada örnekleminde uygulamaya giren öğrencilerin kitapçıkları Fransızca form alanlar ve İngilizce form alanlar şeklinde karşılaştırılmıştır. Buna rağmen bu iki grup arasında DMF içeren maddeler vardır. Farklı kültür aynı dil karşılaştırma çalışmaları için Avustralya ve İngiltere örneklemi seçilmiştir. Farklı kültür, farklı dil olarak ise Türkiye – İngiltere örneklemini seçilerek karşılaştırma yapılmıştır.

Akalın (2014), yaptığı çalışmasında 2007 yılında yapılan kamu personeli seçme sınavının genel yetenek bölümünde bulunan maddelerin cinsiyet değişkenine göre ve mezun olunan bölümlere göre yanlılık içerip içermediğinin araştırılmasını amaçlamıştır. Yapılan analizler sonucunda sözcük bilgisi, okuduğunu anlama, anlatım özellikleri ile ilgili maddelerin kadınlar grubu ve sosyal bilimler grubu lehine olduğunu bunun da kitap okuma oranının kadınlarda daha fazla olmasından kaynaklı olduğunu belirtmiştir. Sayısal bölümde; Kalaycıoğlu (2008)'de belirtildiği gibi matematiksel işlemlerin kadınlar lehine, geometri, grafik ve tablo okuma ve yaşantı ile ilgili maddelerin ise erkekler lehine olduğu belirtilmiştir.

Gök, Kabasakal ve Kelecioğlu (2014) yapılan çalışmada 2009 yılında uygulanan PISA sınavının öğrenci anketi tutum maddelerinin kültüre göre incelenmesi ve DMF içerip içermediğinin araştırılmasıdır. Bu çalışmada da Suna (2012), Sırgancı (2012) ve (Başusta) 2013'nin çalışmalarıyla paralel sonuçlar çıkmıştır. Kültürel ve dilsel farklılıklar arttıkça DMF gösteren maddelerin de sayısı artmaktadır. Farklı dil ve farklı kültür unsurlarına sahip ülkelerde DMF gösteren madde oranı çok yüksek iken, aynı dil aynı aynı kültür düzeyine gelince oldukça düşük bir oran aldığı gözlemlenmiştir.

Kelecioğlu, Karabay, Karabay (2014) yapılan çalışmada 2009 yılında yapılan seviye belirleme sınavında yanlı çıkan maddelerin incelenmesi amaçlanmıştır. Yapılan çalışmada okul türlerine göre ve cinsiyet değişkenine göre DMF unsurlarına bakılmıştır. Altmış dört maddenin incelenmesi sonucunda üç madde kızlar lehine avantaj sağlarken; iki madde erkekler lehine avantaj sağlamıştır. Yapılan çalışmada özel okulların devlet okullarına göre yanlılık avantajının çok daha yüksek olduğu belirtilmiştir.

Eminoğlu Özmercan (2015), tarafından yapılan çalışmada, uluslararası düzeyde uygulanan PISA sınavlarının 2003 ve 2012 yılında yapılan uygulamalarının Türkiye – Kore örnekleminde değerlendirilmesi amacıyla yapılmıştır. Bu araştırma 2003 ve 2012 yıllarında yapılan PISA sınav uygulamasının sonuçlarını karşılaştırmaktadır. DMF gösterme olasılığı yüksek iki ülke olan Türkiye ve Kore örneklemleri karşılaştırılmıştır. Yanlı çıkan maddelerin nedenini uzmanlar; madde içeriğine aşina olmamak olarak belirtmişlerdir. Çeviri kalitesinin yetersi olmasının ve kültürel farklılıkların da olmasının etkisinin olacağını belirtmişlerdir.

Yıldırım (2015), tarafından yapılan çalışmanın amacı, 2012 yılında uygulanan seviye belirleme sınavında yer alan matematik alt testinin cinsiyet ve okul türüne göre yanlılık içerip içermediğini araştırmaktır. Okullar arasında ise DMF farklılıkları çıkmıştır. Ancak yüksek düzeyde değildir şeklinde belirtilmiştir.

Şenferah (2015), tarafından yapılan çalışmada 2010 yılında yapılan seviye belirleme sınavının matematik alt testinde yer alan maddelerdeki yanlılığını cinsiyet ve okul türüne göre

bulunması amaçlanmıştır. Kalaycıoğlu (2008) ve Akalın (2014)'ın bulgularına benzer şekilde geometri sorunun erkekler lehine yanlı olduğu rapor edilmiştir. LR yöntemi ile DMF bulunan bir madde ise TBO DMF içermektedir. Bu sebeple hem erkek öğrenci grubunun hem de kız öğrenci grubunun farklı çözüm yolları geliştirmiş olabileceği gözlemlenmiştir. Okul türü değişkenine göre ise iki madde DMF göstermiştir. Bu maddelerin DMF gösterme sebebi ise okulun imkan ve olanakları olarak gösterilmiştir.

Akcan (2018), tarafından yapılan çalışmada 2016 yılında yapılan lisans yerleştirme sınavının İngilizce testindeki maddelerin cinsiyet ve okul türü değişkenine göre yanlılığının incelenmesi amaçlanmıştır. Cinsiyet değişkenine göre DMF bulunmazken, bir adet madde İngilizce'den Türkçe'ye çeviri sorusunda erkek öğrenciler lehine madde yanlılığı bulunmuştur. Okul türleri meslek lisesi, Anadolu lisesi, özel lise ve imam hatip lisesi olarak sınıflandırma yapılmıştır. Yapılan araştırmada okul türüne göre DMF gösteren maddelerin cinsiyete göre DMF gösteren madde sayısından daha fazla olduğu belirlenmiştir. Okul türüne göre bu denli DMF gösteren maddelerin olması okul türleri arasındaki eğitim veya imkân farklarından olabileceği düşünülmüştür.

2.8.2. Yurtdışında yapılan yanlılık çalışmaları

Linn, Levin ve Hastings (1981) çalışmasında Okuduğunu anlama testindeki öğelerin belirli özelliklerinin, öğrencilerin belirli alt gruplarının okuma başarısına ilişkin yanlı tahminlere yol açabileceği olasılığını araştırmıştır. Başarı testinin okuduğunu anlama bölümündeki madde yanıt verileri Çapa Testi Çalışması veri dosyalarından elde edilmiştir. Öğrenci sınıf düzeyi (beşinci veya altıncı), okulun bulunduğu mahallenin gelir düzeyi (düşük veya orta/üstü) ve öğrencinin ırkı (siyah veya beyaz) olmak üzere üç faktörün birleşimi ile örtüşmeyen sekiz öğrenci alt grubu tanımlanmıştır. Üç parametrelili lojistik model kullanılarak sekiz alt grubun her biri için öğrenci yetenek ve madde parametrelerinin tahminleri ayrı ayrı elde edildi. Yanlılık indeksleri, alt grup çiftleri için madde karakteristik eğrilerindeki farklılıklara dayanarak hesaplandı. Bir maddenin önyargılı olarak etiketlenmesi için bir kriter, yalnızca gelir düzeyinde veya not düzeyinde farklılık gösteren aynı ırkın alt grupları için yanlılık indekslerinin dağılımı kullanılarak geliştirilmiştir. Bu kriter kullanılarak, siyah-beyaz öğrenci alt gruplarının dört bağımsız karşılaştırmasında üç madde tutarlı bir şekilde önyargılı olarak tanımlandı. Yanlı olarak tanımlanan öğelerin içerik ve biçim özelliklerinin, farklı yönlerde yanlı olmayan öğelerle veya öğeler arasında karşılaştırılmaları, sistematik içerik farklılıklarının belirlenmesine yol açmamıştır.

Reynolds, Taylor, Steffensen, Shirey, Anderson'un (1982) çalışmasında kültürel şemalar ve okuduğunu anlama arasındaki ilişki incelenmiştir. 8. sınıf düzeyindeki siyahi ve beyaz öğrencilere ses ve oynama ile ilgili maddeler okutulmuştur. Ses ile ilgili maddelerin siyahiler lehine, oyun oynama ile ilgili maddelerin beyaz öğrenciler lehine yanlı çıktığı anlaşılmıştır. Yanlı çıkma sebebi olarak ise dans ve müziğin siyahi öğrencilerin kültürel yapılarına hitap etmesinden, oyun oynama vb unsurların ise beyaz öğrencilere kültürel olarak hitap etmesinden kaynaklı olduğu düşünülmüştür.

Doolittle ve Cleary (1987), Amerika Birleşik Devletleri'nde üniversiteye girişte için kullanılan American College Testing sınavının matematik alt test maddelerinin cinsiyete göre DMF gösterip göstermediğini belirlemede MH yöntemi kullanmışlardır. Çalışmada algoritmik işlemlerle çözülebilen maddelerin kız öğrenciler lehine; geometri ve akıl yürütme becerisi gerektiren kelime problemleri maddelerinin ise erkek öğrenciler lehine işlediğini belirlemişlerdir.

Gafni (1990) çalışmasında İsrail'de üniversiteye kabul için uygulanan Psikometrik Giriş Testinin (PET) sözel (İbranice ve İngilizce) bölümlerindeki öğeler cinsiyetler arasındaki farklı öğe işleyişi (DMF) açısından incelenmiştir. Nisan 1984'te PET'in 3. Formunu alan 4.354 erkek ve 4.901 kadın, Nisan 1987'de PET'in 17. formunu alan 3.786 erkek ve 3.815 kadın için analizler yapıldı. Üç alt test incelenmiştir: (1) sözel akıl yürütme; (2) İngilizce ve (3) matematiksel akıl yürütme (sözel olmayan bir kontrol testi). Mantel-Haenszel ki kare testi 1987 grubu için DMF'yi inceledi. Sözel ve matematiksel akıl yürütme testlerindeki öğelerin yaklaşık üçte birinin DMF'ye sahip olduğu bulundu, ancak birkaç İngilizce alt test msddesinde bu durum görüldü. DMF sergileyen bazı öğelerin içeriği açıkça kadınsı ve erkeksi ilgi alanlarına ilişkin basmakalıp algılarla ilişkiliydi. Test içeriği üzerindeki etkileri tartışılmıştır.

M. Beller, N. Gafni (1991) The (1991) çalışmasında 1991 Uluslararası Eğitim İlerleme Değerlendirmesi sonuçlarını yeniden analiz ederek ve yorumlayarak 9 ve 13 yaşındakilerin matematik ve fen bilimleri sınavlarındaki performanslarındaki cinsiyet farklılıklarına küresel bir bakış açısı sunmuşlardır. Analizler 13 yaşındakileri test eden 20 ülkede ve 9 yaşındakileri test eden 14 ülkede yapıldı. Her yaş düzeyinde her popülasyondan rastgele 3.300 öğrenci örneği seçilmiş; yarısı matematikte, yarısı fen bilimlerinde değerlendirilmiştir. Hem alt alan düzeyinde hem de toplam puanlardaki matematik değerlendirmesindeki cinsiyet etki boyutlarının, özellikle 9 yaşındakiler arasında küçük olduğu bulunmuştur. Genel olarak, fen bilimleri için cinsiyet etkileri matematik için olanlardan önemli ölçüde daha büyüktü. Analizler seçilen yedi ülkede (Macaristan, İrlanda, İsrail, Kore, İskoçya, İspanya ve Amerika Birleşik Devletleri) gerçekleştirilmiştir.

Hixson and McGlichey (2004) çalışmasında sözlü okuma akıcılık puanlarının bir devlet okuma testinde ve Metropolitan Başarı Testlerinin, MAT 7 bölümünün kişilerin üzerindeki performansı ekonomik ve ırksal gruplar arasında farklı şekilde tahmin edip etmediğini inceleyerek yanlılık sorununu ele almıştır. Bir orta batı devlet okulundan 442 dördüncü sınıf öğrencisinden elde edilen bilgiler çoklu doğrusal regresyon kullanılarak analiz edildi. Her üç değişken de durum değerlendirmesi ve MAT / 7'deki performansın tahmin edilmesine önemli katkı sağlamıştır. Sözlü okuma akıcılığı en büyük katkıyı sağladı. Sonuçlar ayrıca sözlü okuma akıcılık puanlarının durum okuma testinde kesişme yanlılığı gösterdiğini göstermektedir. Aşamalı bir regresyon yöntemi yanlılık kanıtı bulamadı. Hem devlet okuma testinde hem de MAT / 7'de, Afrikalı Amerikalı ve düşük gelirli öğrencilerin test puanları aşırı tahmin edilirken, Kafkas ve yüksek gelirli öğrencilerin puanları düşük tahmin edildi.

Salehi and Tayebi (2012) Bu çalışmada, üç aşamalı lojistik regresyon prosedürü kullanılarak yapılan yüksek oranlı testin okuduğunu anlama alt testinde (35 madde) cinsiyet açısından DMF araştırılmıştır. Araştırmaya katılanlar, Tahran Üniversitesi'nde doktora programına girmek için kısmi bir gereklilik olarak söz konusu testi (UTEPT) alan hem erkek hem de kadın 3.398 test katılımcısı olarak seçilmiştir. Okuduğunu anlama bölümünün 35 maddesinin DMF gösterip göstermediğini tespit etmek için üç aşamalı bir prosedür kullanılarak lojistik regresyon kullanılmıştır. Utept'in okuduğunu anlama alt testinin ne erkek ne de kadınlardan yana yanlı olmadığı sonucuna varılmıştır.

Ahmadi and Jalili (2014), çalışmasında bir EFL okuduğunu anlama testinde DIF kaynaklarını araştırmayı amaçlamıştır. Buna göre, 203 (110 kadın ve 93 erkek) İranlı EFL sınav katılımcısının bir okuduğunu anlama testindeki performansının ortaya çıkan DMF'sini işaretlemek için iki DMF tespit yöntemi, lojistik regresyon (LR) ve madde tepki kuramı kullanılmıştır. Bu bağlamda yedi varsayımsal DMF kaynağı incelenmiştir: Metin aşinalığı, cinsiyet, konu/metin ilgisi, tahmin ve statü, gelir ve eğitim durumunun sosyal değişkenleri. Sadece LR, cinsiyet için ve metin aşinalığı, cinsiyete dayalı metin etkisini destekleyen cinsiyeti önlerken, metin aşinalığı, düşük düzeyde metin aşinalığı olan katılımcılara ters olarak yanlılık gösterebilir. Konuya ilgili olanlar için LR, daha yüksek ilgi düzeyine sahip grubu destekleyen tek bir madde buldu ve IRT modeli de DMF tespit etti. Tahmin ve gelir ile ilgili olarak, LR, DMF'nin düşük tahmin edicileri ve yüksek gelir grubunu desteklediğini gösterirken, IRT, aksine, DMF'nin yüksek tahmin edicileri ve düşük gelir grubunu desteklediğini gösterdi. Statü ve eğitim için her iki yöntem de buna bağlı olarak statüsü iyi olanlar ve eğitilmiş gruplar için DMF göstermiştir. Son olarak, diferansiyel testin

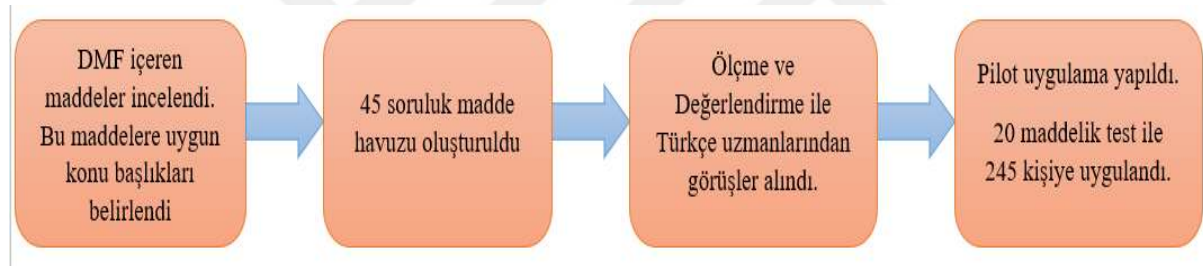
işleyiş sonucu sadece üç DMF kaynağının (cinsiyet, gelir ve ilgililer) test seviyesine aktarıldığını açıkça ortaya koymuştur. Bulgular, DMF kaynaklarının orantılı bir etkisini destekleyebilir.

ÜÇÜNCÜ BÖLÜM

YÖNTEM

3.1.Araştırmanın Modeli

Bu araştırmada 8.sınıf düzeyindeki öğrencilere yönelik geliştirilen Türkçe okuduğunu anlama testi ile, madde yanlışlığının test uygulanmadan önce uzman görüşü alınarak ve yapılan çalışmalar ile tespit edilip edilemeyeceğinin belirlenmesi beklenmiştir. Bu yönüyle araştırma betimsel araştırma türünde bir çalışmadır. Bu araştırma iki aşama halinde tasarlanmıştır. Bu aşamalar aşağıdaki şekillerde gösterilmiştir;



Şekil 3. Araştırmanın birinci aşaması adımları



Şekil 4. Araştırmanın ikinci aşaması adımları

3.2.Çalışma Grubu

Araştırmanın çalışma grubu olarak uzman grubu ve öğrenci grubu ile çalışılmıştır. Çalışma kapsamında birinci aşamada, öncelikle beş kişilik bir ölçme ve değerlendirme uzman grubu, iki kişilik Türkçe uzman grubu ve iki kişilik Türkçe öğretmen grubu ile çalışılmıştır. Daha sonra 245 kişilik öğrenci grubu ile çalışılmıştır. Araştırmanın ikinci aşamasında 881 kişilik öğrenci grubu ile yapılan çalışmadan sonra araştırmanın bu aşamasında madde yanlılığı alanında çalışan dört kişilik bir uzman grubu ile çalışılmıştır. Araştırmanın evreni olarak Türkiye’deki 8.sınıf düzeyindeki öğrenciler düşünülebilir. Örneklem olarak ise Güneydoğu Anadolu Bölgesi’nde yer alan bir ilin ilçelerindeki seçilen okullar olduğu ifade edilebilir. Örneklem seçimi olarak ise seçkisiz olmayan tabakalı amaçsal örnekleme yöntemi kullanılmıştır. Bu yöntem, ilgilenilen belli alt grupların özelliklerini göstermek, betimlemek ve bunlar arasında karşılaştırmalara olanak tanımak amacıyla tercih edilir (Büyüköztürk vd., 2017:94). Çalışmaya katılan öğrencilerin cinsiyet ve okul ismine göre katılımı demografik olarak şu şekildedir.

Tablo 2. Örneklem Tablosu

Okul Adı	Kız Öğrenci Sayısı	Erkek Öğrenci Sayısı	Toplam
Vali Lütfullah Bilgin Ortaokulu	152	154	306
Nurel Enver Taner Ortaokulu	87	99	186
Kocatepe Ortaokulu	91	108	199
Turgut Özal Ortaokulu	101	120	221
Batıkent Ortaokulu	38	42	80
Şehit Samet Kaymakçı İmam Hatip Ortaokulu	37	47	84
Feridun Oral Ortaokulu	14	21	35
Kazıklı Ortaokulu	10	5	15
			Toplam:1126

Tablo 2’de görüldüğü üzere toplamda 1126 öğrenciye ulaşılmıştır. Bu öğrencilerin 245 kişilik çalışma grubuna pilot uygulama aşamasında 881 kişilik çalışma grubuna ise ana uygulamada ulaşılmıştır.

3.3. Veri Toplama Araçları

Veri toplama aracı olarak 8.sınıf düzeyine uygun Türkçe okuduğunu anlama testi geliştirilmiştir. Test geliştirme süreci birbirini izleyen adımlardan oluşur ve bu sıralama madde yazarlarına veya test geliştiricilerine göre farklı şekillerde oluşturulabilir (Baykul, 2015:269). Bu adımlar sırasıyla; Testin hangi amaçlar kullanılacağına belirlenmesi: literatür taramasından hareketle madde yanlışlığı olasılığı olan içerikteki yazılan maddelerdir. Testle ölçülecek niteliklerin saptanması; geliştirilen Türkçe testi okuduğunu anlama amacıyla geliştirilmiştir. Maddelerin yazılması; maddeler yazım olarak MEB (2019) Türkçe dersi öğretim programı (İlkokul ve Ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. Sınıflar) öğretim programı içeriğindeki okuma kazanımlarına uygun olarak yazılmıştır. Her madde program içerisinde yer alan okuma kazanımlarına göre yazılmıştır. Yazılan maddeler alan uzmanlarından ve Türkçe alan öğretmenlerinden görüş alınarak yazılmıştır. Uygulanan testin madde olarak kazanımları aşağıda yer alan Tablo 3'te gösterilmiştir;

Tablo 3. Madde Kazanımları Tablosu

Madde No	Kazanım
1	Metnin ana fikrini/ana duygusunu belirler.
2	Görsellerle ilgili soruları cevaplar.
3	Metindeki yardımcı fikirleri belirler.
6	Bağlamdan yararlanarak bilmediği kelime ve kelime gruplarının anlamını tahmin eder.
7	Grafik, tablo ve çizelgeyle sunulan bilgileri yorumlar.
10	Deyim, atasözü ve özdeyişlerin metne katkısını belirler.
11	Metnin ana fikrini/ana duygusunu belirler.
12	Grafik, tablo ve çizelgeyle sunulan bilgileri yorumlar.
13	Metnin ana fikrini/ana duygusunu belirler.
14	Grafik, tablo ve çizelgeyle sunulan bilgileri yorumlar.
15	Metindeki yardımcı fikirleri belirler.
17	Metnin içeriğini yorumlar.
18	Görsellerle ilgili soruları cevaplar.
19	Metinle ilgili soruları cevaplar.
20	Metindeki yardımcı fikirleri belirler.

Maddelerin redaksiyonu; testi geliştirmek amacıyla öncelikle 45 maddelik bir havuz oluşturulmuştur. Oluşturulan maddeler beş kişilik ölçme ve değerlendirme uzmanlarına, iki

kişilik Türkçe alan uzmanlarına ve iki kişilik Türkçe alan öğretmenlerine gönderilerek uzman görüşü alınmıştır. Deneme formunun hazırlanması: alınan uzman görüşleri ve değerlendirmeleri sonrasında deneme formu olarak 20 maddelik bir Türkçe okuduğunu anlama testi geliştirilmiştir. Deneme uygulamasının yapılması; toplanan veriler analiz edilir ve testten çıkarılması gereken maddeler testten çıkarılır. Nihai testin oluşturulması; yapılan analizlerden sonra test son haline ulaşır ve ana uygulama için nihai test olarak kullanılır (Baykul, 2015:269).

3.4. Veri Toplama Süreci

Oluşturulan nihai test Güneydoğu Anadolu Bölgesi'nde yer alan bir ilin ilçelerindeki seçilen okullarda uygulanmıştır. Bu aşamada çalışma grubu olarak öğrenci grubu ve uzman grubu ile çalışılmıştır. Uygulama ile ilgili okul yönetimlerinden izinler alınmıştır. Okul öğretmenlerine uygulama ile ilgili bilgi verilmiştir. Uygulamalar 2020-2021 bahar döneminde yüz yüze yapılmıştır. Okul yönetimleri tarafından belirlenen gün ve saatlerde okula gidilerek çalışmalar tamamlanmıştır. Uygulama toplamda iki aşamadan oluşmuştur. Öncelikle 245 öğrenci ile deneme uygulaması yapılmıştır. Uygulama sonrasında toplanan verilerle gerekli madde analizleri yapılmıştır. Sonrasında ise 15 soruluk nihai test oluşturularak, 881 kişilik öğrenci sayısı ile ana uygulama yapılmıştır. Veriler toplandıktan sonra ise nihai testte yer alan hangi maddelerin, hangi düzeyde yanlı olabileceği ile alakalı uzman görüş formu uzman grubuna mail yoluyla iletilmiştir.

3.5. Verilerin Analizi

Verilerin analizi noktasında bütün veriler toplandıktan sonra öncelikle R paket programı ile betimsel istatistik analizleri yapılmıştır. Çarpıklık ve basıklık katsayılarına bakılarak dağılımın normal olduğu gözlemlenmiştir. Uç ve kayıp değerlere bakılmıştır. Normallik ve güvenilirlik analizleri yapılmıştır. Daha sonra AFA yapılmıştır. Büyüköztürk vd., (2016)'ne göre faktör yükü .30 altında olan maddeler testten çıkarılmıştır. Nihai teste ulaşıncaya madde istatistikleri olan madde güçlük indeksi ve veriler 1-0 olarak kodlandığından dolayı ayırt edicilik olarak basit yöntem olarak adlandırılan alt – üst grup yöntemi kullanılmıştır. Son olarak ise “DifR” paketi ile KTK belirleme yöntemlerinden olan MH ve LR analizleri yapılmıştır. Yapılan analizlerde referans grubu= kızlar, odak grubu ise= erkekler olarak analiz yapılmıştır. R paket programının özellikleri ile MH ve LR sonuçları karşılaştırması yapılmıştır. 0.05 düzeyinde ise 1. tip hata kabul edilmiştir. I. Tip hata, maddenin gerçekte DMF içermediği halde, madde DMF içeriyor şeklinde kararın verilmesi durumuna denir. Maddelerin DMF'li olup olmama durumları belirlenirken istenilen, güçlük oranlarının yüksek, I. tip hata oranlarının ise düşük olmasıdır (Sünbül ve Sünbül, 2016).

DÖRDÜNCÜ BÖLÜM

BULGULAR

Araştırmaya ait bulgular ve bulgulara ilişkin yorumlara sırasıyla aşağıda yer verilmiştir. Bu aşamada öncelikle 245 kişilik deneme uygulamasının betimsel istatistik analizlerine ilişkin verilere yer verilmiştir.

4.1.Betimsel İstatistik Analiz Sonuçları

Tablo 4. Betimsel İstatistik Analiz Sonuçları

	$\bar{X}$	Medyan	Mod	Varyans	Standart Sapma	Çarpıklık	Basıklık
Toplam	11.9	12.0	16.0	14.9	3.86	-0.55	-0.34

Tablo 7 incelendiğinde deneme uygulama yapılan 245 öğrencinin Türkçe okuduğunu anlama testinin aritmetik ortalaması 11.9, medyan değeri 12.0, mod değeri 16.0, varyansı 14.9, standart sapması 3.86, basıklık değeri -0.55, çarpıklık değeri -0.34 olarak tespit edilmiştir. Çarpıklık ve basıklık katsayısının sıfır olduğu bir dağılımın standart normal dağılıma sahip olduğu söylenebilir. Çarpıklık ve basıklık katsayısının ± 1.0 uç değerler arasında kalması, puanların normalden fazla bir sapma olmadığı şeklinde açıklanabilir (Büyüköztürk vd., 2016:16).

4.2.Açımlayıcı Faktör Analizi Sonuçları

Türkçe okuduğunu anlama testi için Açımlayıcı Faktör Analizi uygulanmıştır. Çok değişkenli istatistiklerde, Açımlayıcı faktör analizi (AFA), nispeten büyük bir değişken kümesinin altında yatan yapıyı ortaya çıkarmak için kullanılan istatistiksel bir yöntemdir. AFA, temel amacı ölçülen değişkenler arasındaki temel ilişkileri tanımlamak olan faktör analizinde kullanılan bir tekniktir (Norris and Megan, 2009:8-20). AFA, ölçülen değişkenin herhangi bir faktörle ilişkili olabileceğini varsayar. Bir ölçek geliştirirken, araştırmacılar doğrulayıcı faktör analizine (DFA) geçmeden önce önce AFA'yı kullanmalıdır (Worthington and Whittaker, 2006:806-838). AFA, bir dizi ölçülen değişken için altta yatan faktörleri/yapıları belirlemek için esastır (Suhr,2006:1-17). Toplanan veriler 1-0 (doğru-yanlış) olarak kodlandığı için tetrakorik korelasyon katsayısı kullanılmıştır. Tetrakorik korelasyon, ikili değişkenler bakımından değerlendiriciler arasındaki uyumu belirlemede kullanılan katsayılardan birisidir. Gözlenen her iki değişkenin de ikili (dichotom veya binary) değişken olduğu durumda

uygulanan polikorik korelasyonun özel bir halidir (Keskin, 2018:11-12). Bu konudaki en belirgin husus, korelasyonların hesaplandığı örneklem büyüklüğüdür. Gözlem sayısı arttıkça, elde edilen korelasyonların güvenilirliği artmaktadır. 50 kişilik bir örneklem korelasyon katsayılarının güvenilirliğini düşürürken, 1000 kişilik bir örneklem grubu faktör analizi yapabilmek için fazlasıyla yeterlidir. Örneklem büyüklüğü (Comrey and Lee, 1992) tarafından şöyle açıklanmıştır 50 kişilik bir örneklem çok zayıf; 100 kişilik bir örneklem zayıf; 200 kişilik bir örneklem orta; 300 kişilik bir örneklem iyi; 500 kişilik bir örneklem çok iyi; ve 1000 veya daha fazla kişilik bir örneklem mükemmel olarak adlandırılabilir İlişki katsayıları, küçük örneklemelerden kestirildiyse daha az güvenilir olma eğilimindedir. Bu yüzden, örneklem büyüklüğünün, ilişkilerin güvenilir bir şekilde kestirilebilmesini sağlayacak büyüklükte olması önemlidir. Literatürde, özellikle faktörler güçlü ve belirgin olduğunda ve değişken sayısı fazla büyük olmadığında, 100 ile 200 arasındaki örneklem büyüklüğünün yeterli olduğu belirtilmektedir (Büyüköztürk, 2002:480). Bu aşamada açımlayıcı faktör analizi sonuçlarını sırasıyla vermek yerinde olacaktır. İlk olarak faktör yükleri tablosu verilecektir.

Tablo 5. Faktör Yükleri Tablosu

Faktör Yükleri	1
M1	0,39
M2	0,51
M3	0,51
M6	0,69
M7	0,68
M10	0,68
M11	0,36
M12	0,58
M13	0,58
M14	0,76
M15	0,46
M17	0,58

Faktör Yükleri	1
M18	0,47
M19	0,56
M20	0,76

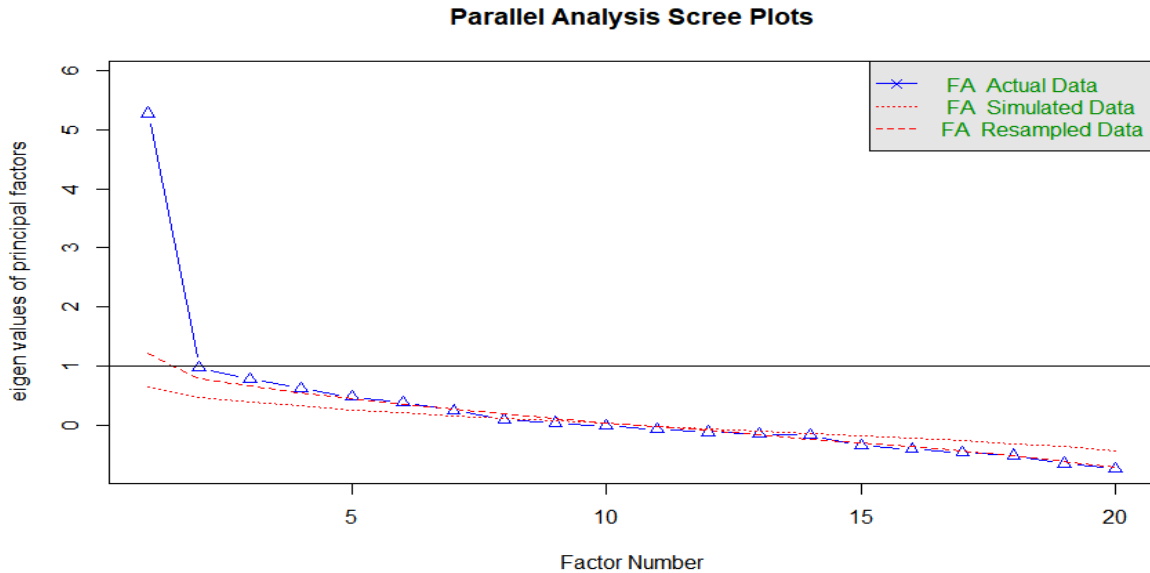
R paket programı, oluşturulan kodlar ile yapılan analizler sonucu modelin tek faktörlü olduğu görülmektedir. Açıklanan varyans ise %34 olarak rapor edilmiştir. Bir maddenin faktör yükü değerinin düşük olması demek o faktörle güçlü bir ilişkisinin olmadığını gösterir. Literatürde bir maddenin faktör yük değerinin en az 0.30 olması önerilir ve bunu savunan araştırmacılar çoğunluktadır (Büyüköztürk vd., 2016:194). Bu ifadeden hareketle faktör yük değeri 0.30 altında olan M4, M5, M8, M9 ve M16 testten çıkarılmıştır. Bu madde sayısı oranı ise toplam testin %25 dilimlik oranına denk gelmektedir. Faktör yükleri tablosundan sonra ise özdeğerler analiz tablosuna bakmak yerinde olacaktır.

Tablo 6. Özdeğer Analiz Sonuçları

Özdeğer Analiz Sonuçları	
Faktör V1	Özdeğer
1	5.886
2	1.757
3	1.610
6	1.157
7	0.983
10	0.751
11	0.637
12	0.604
13	0.480
14	0.460
15	0.388
17	0.225
18	0.162
19	0.119
20	0.000

1954'te Guttman, dikkate değer faktörlerin öz değerlerinin 1.0'dan büyük olması gerektiğini ifade etti. Bu, çoğu istatistiksel paketteki faktör sayısını belirlemek için varsayılan karar verme stratejisidir. Faktörler, tanım gereği, ölçülen değişkenlerin toplamı olarak oluşturulan gizli yapılardır ve bu nedenle, tek bir ölçülen değişkenden daha fazlasını içermelidir. Bir faktör tek bir ölçülen değişkenden oluşuyorsa, o ölçülen değişkenin örüntü/yapı katsayısı 1.0 (veya -1.0) olsa ve bu faktördeki diğer tüm değişkenlerin örüntü/yapı katsayıları .0 olsa bile, faktörün bir özdeğeri olacaktır. Bu nedenle, dikkate değer faktörlerin (yani ölçülen değişkenlerin toplamlarını temsil eden yapılar) öz değerlerinin 1.0'dan büyük olması mantıklı görünmesine vesile olmaktadır (Thompson, 2004:32). Ama Büyüköztürk vd., (2016)'ne göre ise özdeğerlerin azalan eğilimlerini gösteren yamaç – birikinti grafiği (Scree Plot) çizilerek eğimin kaybolduğu ya da çok küçük noktaya kadar olan özdeğerler kabul edilerek modelin faktör sayısı ifade edilebilir. 1 değerinin üzerinde dört faktör var gibi görünse de düşüş çok net görülmektedir. Bununla ilgili olarak şekil 5'te yer alan yamaç – birikinti grafiğine bakmak yerinde olacaktır. Bu sebeple yukarıdaki tablo 6'da yer alan özdeğerler durumuna bakılacak olursa modelin özdeğerlerindeki yüksek düşüşün birinci faktörden sonra olduğu görülmektedir, tek faktörün olduğu ve modelin tek faktörlü olduğu görülmektedir. Son olarak ise yamaç – birikinti grafiğine bakmak yerinde olacaktır.

Yamaç – Birikinti Grafiği (Scree Plot)



Şekil 5. Yamaç – Birikinti Grafiği (Scree Plot)

Özdeğerler, belirli bir faktör içinde temsil edilen bilgi miktarını karakterize eder. Faktörler art arda daha küçük özdeğerlere sahiptir. Bir eğim grafiği, yatay eksen oluşturarak özdeğer sayılarıyla birlikte, dikey erişimdeki özdeğer büyüklüklerini grafikler. Faktör çıkarımı, bir "eğim, düşüşün" olduğu noktada durdurulmalıdır. Özdeğerlerin düzleştiği grafiğin "dirseği" bulunur ve bu noktanın solundaki faktörler veya bileşenler anlamlı olarak tutulmalıdır (Thompson,2004:32-33). Bu tanıma göre özdeğerlere bakıldığında ve scree plot grafiğine bakıldığında düşüşün birinci faktörden sonra dik olduğu veya diğer bir tanımla dirseğin olduğu yerin birinci faktörden sonra olduğu görülmektedir. Bu ifadeden hareketle modelin tek faktörlü olduğu görülmektedir.

Örneklem büyüklüğü açısından faktör analizi için veri yapısının uygunluğunu test etmeye yönelik bir başka ölçüt ise Kaiser-Meyer-Olkin testi sonuçlarıdır. Bartlett'in Küresellik Testi sonuçlarına bakıldığında KMO değerinin 0.7, $\chi^2=7089$ serbestlik derecesi=190 olarak rapor edilmiştir. Analiz sonuçlarına göre Kaiser- Barlett Küresellik testinin de anlamlı ($p<.001$) olduğu görülmektedir. Kaiser-Meyer-Olkin değerinin yüksek olması, ölçekteki her bir değişkenin, diğer değişkenler tarafından mükemmel bir şekilde tahmin edilebileceği anlamına gelir. Kaiser-Meyer-Olkin testi sonucunda değerin 0.50'den düşük olması durumunda faktör analizine devam edilemez yorumu yapılabilir (Büyüköztürk vd., 2016:207). Kaiser-Meyer-Olkin testinde örneklem büyüklüğü için değerler;

0.50–0.60 arası ise kötü

0.60–0.70 arası ise zayıf

0.70–0.80 arası ise orta

0.80–0.90 arası ise iyi (Leech, Barret and Morgan, 2015;118-119)

0.90 ve üzeri ise mükemmel denilebilir (0.50 ve altında ise faktör analizi için kabul edilemez (Dziuban and Shirkey 1974:358-361).

Açımlayıcı faktör analizinden sonra ise güvenirlik analizi yapılmıştır. Bu testin güvenirlik katsayısı yöntemi olarak KR-20 kullanılmıştır. KR-20 madde güçlükleri hesaplanabiliyorsa kullanılır. Hesaplanamadığı veya bilinmediği durumlarda KR-21 kullanılır (Güler, 2017:52). Yapılan çalışmada madde güçlükleri hesaplanabildiğinden dolayı KR-20

kullanılmıştır. KR-20’de kabul edilen sınır .70 olarak belirlenmiştir (Pallant, 2020:113). Bu tanımlardan hareketle araştırmada KR-20 yöntemi kullanılmıştır. KR-20 ölçüm sonucu .79 olarak rapor edilmiştir. Kabul edilebilir alt değerin de .70 ve üzeri olabileceği belirtilmiştir. Sonuç olarak uygulanan test KR-20 güvenirlik katsayısı iyi çıkmıştır ve kabul edilebilir alt değerin üzerinde çıkarak güvenirlik düzeyinin yeterli düzeyde olduğu gözlemlenmiştir. Bu aşamada çoklu doğrusallık değeri olan VIF sonuçlarının verilmesi yerinde olacaktır.

Tablo 7. Çoklu Doğrusallık Varsayımı Analiz Sonuçları

Madde No	VIF	Tolerans Değeri
M1	1.09	0.920
M2	1.08	0.923
M3	1.23	0.810
M6	1.27	0.785
M7	1.27	0.789
M10	1.23	0.814
M11	1.06	0.944
M12	1.17	0.856
M13	1.13	0.881
M14	1.41	0.708
M15	1.11	0.899
M17	1.25	0.800
M18	1.12	0.892
M19	1.21	0.830
M20	1.41	0.711

VIF için en küçük değer 1.0 olmalıdır ve bu tam doğrusallık olmadığını gösterir. Tipik olarak pratikte tahmin ediciler arasında az miktarda bir doğrusallık vardır. Genel bir kural olarak, 5 veya 10’u aşan bir VIF değeri, sorunlu bir doğrusallık miktarını gösterir (James, Witten, Hastie and Tibshirani, 2014:102). Tanımdan hareketle VIF değerinin $1 < VIF < 5$ olması gerekmektedir. Bu tanıma göre yukarıdaki tablo 13’te verilen analiz sonuçları için çoklu doğrusallık varsayımının karşılandığı görülmektedir. Çok düşük olan tolerans değerleri (.1’den daha düşük), değişkenin modeldeki diğer değişkenler ile yüksek derecede korelasyonlara sahip

olduğunu gösterir (Pallant, 2015:184-185). Çalışmada 1 ve 0 gibi kategorik ölçümler kullanılmışsa, doğrusallık varsayımının ihlal edilmesi nedeniyle sonuçlar yanıltıcı olabilir (Tabachnick and Fidell, 2012:352). Bu verilen tanımlara göre ise yukarıdaki tablo 18’de yer alan tolerans değerlerinin anlamlı olduğu ve değişkenler arasındaki korelasyon değerlerinin yüksek olduğu görülmektedir. LR DMF için etki boyutunun incelenmesi MH DMF yöntemi gibi benzer bir DMF yönteminden elde edilen sonuçlarla karşılaştırılması gerekir.

4.3.Madde istatistikleri

Bu bölümde madde istatistikleri olarak madde güçlük indeksi değerleri ve madde ayırt edicilik indeksleri verilecektir.

Tablo 8. Madde İstatistik Analiz Sonuçları

Madde No	Madde Güçlük İndeksi	Madde Ayırt Edicilik İndeksi (Pbis)
1	.57	.51
2	.88	.30
3	.59	.68
6	.61	.68
7	.68	.62
10	.62	.64
11	.46	.38
12	.84	.35
13	.45	.50
14	.57	.74
15	.33	.48
17	.58	.71
18	.77	.46
19	.49	.64
20	.69	.77

Tablo 9’a göre madde güçlük indeksleri ve alt – üst grup yöntemiyle madde ayırt edicilik indeksi sonuçları yukarıda verilmiştir. Madde güçlük indeksi bir maddeyi doğru cevaplayan birey sayısının testin uygulandığı bütün bireylerin sayısına oranı olarak tanımlanabilir (Güler,

2017:130). Madde güçlüğü 1 ile 0 arasında değerler alır. Madde güçlüğü için .50 orta değer sayılır. Madde 0 değerine yaklaştıkça zorlaşır, 1 değerine yaklaştıkça ise madde kolaylaşır (Güler, 2017:131). Bu tanımdan hareketle testteki en kolay maddenin 0.88 ile M2 olduğu, en zor maddenin ise 0.33 ile 15. Madde olduğu görülmektedir. M1, M3, M11, M13, M14, M17, M19 ise orta düzeyde güçlük ile ideal bir güçlük değerine sahiptir. Testin geneline bakıldığında madde güçlük indeksi olarak uygun düzeyde olduğu görülmektedir. Madde güçlük indeksi formülü şu şekildedir;

$$P_j = \frac{\text{Doğru cevaplayan öğrenci sayısı}}{\text{Toplam öğrenci sayısı}}$$

(Güler, 2017).

Diğer bir ifade ile;

$$P_j = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı} + \text{alt grupta doğru cevaplayan öğrenci sayısı}}{\text{Toplam öğrenci sayısı}}$$

(Güler,2017).

Madde ayırt edicilik indeksi ise bilenlerle bilmeyenlerin birbirinden ayrıldığı indeks olarak tanımlanabilir. Bu indeks bir maddenin en önemli istatistiğidir. Maddenin güvenilirliğinin ölçüsünü verir. Maddenin testten atılıp atılmayacağına karar vermede kullanılan madde istatistiğidir (Güler, 2017). Madde ayırt ediciliği -1 ile +1 arasında değerler alabilir. Bu durumda +1 mükemmel ayırt edicilik olarak adlandırılırken -1 ise tersine bir ayırt ediciliğe işaret eder. Madde ayırt edicilik indeksi formülleri şu şekildedir;

$$r_{jx} = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı} - \text{Alt grupta doğru cevaplayan öğrenci sayısı}}{\text{Toplam öğrenci sayısı} / 2}$$

ya da

$$r_{jx} = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı} - \text{Alt grupta doğru cevaplayan öğrenci sayısı}}{\text{alt gruptaki öğrenci sayısı}}$$

ya da

$$r_{jx} = \frac{\text{üst grupta doğru cevaplayan öğrenci sayısı} - \text{Alt grupta doğru cevaplayan öğrenci sayısı}}{\text{üst gruptaki öğrenci sayısı}}$$

(Güler, 2017).

Madde ayırt edicilik indeksi parametreleri şu şekildedir;

$-1 \leq r_{jx} \leq 0$ ise madde bilenler lehine değil bilmeyenler lehine ayırt edicilik göstermiştir.

$0 \leq r_{jx} \leq 0.20$ ise maddenin teste alınması istenen bir durum değildir. Bu düzeydeki bir madde testten çıkarılmalıdır.

$0.20 \leq r_{jx} \leq 0.30$ ise maddenin orta düzeyde ayırt ediciliğinin olduğunu belirtir.

$0.30 \leq r_{jx}$ ise maddenin yeterli düzeyde ayırt ediciliğinin olduğunu belirtir (Güler, 2017: 131-133).

Bu tanımdan hareketle en az ayırt edicilik değerinin 0.30 ile M2’de olduğu, en çok ayırt edicilik değerinin ise 0.55 ile M20 olduğu görülmektedir. Alt – üst grup yönteminde en yüksek puan alan öğrenciler ile en düşük puan alan öğrencilerin oluşturduğu gruplar analize alınır. Öğrenci sayısının 300 – 400 ya da daha çok olduğu durumlarda bu yöntemin kullanılması beklenir (Güler, 2017).

Madde kalitesini değerlendirmek için madde ayırt edicilik indeksi kullanılmalıdır; güçlük indeksi değerleri madde zorluğunu değerlendirmek için kullanılmalıdır (Varma, 2003:7). Testin geneline bakıldığında madde güçlük indeksi değerlerinde ve madde ayırt edicilik değerlerinde sorun olmadığı, yeterli düzeylerde olduğu görülmektedir.

4.4.Mantel – Haenzsel Analizi Sonuçları

Açımlayıcı faktör analizi, güvenirlik analizi ve madde istatistikleri için yapılan analizlerden sonra R paket programı “DifR” paketi ile klasik test kuramına göre yanlılık kestirim yöntemlerinden olan Mantel Haenzsel yöntemi kullanılmıştır. Maddelerin cinsiyete göre yanlılık gösterip göstermediği analiz edilmiştir. Analizlerde referans grup olarak kız öğrenciler odak grup olarak ise erkek öğrenciler analize dahil edilmiştir. Çıkan sonuç tablo 10’da gösterilmiştir;

Tablo 9. Mantel – Haenzsel Analizi Sonuçları

Madde No	AlphaMH	Delta ΔMH	Düzy
M1	0.9039	0.2374	A
M2	0.4027	2.1374	C
M3	1.2717	-0.5649	A
M6	0.8574	0.3615	A
M7	1.4892	-0.9358	A

M10	0.9855	0.0343	A
M11	0.8241	0.4546	A
M12	0.9059	0.2323	A
M13	1.4056	-0.8001	A
M14	1.9822	-1.6078	C
M15	1.3409	-0.6893	A
M17	0.9373	0.1521	A
M18	0.8324	0.4312	A
M19	0.7525	0.6682	A
M20	0.6318	1.0790	B

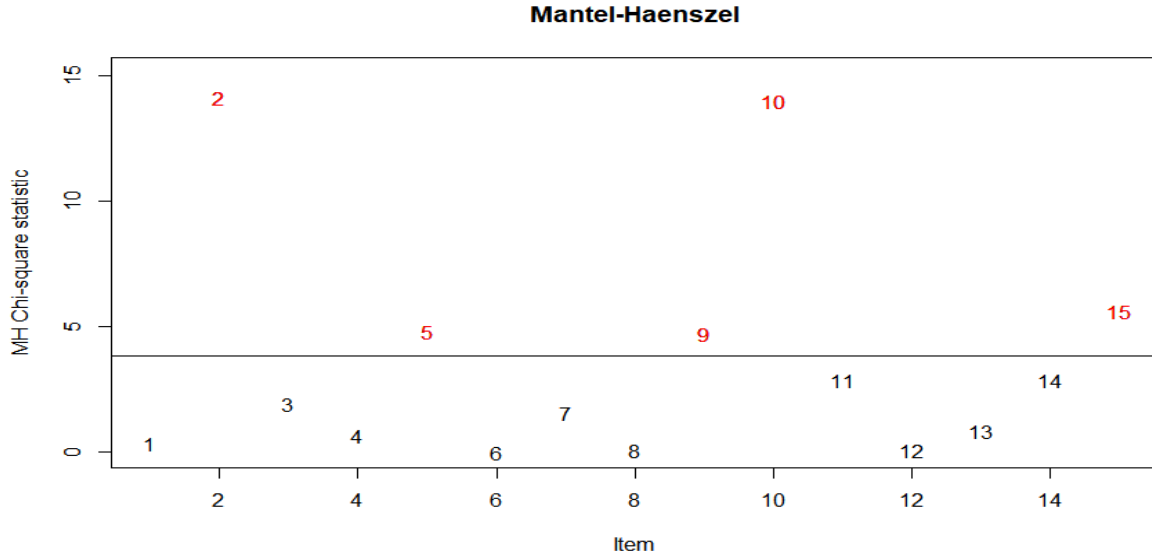
Analiz sonuçlarında da görüleceği üzere yanlı maddeler A, B, C diye sınıflandırılmıştır. Zwick and Erickson (1989) Delta (Δ) olarak adlandırdığı sıralamada şöyle bir durum olduğunu belirtmiştir;

$\Delta < 1$ ise A düzeyindedir yani yok denilebilir veya ihmal edilebilir düzeydedir.

$1 < \Delta < 1.5$ ise B düzeyinde yani orta düzeyde DMF vardır.

$1.5 < \Delta$ ise C düzeyindedir yani yüksek düzeyde DMF vardır demektir.

M2, M7, M13, M14, M20 maddelerinde yanlılık bulunmuştur. Yukarıdaki tabloya bakıldığında ise M2'nin C düzeyi kızlar lehine, M7'nin A düzeyi erkekler lehine, M13'ün A düzeyi erkekler lehine M14'ün C düzeyi erkekler lehine, M20'nin B düzeyi kızlar lehine yanlı çıktığı görülmektedir. MH yöntemine göre 10 maddede ise A düzeyinde yanlılık bulunmuştur. 15 maddelik bir testin 5 maddesinde yanlı madde bulunmuştur. Genel test oranına bakılırsa %33'lük bir madde oranı yanlı çıkmıştır. Bununla alakalı R paket programı ile yapılan MH analizinin görsel sonuçları şekil 6'da verilmiştir.



Şekil 6. Mantel Haenszel Analizi Sonuç Şekli

Şekil 6’da görüldüğü üzere maddeler 15 madde üzerinden gösterilmiştir. Yanlı maddeler sınırın üstünde yer alarak kırmızı renk ile gösterilmiştir. Maddelerde kayma unsuru olduğundan kaynaklı olarak yanlı maddeler çıkarılan maddelerden sonra karşılık olarak kullanılmıştır. Hangi maddenin testte hangi maddeye karşılık geldiği eşitlik işareti ile ifade edilmiştir. M2, M5=M7, M9=M13, M10=M14 ve M15=M20 yanlı olarak R paket programından görselleştirilmiştir. Bu oran madde sayısının %33’lük dilimini kapsamaktadır. Yazılan 15 maddenin %33’lük diliminde yanlı madde çıkması hazırlanan ve uygulanan test için bir sorun olarak göze çarpmaktadır. Bu aşamadan sonra LR analiz sonuçları aşağıda verilmiştir.

4.5.Lojistik Regresyon Analiz Sonuçları

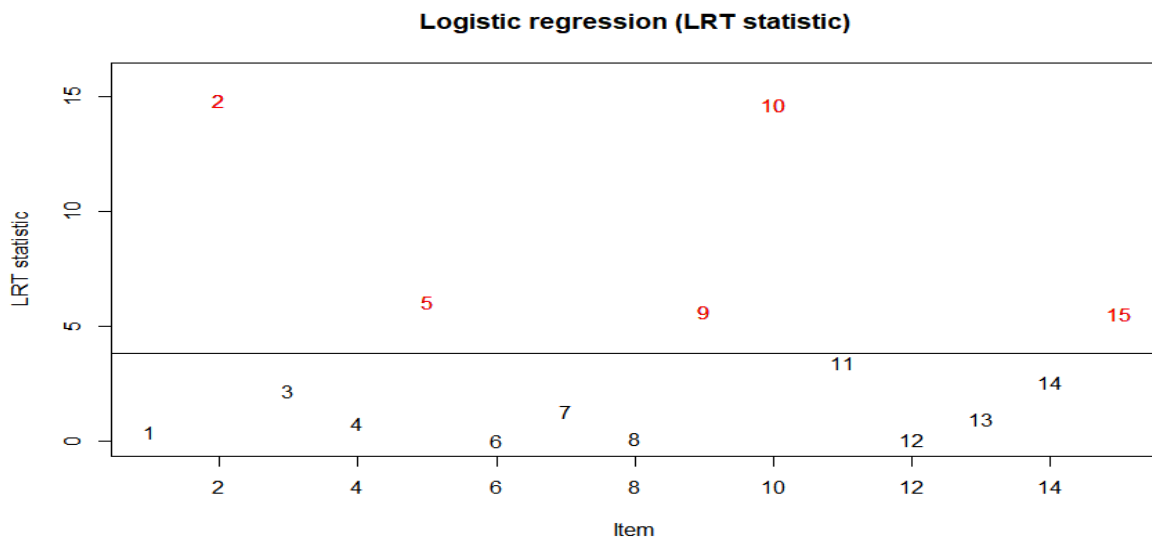
Klasik test kuramına göre yanlılık kestirim yöntemlerinden bir diğeri ise lojistik regresyon yöntemidir. Mantel – Haenszel yönteminde kullanılan parametreler bu aşamada da aynı şekilde kullanılmıştır. Lojistik regresyon yöntemi sonuçları tablo 11’de verilmiştir;

Tablo 10. Lojistik Regresyon Analizi Sonuçları

Madde No	İstatistik	Güçlük Değeri	Anlamlılık Düzeyi
M1	0.4150	0.5194	
M2	14.7991	0.0001	***
M3	2.1848	0.1394	
M6	0.7702	0.3801	

M7	6.0683	0.0138	*
M10	0.0134	0.9079	
M11	1.3256	0.2496	
M12	0.1328	0.7155	
M13	5.6325	0.0176	*
M14	14.6370	0.0001	***
M15	3.4410	0.0636	.
M17	0.1014	0.7502	
M18	0.9835	0.3213	
M19	2.5520	0.1102	
M20	5.5563	0.0184	*
Anlamlılık Düzeyi: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1			

Lojistik regresyon analizi için R paket programı ve “difR” paketi kullanılmıştır. Lojistik regresyon sonuçları tablo 11’ de verilmiştir. R paket programı ile yapılan Lojistik Regresyon sonuçlarına göre M2, M7, M13, M14, M20’de madde yanlılığı olduğu rapor edilmiştir. MH yöntemine göre 10 maddede ise A düzeyinde yanlılık bulunmuştur. 15 maddelik bir testin 5 maddesinde yanlı madde bulunmuştur. Genel test oranına bakılırsa %33’lük bir madde oranı yanlı çıkmıştır. Bir maddenin P=güçlük değeri, o maddeyi doğru yanıtlayan öğrencilerin oranını verir ve madde zorluğunun bir göstergesidir. Yanlı çıkan maddelerde P değerinin düşük olduğu gözlemlenmiştir. Bununla alakalı R paket programı ile yapılan LR analizinin görsel çıktısı aşağıda verilmiştir.



Şekil 7. Lojistik Regresyon Analizi Sonuç Şekli

Şekil 7’de de görüldüğü üzere maddeler 15 madde üzerinden gösterilmiştir. Yanlı maddeler sınırın üstünde yer alarak kırmızı renk ile gösterilmiştir. M2, M5=M7, M9=M13, M10=M14 ve M15=M20 yanlı olarak R paket programında görselleştirilmiştir. Aşağıda ise Zumbo and Thomas ve Jodoin and Gierl’in tanımladığı LR istatistiklerine göre LR sonuçları verilmiştir.

Tablo 11. Zumbo and Thomas ve Jodoin and Gierl indeksi analiz sonuçları

	R^2_{Δ}	ZT	JG
M1	0.0005	A	A
M2	0.0262	A	A
M3	0.0025	A	A
M6	0.0008	A	A
M7	0.0070	A	A
M10	0.0000	A	A
M11	0.0018	A	A
M12	0.0002	A	A
M13	0.0071	A	A
M14	0.0148	A	A
M15	0.0045	A	A
M17	0.0001	A	A
M18	0.0013	A	A
M19	0.0029	A	A
M20	0.0054	A	A

R paket programı ile yapılan analiz sonuçları tablo 12’de yer almaktadır. ZT= Zumbo ve Thomas, bir etki büyüklüğü ölçüsü veren R^2 ’nin ağırlıklı en küçük kareler tahminini bölümlere ayırmaya dayalı olarak LR prosedürü için DMF’nin etki büyüklüğünü ölçmek için bir indeks geliştirdi. Bu indeks, ilk olarak, LR modelindeki her terim için (yani, test puanı, grup üyeliği, test puanı-grup üyeliği etkileşimi) R^2 uygunluk DIF ölçüsünü hesaplayarak ve ardından R^2 ’nin her biri için bölümlenmesiyle elde edilir. Zumbo ve Thomas (1996) tarafından ağırlıklı en küçük kareler etki büyüklüğü, R^2_{Δ} için bir sınıflandırma sistemi önerildi; ihmal edilebilir DMF<.13’ün altında olmalıdır ve A düzeyidir. Orta düzeyde DMF, .13<DMF <.26 arasındadır ve B düzeyidir. Yüksek düzeyde DMF, .26’nın üzerinde olmalıdır ve C düzeyidir. LR DMF için

etki büyüklüğü, R^2_{Δ} , mevcut çalışmada analiz edilmektedir. Sonuç, tek boyutlu DMF'nin büyüklüğünü ölçen grup üyeliği ile ilişkili bir etki büyüklüğü ölçüsüdür. Tek boyutlu olmayan DMF'nin büyüklüğünü ölçen grup üyeliği etkileşimi için R^2 'nin çıkarılmasıyla toplam puan-grup üyelik terimi için ikinci bir DMF etki büyüklüğü üretilir. $R^2\Delta$, DMF'li $R^2\Delta$ öğeleri tanımlamak için LR anlamlılık testi ile kullanılabilir. JG= Jodoin ve Gierl (2002) R^2_{Δ} 'yi yorumlamak için deneysel olarak yönergeler oluşturmuştur. Model uyumu için ki-kare testi istatistiksel olarak anlamlı olmadığında veya $R^2_{\Delta} < 0,035$ olduğunda bir madde ihmal edilebilir veya A düzeyinde DMF'ye sahiptir. Ki kare testi istatistiksel olarak anlamlı olduğunda ve $0,035 \leq R^2_{\Delta} < 0,070$ olduğunda bir madde orta veya B düzeyinde DMF'ye sahiptir. Ki-kare testi istatistiksel olarak anlamlı olduğunda ve $R^2_{\Delta} \geq 0,070$ olduğunda bir maddenin büyük veya C düzeyinde DMF'si vardır. Bu yönergeler hem tek tip hem de tek biçimli olmayan DIF için geçerlidir (Ukanda vd., 2019: 878-887). LR DMF için etki büyüklüğü, $R^2\Delta$, mevcut çalışmada analiz edilmektedir. Etki büyüklüğü, R^2_{Δ} , LR DIF için doğru ve kararlı bir istatistik ise, o zaman LR DMF ve MH DMF yöntemlerinin birlikte kullanılması araştırmacılar için faydalı olabilir. Bir diğer lojistik regresyon varsayımı çoklu doğrusal varsayımdır. Aşağıda ise tablo 13'te sonuçlar verilmiştir.

4.6.Mantel – Haenszel ve Lojistik Regresyon DMF Karşılaştırma

Klasik Test Kuramı'nın bir maddenin değişen madde fonksiyonu içerip – içermediği ile ilgili yani yanlılık olup olmadığı ile ilgili kestirim yapma yöntemleri olarak mantel – haenszel ve lojistik regresyon kullanılabilir. Bu iki yöntemin birbiriyle uyumlu olması ve ortak karar verebilmesi istenen bir durumdur. Gök, Kelecioğlu ve Doğan (2010) çalışmasında MH ve LR yöntemlerinin karşılaştırılmasında yöntemlerin düşük düzeyde uyum gösterdiğini rapor etmişlerdir. Bu iki kestirim yöntemi yapılan araştırmada iyi bir uyum göstermiştir. R paket programı ile bu iki kestirim yönteminin karşılaştırma tablosu da verilmiştir. Karşılaştırma sonuçları şu şekildedir;

Tablo 12. MH – LR Karşılaştırma Tablosu

Madde	M-H	LR	#DMF
M1	NoDMF	NoDMF	0/2
M2	DMF	DMF	2/2
M3	NoDMF	NoDMF	0/2
M6	NoDMF	NoDMF	0/2
M7	DMF	DMF	2/2

M10	NoDMF	NoDMF	0/2
M11	NoDMF	NoDMF	0/2
M12	NoDMF	NoDMF	0/2
M13	DMF	DMF	2/2
M14	DMF	DMF	2/2
M15	NoDMF	NoDMF	0/2
M17	NoDMF	NoDMF	0/2
M18	NoDMF	NoDMF	0/2
M19	NoDMF	NoDMF	0/2
M20	DMF	DMF	2/2

Tablo 14’de görüldüğü üzere beş maddede iki yöntemin de DMF tespit ettiği görülmektedir. Tabloda NoDMF ifadesi maddelerin DMF göstermediğini ifade etmektedir. LR ve MH yöntem çıktılarına bakıldığında da bu yanlı maddelerin ortak bulunması hususunda %100 uyum gösterdikleri görülmektedir. M2, M7, M13, M14 ve M20’ de DMF bulunan maddelerdeki uyum kadar DMF bulunmayan maddelerdeki uyumu da net bir şekilde göze çarpmaktadır. MH ve LR yöntemlerinin uyumunun karşılaştırılması tablosundan sonra uzman görüş tablosunun verilmesi yerinde olacaktır.

4.7.Uzman Görüş Tablosu

Aşağıda verilen tabloda analizler yapılmadan önce madde yanlılığı alanında araştırma yapmış olan dört kişilik bir ölçme ve değerlendirme uzman grubuna ulaşılarak hangi maddelerin yanlı olduğu ve yanlılık düzeylerini belirtmeleri istenmiştir. Uzmanlar U1, U2, U3, U4 olarak kodlanmıştır. Açıklama olan kısımda ise uzmanların yanlılık düzeyi ve hangi cinsiyete yanlılık oluşturma ihtimali olduğunu belirttiği bölüm ise açıklama başlığı altında toplanmıştır. Yapılan çalışmada hangi gruba ne düzeyde yanlılık sağlayabileceği düşünülen maddeler Ö(Öngörülen) şeklinde kodlanmıştır. Madde yanlılığı olduğu düşünülen maddeler ve düzeyleri ile işaretlenmiştir. Tabloda cinsiyetlerin altında yer alan A düzeyi maddeler için yanlılık yorumlaması yapılırken yok denilebilir veya ihmal edilebilir düzeyi, B düzeyi maddeler için yanlılık yorumlaması yapılırken orta düzeyi, C düzeyi maddeler için yanlılık yorumlaması yapılırken yüksek olduğu düzeyi ifade etmektedir. Bu tabloda yer alan düzeylerden A düzeyi sorunsuz, ihmal edilebilir olarak yorumlanabilir. B düzeyi maddenin düzeltilmesi gerektiğini ifade ederken, C düzeyi ise maddenin testten çıkarılması gerektiği şeklinde yorumlanabilir. Yanlılık önemli düzeydeyse madde testten çıkarılmalıdır (Lord, 1980).

Tablo 13. Uzman Görüş Tablosu

		KIZLAR			ERKEKLER			Açıklama
M.N.	Yok	A	B	C	A	B	C	
M1	U1 U2 U3 U4	(Ö)						
M2	U1		(Ö) U2 U3 U4					U2: El sanatlarına kızların daha ilgili olabileceği varsayımını kabul edebiliriz. U3: Konudan ve seçeneklerden dolayı kız grubuna yanlı olabileceği kanısındayım. U4: Bu soru sanat ile ilgili olduğundan kızların daha yatkın olduğunu düşünüyorum.
M3					U1	U2 U3 U4	(Ö)	U1: Erkeklerin futbol konusuna aha fazla ilgisi olabileceğinden erkekler lehine yanlı olduğunu düşünüyorum. U2: Metin içeriği erkeklerin daha aşına oldukları durumları içeriyor. U3: Erkek öğrencilerin metin okurken kendi yaşantısıyla ilişkilendirebilir. U4: Erkek öğrencilere daha çok hitap eden bir spor olduğunu düşünüyorum.
M6	(Ö) U1 U2 U4	U3						U3: Kelime sorusu olduğundan düşük düzeyde kızlar lehine yanlılık oluşturabilir.
M7	U1 U2 U3 U4				(Ö)			
M10	(Ö)							

	U1 U2 U3 U4							
M11	U1 U2 U3 U4	(Ö)						
M12	(Ö) U1 U2 U3 U4							
M13	U1 U2 U4		U3			(Ö)		U3: Metindeki konunun şarkı söylemek ve kendini ifade etme temalarına yönelmesinden ve dolayı kızlara yönelik yanlı olabileceğini düşünüyorum.
M14	U2 U4				U3	(Ö)	U1	U1: Toplumsal roller/ilgiler gereği erkekler ekonomiyle daha çok ilgili olduğundan erkekler lehine olduğunu düşünüyorum. U3: Para, döviz, dolar gibi kavramlar erkek öğrencilerin daha çok ilgisini çekebilir.
M15	(Ö) U1 U2 U3 U4							
M17	U1 U4				U2	U3 (Ö)		U2: Savaş içeriği erkek öğrencilere daha aşına gelebilir. U3: Küçük yaşlardan itibaren erkek çocuklara savaşa dayalı oyun ve oyuncakların empoze edilmesi, kızların bu konuda ilgisinin düşük olmasından kaynaklı olarak yanlı olabilir.

M18	U1 U2 U4		(Ö)		U3			U3: Kız çocuklarına göre erkek çocukları politik konulara daha eğilimli olmakta, vatan millet gibi konularla daha çok ilgilenmektedir.
M19	U1 U2 U3 U4		(Ö)					
M20					U2 U4	(Ö) U3	U1	U1: Bilgisayar oyunları daha çok erkek öğrencilerin gerek bireysel gerekse grup olarak yaptıkları bir etkinliktir. U2: İçerik ve bağlam erkeklere daha aşina gelebilir. U3: Erkek öğrencilerin bilgisayar üzerinde fazla durması ve bu konuda daha fazla çevresinden uyarı alıyor olması etkili olabilir. U4: Erkeklerin konuyla daha çok ilgili olduğunu düşünüyorum.

Tablo 15’te incelenen görüş ve fikirler analiz yapılmadan önce uzmanlardan alınan görüşleri içermektedir. Bu tabloda görüldüğü üzere uzman görüşleri konusunda genel olarak bir fikir birliğinin olduğu söylenebilir. Cinsiyet konusunda %100 olarak ortak bir görüş bildirdikleri görülmekte ama düzey konusunda farklı fikirler olduğu görülmektedir ve bu öngörülen bir durumdur. M1, M7, M10, M11, M12, M15 ve M19’da düzeyin yok seviyesinde olduğunu söyleyen uzmanların hepsi bu fikirde birleşmişlerdir. M2 için ise sadece bir uzman görüş birliğinin dışında kalmıştır. M3 için ise sadece düzey konusunda ortak bir fikir birliğine varılamamıştır. M6 için sadece bir uzman A düzeyi olduğunu söyleyerek diğer uzmanlardan farklı bir görüş belirtmiştir. M13 için sadece bir uzman farklı görüş belirtmiştir. M14 ve M17 uzmanların ortak bir düzeyde görüş bildirmediikleri iki madde olarak göze çarpmaktadır. Uzmanların yarısı yanlılığın olmadığını düşünürken diğer yarısı ise olduğunu düşünüyor. Olduğunu düşünen uzmanlar ise düzeyi konusunda ortak bir görüşte karar kılammıştır. M18 için ise sadece bir uzman yanlılık olabileceğini belirtmektedir. M20 için ise bütün uzmanlar erkekler lehine yanlı olabileceğini belirtirken düzeyi konusunda bir ortak görüşte

buluşamamıştır. Bu aşamada ise karşılaştırmalı analiz sonuç tablosunun verilmesi yerinde olacaktır.

Tablo 14. Karşılaştırmalı Analiz Sonuç Tablosu

Madde No	Soru İçeriği	Öngörülen	Uzman Görüşleri	MH – LR Sonuçları
2	Ebru Sanatı	B düzeyi kızlar lehine	U2, U3, U4 B düzeyi kızlar lehine	C düzeyi kızlar lehine
7	Grafik ve Tablo	B düzeyi erkekler lehine	DMF yok ortak görüş	A düzeyi erkekler lehine
13	Konser – Günlük Yaşam	A düzeyi erkekler lehine	U3 B düzeyi kızlar	A düzeyi erkekler lehine
14	Döviz Kuru Grafiği	B düzeyi erkekler	U3 A düzeyi U1 C düzeyi erkekler	C düzeyi erkekler lehine
20	Bilgisayar Oyunları	B düzeyi erkekler	U2, U4 A düzeyi, U3 B düzeyi, U1 C düzeyi erkekler	B düzeyi kızlar lehine

BEŞİNCİ BÖLÜM

YORUM

Testte yer alan maddelerin her türlü uygulama durumunda hatalardan arınık olması istenir. Bu teorik olarak mümkündür ama pratik olarak olanaksızdır. Yanlılık ise teoride kaldırılması mümkün olan bir kavramdır. Ama durum pratik yönünde görüldüğü gibi değildir. Bundan kaynaklı olarak maddelerin ve dolayısıyla testin olabildiğince yanlılık ve diğer hatalardan arınık bir şekilde geliştirilmesi ve uygulanması istenir. Testin her türlü uygulamada kişilere veya gruplara cinsiyet olarak olsun, sosyoekonomik düzey olarak olsun hiçbir şekilde bir gruba veya bireylere avantaj veya dezavantaj sağlamaması istenir. Kişilerin aynı yetenek düzeyine sahip olduğu düşünülerek bu maddelerin avantaj veya dezavantaj sağlamaması

beklenen ve istenen bir özelliktir. Çünkü Türkiye gibi genç nüfusun çok olduğu ülkelerde seçme işlemi çok itinalı bir şekilde yapılmalıdır. Bir maddenin dahi bu durumlarda önemi artmaktadır. Merkezi sınavların ise yılda sadece bir kez yapıldığı göz önüne alınırsa bu konuda madde yanlılığı kişilerin sonraki eğitim ve sosyal hayatını etkileyebilmektedir. Yapılan çalışmalar incelendiğinde madde yanlılığı ile ilgili çalışmaların seçme ve yerleştirme işlemi sonuçlandıktan sonra yapıldığı gözlemlenmiştir. Bu durum ise bir sorun olarak göze çarpmaktadır. Bu amaçla yapılan bu çalışmada ise 8.sınıf düzeyinde eğitim gören öğrencilerle çalışma olanağı bulunmuştur. Uygulamadan önce ilgili araştırmalar incelenerek yanlı madde konuları ve soru tipleri incelenmiştir. Buna dayalı olarak kasıtlı bir şekilde yanlı maddeler yazılmaya çalışılmıştır. Yazılan sorular beş kişilik bir ölçme ve değerlendirme uzman grubu, iki kişilik Türkçe bölümü uzmanı ve iki kişilik Türkçe alan öğretmenleri tarafından değerlendirilmiştir. Alınan düzeltme ve onaylardan sonra deneme testi oluşturulmuştur. Uygulanacak test İl Milli Eğitim Müdürlüğü'nden izin alınarak belirlenen okullarda önce 245 kişilik bir deneme uygulaması olarak daha sonra ise 881 kişilik ana uygulama olarak uygulanmıştır. 245 kişilik deneme uygulamasından sonra ise AFA yapılarak faktör yapısının tek faktörlü olma durumu incelenmiştir. AFA sonucunda Büyüköztürk, Çokluk, Şekercioğlu, (2016)'na göre faktör yükü .30 altında olan beş madde testten çıkarılmıştır. Beş madde testten çıkarılınca kalan maddeler üzerinden betimsel istatistikler, madde istatistikleri, güvenirlik analizleri yapılarak nihai test oluşturulmuştur. 15 sorudan oluşan nihai test ise 881 kişilik örnekleme ulaşılarak ana uygulama olarak yapılmıştır. Toplanan verilerden bir öğrenci cinsiyet değişkenini işaretlemediği için testten çıkarılmıştır. Daha sonra toplanan veriler üzerinde herhangi bir işlem yapılmadan önce hangi maddelerin, hangi düzeyde yanlı olduğu ve sebepleriyle açıklamaları ile ilgili dört kişilik bir ölçme ve değerlendirme uzmanından görüş alınmıştır. Alınan görüşler ise Uzman Görüş Tablosu haline getirilmiştir. Uzmanlar U1, U2, U3 ve U4 olarak kodlanmıştır ve maddeler yazılırken düşünülen yanlılık düzeyi ve cinsiyeti (Ö) şeklinde kodlanarak tablo haline getirilmiştir. “Uzman Görüş Tablosu'ndan” da anlaşılabacağı üzere madde incelemesinde uzmanların cinsiyet olarak %100 uyumlu, düzey olarak ise yüksek düzeyde birliktelik sağladıkları görülmektedir. M1, M7, M10, M11, M12, M15 ve M19'da düzey olarak “DMF yoktur” şeklinde ifade eden uzmanların hepsi bu fikirde birleşmişlerdir. Yanlı olduğunu düşündükleri maddelerde ise cinsiyet konusunda %100 oranına uyum sağladıkları fakat düzey konusunda uyuşmazlık yaşadıkları görülmüştür. Toplanan veriler R paket programı ve DifR paketi ile Klasik Test Kuramı madde yanlılığı kestirme yöntemlerinden olan Mantel – Haenszel ve Lojistik Regresyon teknikleri ile analiz edilmiştir. Analizde referans grup olarak kız öğrenciler odak grup olarak ise erkek öğrenciler kabul edilmiştir. MH ve LR

sonuçlarına göre M2, M7, M13, M14, M20 yanlı çıkmıştır. MH ve LR yöntemlerinin %100 oranla çok iyi uyum sağlandığı görülmektedir. Yanlı çıkan maddelerden M2 ebru sanatı ile ilgili bir soru olarak göze çarpmaktadır. Yapılan araştırmalardan Ateşok Deveci (2008)' nin çalışmasında sanat ile ilgili maddelerin kızlar lehine yanlı çıkabileceği rapor edilmiştir. Bu durumda yazılan maddelerin bu cinsiyete yanlı çıkması beklenen bir değerdir. M2 için uzman görüşleri olarak U2, U3, U4 B düzeyi kızlar lehine diyerek doğru bir oran vermişlerdir. Aynı tabloda yer alan “öngörülen” düzeyi ile uzmanların %75'i aynı fikri savunmuştur ve doğru cinsiyet ve düzeyde ortak görüş belirtilmiştir. MH ve LR sonuçlarına göre madde kızlar lehine C düzeyinde DMF göstermektedir. M7 ise tablo ve grafik sorusu olarak görülmektedir. Akalın (2014) çalışmasında grafik ve tablo yorumlama sorularının erkekler tarafına yanlı çıkabileceğini belirtmiştir. Öngörülen düzey ise B düzeyi erkekler lehinedir. Uzmanlar bu madde için ise DMF yok diyerek görüş belirtmişlerdir. M7 konusunda “öngörülen” düzey B düzeyi olarak işaretlenmiştir. Uzmanlar ise %100 oranında “DMF yok” diyerek ortak bir noktada buluşmuşlardır. MH ve LR sonuçlarına göre ise M7’de erkekler lehine yanlılık bulunmuştur. M13; madde sanat sorusu gibi görünse de daha çok günlük hayat, yaşantı konuları ele alınarak yazılmıştır. Akalın (2014) çalışmasında günlük yaşam, yaşantı ile ilgili maddelerin erkekler lehine yanlı çıkabileceğini belirtmiştir. Yaşanılan coğrafya itibari ile konser, yaşantı unsurlarının erkekler lehine yanlı olabileceği varsayımı kabul edilerek madde yazılmıştır. MH ve LR sonuçlarına göre madde erkekler lehine DMF göstermektedir. M13 için ise sadece U3 B düzeyi olarak görüş belirtmiştir. Diğer uzmanlar bu madde ile alakalı ise %75 oranında “DMF yok” diyerek ortak görüş belirtmiştir. Ateşok Deveci (2008) çalışmasında döviz kuru maddelerinin erkekler lehine yanlı çıktığını rapor etmiştir. M14; bu madde için öngörülen düzey B düzeyi erkekler lehine olarak işaretlenmiştir. U3 A düzeyinde U1 ise C düzeyinde olduğu konusunda görüş belirtmiştir. U2 ve U4 ise “DMF yok” diyerek görüş belirtmiştir. Bu madde incelendiği zaman cinsiyetin doğru tahmin edildiği düzeyin ise sadece U1 tarafından doğru bulunduğu gözlemlenmektedir. MH ve LR sonuçlarına göre madde erkekler lehine C düzeyinde DMF göstermektedir. M20’nin konusu ise bilgisayar ve bilgisayar oyunları olarak görülmektedir. Yapılan araştırmalar neticesinde bilgisayar oyunlarının daha çok ilgi çektiğinden kaynaklı olarak “Öngörülen” düzey B olarak erkekler lehine olduğu düşünülmüştür. U2, U4 A düzeyi erkekler lehine, U3 B düzeyi erkekler lehine, U1 C düzeyi erkekler lehine DMF gösterdiği konusunda görüş belirtmiştir. Ancak MH ve LR sonuçlarına göre madde B düzeyinde kızlar lehine DMF göstermektedir. Yılmaz (2009) ise çalışmasında, çeşitli araştırma bulgularına göre; kız ve erkek öğrenciler akademik başarılarına göre karşılaştırıldığında, 10 yaşından itibaren kız öğrencilerin okuduğunu anlama gibi çeşitli

becerilerin kullanıldığı sözel alanlarda erkek öğrencilere göre daha başarılı olduklarını ifade etmektedir. 15 soruluk bir testin %33'e denk gelen beş madde ile DMF oluşturabilecek maddelerin yazılması, uzmanların ortak bir noktada fikir beyan ederek, cinsiyeti %80 oranında doğru tespit etmesi, düzey konusunda ise farklılıkların olması gayet olağan bir durumdur. Genel olarak bakıldığında uzmanların ve öngörülen düzeylerin görüş olarak ortak bir noktada bulunduğu ifade edilebilir. Tüm bu sonuçlardan hareketle test uygulanmadan önce madde yanlılığı düzeyi tam olarak kestirime sahip olmasa dahi cinsiyete göre DMF içerip içermediği bulunabilir şeklinde ifade edilebilir.

ALTINCI BÖLÜM

SONUÇ VE ÖNERİLER

- Bu araştırma LGS sınavına girecek 8.sınıf düzeyindeki öğrencilere yönelik Türkçe okuduğunu anlama testinin sınav uygulanmadan önce DMF içerip içermediğinin bulunması üzerine uygulanan bir çalışmadır.
- Deneme uygulaması ile 245 kişilik bir öğrenci grubuyla, ana uygulama için ise 881 adet öğrenci grubuyla çalışma olanağı bulunmuştur.
- 245 öğrenci ile tek boyutluluk varsayımını karşılamak için AFA analizleri yapılarak beş madde testten çıkarılmıştır.
- Nihai testin oluşturulduğu 15 maddelik test ile ana uygulama yapılmıştır.
- Uygulama yapıldıktan sonra uzmanlara hangi maddelerin, hangi cinsiyete göre ve ne düzeyde DMF içerip içermediğinin belirtilmesi istenmiştir.
- Toplanan uzman görüşleri alınarak tablo haline getirilmiştir.
- Uzman görüşleri de tablo haline getirilince, uygulama sonucunda toplanan veriler R paket programı ile “DifR” paketi kullanılarak, klasik test kuramı yanlılık kestirme yöntemlerinden MH ve LR ile analiz yapılmıştır.
- M2, M7, M13, M14 ve M20 DMF içeren maddeler olarak tespit edilmiştir. Uzmanların ise bir madde haricinde cinsiyeti doğru tahmin ettikleri ama düzey konusunda uyuşmazlık yaşadıkları görülmüştür.
- 15 soruluk testte %33 oranla 5 maddenin yanlı çıkması test için büyük bir problem olarak göze çarpmaktadır.
- Ama istenen durumun bu olduğu düşünüldüğünde bu konuda istenen sonuçların çıktığı gözlemlenmektedir.

- Karşılaştırmalı analiz sonuç tablosu aşağıda verilmiştir.

ÖNERİLER

- Bu çalışma 8.sınıf Türkçe alanından ziyade, Matematik, Sosyal Bilgiler ve Fen Bilimleri alanlarında yapılabilir.
- Maddeler yazılmadan önce literatür taraması yapılarak yanlışlık oluşturan konu başlıkları ele alınarak seviyeye uygun maddeler yazılabilir.
- Benzer çalışma 12. sınıf düzeyinde üniversite sınavına yönelik yapılabilir.
- Herhangi bir uygulama öncesinde uzman görüşü alınırsa yanlış maddeler tespit edilerek yapılan uygulama en azından bu durum için – hatalardan arınık bir hale getirilebilir.
- Başka bölge, başka şehirlerde daha büyük veya daha küçük gruplarla örneklem büyüklüğü de test edilerek uygulama yapılabilir.
- Araştırma yapılırken daha fazla uzmandan görüş alınabilir.

KAYNAKÇA

- Agresti, Alan (2002). Categorical data analysis. New Jersey: John Wiley and Sons, Inc. pp. 231–232.
- Ahmadi, A. and Jalili, T. (2014). A confirmatory study of differential item functioning on EFL reading comprehension. *Applied Research on English Language*. 3 (2), 55-68.
- Akarsu, B. (2018). Ölçme ve değerlendirme. İstanbul: Cinius Yayınları.
- Arıkan, S., Çelen, Ü., Gülleroğlu, H.D., Gültekin, S., Kilmen, S., Köse, İ.A. (2014). Çıkrıkçı Demirtaşlı, N. (ed.), Eğitimde ölçme ve değerlendirme. Ankara: Edge Akademi.
- Atılgan, H., Kan, A., Aydın, B. (2019). Atılgan, H. (ed.), Eğitimde ölçme ve değerlendirme. Ankara: Anı Yayıncılık.
- Avcu, A. (2021). Test geliştirmede modern yaklaşımlar. Ankara: Nobel Yayıncılık.
- Aydın, İ. (2013). Eğitim ve öğretimde etik. Ankara: Pegem Akademi.
- Aydın, İ. (2015). Alternatif okullar. Ankara: Pegem Akademi.
- Bakan Kalaycıoğlu, D. (2008). Öğrenci Seçme Sınavı'nın madde yanlılığı açısından incelenmesi (Doktora Tezi). Ankara: Hacettepe Üniversitesi.
- Balcı, A. (2016). Sosyal bilimlerde araştırma yöntem, teknik ve ilkeler. Ankara: Pegem Akademi.
- Baykul, Y. (2015). Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması. Ankara: Pegem Akademi.
- Beller, M. ve Gafni, N. (1996). 1991 International assessment of educational progress in mathematics and sciences: The gender differences perspective. *Journal of Educational Psychology*, 88(2), 365-377.
- Berberoğlu, G. (2006). Sınıf içi ölçme ve değerlendirme teknikleri, İstanbul: Morpa Kültür.
- Berger, M. Tutz, G. (2015). Detection of uniform and non-uniform differential item functioning by item focussed trees. Cornell University. [arXiv:1511.07178v1](https://arxiv.org/abs/1511.07178v1)
- Bobbitt, J.F. (2017). Eğitim programı (M.E. Rüzgar, Çev.). Ankara: Pegem Akademi.

Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. Kuram ve Uygulamada Eğitim Yönetimi. 32 (32), 470 – 483

Büyüköztürk, Ş. (2017). Sosyal bilimler için veri analizi el kitabı. Ankara: Pegem Akademi.

Büyüköztürk, Ş., Çokluk, Ö., Köklü, N. (2017). Sosyal bilimler için istatistik. Ankara: Pegem Akademi.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz, Ş., Demirel, F. (2017). Bilimsel araştırma yöntemleri. Ankara: Pegem Akademi.

Camilli G. and Shepard L. A. (1994). Methods for identifying biased test items (volume 4). California: SAGE Publications Inc.

Camilli, G. (2006). Test fairness: In R. L. (Ed.), Educational Measurement (4th ed., pp. 220–256). Westport, CT: American Council on Education.

Cohen, R. J. and Swerdlik, M.E. (2020). Psikolojik test ve değerlendirme (E. Tavşancıl, Çev.). Ankara: Nobel Yayıncılık.

Comrey, A.L. and Lee H. B. (1992) A first course in factor analysis. New Jersey: Lawrence Erlbaum Associates, Inc.

Crocker, L. and Algina, J. (2008). Introduction to classical and modern test theory. Ohio: Cengage Learning.

Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2016). Sosyal bilimler için çok değişkenli SPSS ve LISREL uygulamaları. Ankara: Pegem Akademi Yayınları.

Day N.E., Byar D.P. (1979). "Testing hypotheses in case-control studies-equivalence of Mantel–Haenszel statistics and logit score tests". Biometrics. 35 (3): 623–630.

Demirel, Ö. (2012). Eğitimde program geliştirme kuramdan uygulamaya (18.baskı). Ankara: Pegem Akademi.

Devellis, R.F. (2014). Ölçek geliştirme kuram ve uygulamalar. (T. Totan, Çev.). Ankara: Nobel Yayınevi.

Doğan, N. ve Öğretmen, T. (2008). Değişen madde fonksiyonunu belirlemede Mantel – Haenszel, ki-kare ve lojistik regresyon tekniklerinin karşılaştırılması. Eğitim ve bilim, 33(148), 100-112.

Dorans, N. J., and Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland and H. Wainer (Eds.), *Differential item functioning* (pp. 35–66). Hillsdale, NJ: Erlbaum.

Dziuban, Charles D.; Shirkey, Edwin C. (1974). "When is a correlation matrix appropriate for factor analysis? Some decision rules". *Psychological Bulletin*. 81 (6): 358–361.

Embretson, S.E. and Reise, S.P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum.

Erkuş A. (2016). *Psikolojide ölçme ve ölçek geliştirme I*. Ankara: Pegem Akademi.

Gafni, N. (1991). Differential item functioning: performance by sex on reading comprehension tests. *Paper presented at the 9th ACROLT Symposium* (ss 2-9). Jerusalem, Israel.

Gierl, M., Jodoin, G. M. and Ackerman, T. A. (2000). Performance of Mantel-Haenszel, Simultaneous item bias Test, and logistic regression when the proportion of DIF items is Large. Paper Presented at the Annual Meeting of the American educational Research Association (AERA). New Orleans, Louisiana, USA.

Gök, B., Kelecioğlu, H. ve Doğan, N. (2010). Değişen madde fonksiyonunu belirlemede Mantel– Haenszel ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim Dergisi* 35(156), 3-16

Güler, N. (2017). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.

Hixson M.D. and Glinchey, M.T. (2004). The relationship between race, Income, and oral reading fluency and performance on two reading comprehension measures. *Journal of psychoeducational Assessment*. 22 (4), 351-364.

Holland, P. W. and Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Eds.), *Test Validity* (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum.

Holland, P. W., and Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer and H. I. Braun (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ: Erlbaum.

Holland, P.W., ve Thayer, D.T. (1986). Differential item performance and the Mantel-Haenszel procedure (Technical Report No. 86–69). Princeton, NJ: Educational Testing Service.

Hosmer, David W.; Lemeshow, Stanley (2000). *Applied logistic regression* (2nd ed.). Wiley.

- Hovardaoğlu, S. (1994). Davranış bilimleri için istatistik. Ankara: Hatiboğlu Yayıncılık.
- James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). An introduction to statistical learning with applications in R. Springer Texts in Statistics. Vol: 103.
- Jodoin, M. G. and Gierl, M. J. (2001). Evaluating power and type I error rates using an effect size with the logistic regression procedure for DIF. *Applied Measurement in Education*, 14, 329-349.
- Jodoin, M.G. and Gierl, M.J. Size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education*, 14(4), 329-349.
- Kalaycıoğlu, D. B., and Kelecioğlu, H. (2011). Öğrenci seçme sınavının madde yanlılığı açısından incelenmesi. *Eğitim ve Bilim*, 36(161), 3-13.
- Karakaya, İ. (2012). Seviye belirleme sınavı'ndaki fen ve teknoloji ile matematik alt testlerinin madde yanlılığı açısından incelenmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(1), 215-229.
- Karakaya, İ. ve Kutlu, Ö. (2012). Seviye belirleme sınavındaki Türkçe alt testlerinin madde yanlılığının incelenmesi. *Eğitim ve Bilim*, 37(165), 342-368.
- Kelecioğlu, H., Karabay, B. ve Karabay E. (2014). Seviye belirleme sınavının madde yanlılığı açısından incelenmesi. *İlköğretim Online*, 13(3), 934-953.
- Keskin, S. (2018). Tetrakorik ve polikorik korelasyon katsayıları. *Sağlık bilimlerinde güncel akademik çalışmalar*. Vol:2 ss. 11-14
- Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). New York: Guilford.
- Koğar, H. (2021). R ile geçerlik ve güvenirlik analizleri. Ankara: Pegem Akademi.
- Leech, N.L., Barrett, K.C., Morgan, G.A. (2015). IBM SPSS for intermediate statistics: use and interpretation, Fifth Edition. New York: Routledge Publishing.
- Linn, R.L., Levine, M.V., Hastings, C.N. (1981). Item bias a test of reading comprehension. *Applied Psychological Measurement*. 5 (2), 159-173.
- Lord, F. M. (1980). Applications of item response theory to practical testing problems. Routledge Publisher.

Mantel, N. and Haenszel, W. M. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of National Cancer Institute*, 22:719- 748.

MEB (2019). Türkçe dersi öğretim programı (İlkokul ve Ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. Sınıflar). 01.01.2020 tarihinde <https://mufredat.meb.gov.tr/Programlar.aspx> adresinden alınmıştır.

MEB. (2021). Liselere Geçiş Sistemi (LGS) Merkezi sınavla yerleşen öğrencilerin performansı raporu 20.11.2021 tarihinde <http://odsgm.meb.gov.tr/www/raporlar/kategori/94> adresinden alınmıştır.

Mertler C.A. and Vannatta Reinhart, R. (2017). Advanced and multivariate statistical methods practical application and interpretation. New York: Routludge Publisher.

Nathan Mantel (September 1963). "Chi-Square tests with one degree of freedom, extensions of the Mantel–Haenszel procedure". *Journal of the American Statistical Association*. 58 (303): 690–700.

National Council on Measurement in National Council on Measurement in Education http://www.ncme.org/ncme/NCME/Resource_Center/Glossary/NCME/Resource_enter/Glossary1.aspx?hkey=4bb87415-44dc-4088-9ed9-e8515326a061#anchorC
<https://www.ncme.org/home>

Nitko, A.J. and Brookhart, S.M. (2013) Educational assessment of students. London: Pearson Education Limited.

Norris, Megan; Lecavalier, Luc (17 July 2009). "Evaluating the use of exploratory factor analysis in developmental disability psychological research". *Journal of Autism and Developmental Disorders*. 40 (1): 8–20.

Oral, B ve Yazar, T. (2017). Eğitimde program geliştirme ve değerlendirme. Ankara: Pegem Akademi.

Osterlind, S. (1983). Test item bias. Newbury Park: Sage Publications.

Osterlind, S. J., and Everson, H. T. (2009). Differentiatial item functioning. Thousand Oaks: CA: Sage.

Özçelik, D.A. (2013). Test hazırlama kılavuzu. Ankara: Pegem Akademi.

Özçelik, D.A. (2016). Ölçme ve değerlendirme. Ankara: Pegem Akademi.

Özdemir, E. (2013). Eleştirel okuma. Ankara: Bilgi Yayınevi.

- Pallant, J. (2015). SPSS survival manuel. Berkshire: McGrawhill
- Pallant, J. (2020). SPSS kullanma kılavuzu (S. Balcı, ve B. Ahi, Çev.). Ankara: Anı Yayıncılık.
- Perrone, M. (2006). Differential item functioning and item bias: critical considerations in test fairness. Teachers College, Columbia University Working Papers in TESOL and Applied Linguistics, 6, 1-3. Response Theory. Sage Publications, California.
- Reynolds, R.E., Taylor, M.A., Steffensen, M.S., Shirey, L.L., Anderson, R.C. (1982). Cultural schemata and reading comprehension. *Reading Research Quarterly*. 17 (3), 353-366. <https://doi.org/10.2307/747524>
- Rudner, Lawrence M., Convey, John J. (1978). An evaluation of select approaches for biased item identification. An Investigation of Item Bias in a Test of Reading:Comprehension. *Technical report No. 163*. Washington,D.C. (1-37).
- Salehi, M. and Tayebi, A. (2012). Differential item functioning (DIF) in terms of gender in the reading comprehension subtest of a high-stakes test. *Iranian Journal of Applied Language Studies*. 4 (1), 135-168.
- Suhr, D. D. (2006). Exploratory or confirmatory factor analysis (pp. 1-17). Cary: SAS Institute.
- Sünbül, S. Ö., and Sünbül, Ö. (2016). Değişen madde fonksiyonunun belirlenmesinde kullanılan yöntemlerde I. tip hata ve güç çalışması. *Elementary Education Online*, 15(3), 882-897.
- Swaminathan, H. ve Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361-370.
- Şenferah, S. (2015). 2010 Seviye belirleme sınavı matematik alt testli için değişen madde fonksiyonlarının ve madde yanlılığının incelenmesi (Yüksek Lisans Tezi). Ankara: Gazi Üniversitesi.
- Tabachnick, B. G, and Fidell, L.S. (2012). Using multivariate statistics (Sixth Edition). Pearson Education Limited. Edinburgh.
- Thompson, B. (2004) Exploratory and confirmatory factor analysis, American Psychological Association • Washington, DC
- Tolles, Juliana; Meurer, William J (2016). "Logistic regression relating patient characteristics to outcomes". *JAMA*. 316 (5): 533–534

Traub, R. (1997). Classical test theory in historical perspective. Educational Measurement: Issues and Practice 16 (4), 8–14.

Turgut, M. F. ve Baykul, Y. (2015). Eğitimde ölçme ve değerlendirme. Ankara: Pegem Akademi.

Tyler, R.W. (2014). Eğitim programlarının ve öğretimin temel ilkeleri ((M.E. Rüzgar ve B. Aslan, Çev.). Ankara: Pegem Akademi.

Ukanda, F., Othuon, L., Agak, J. and Oleche, Paul. (2019). Effectiveness of mantel haenszel and logistic regression statistics in detecting differential item functioning under different conditions of sample size, ability distribution and test length. *American Journal of Educational Research*. 7 (11): 878-887

Varma, S. (2003). Preliminary item statistics using point-biserial correlation and P-values. Educational Data System Inc. California (1-7)

Walker, S. H. and Duncan, D.B. (1967). Estimation of the probability of an event as a function of several independent variables. Oxford University Press. 54 (1/2): 167 – 179.

Wang, W., and Su, Y. Factors influencing the Mantel and generalized Mantel-Haenszel methods for the assessment of Differential Item Functioning in polytomous items. *Applied Psychological Measurement*, 2004, 28(6), 450-480.

William G. Cochran (December 1954). "Some methods for strengthening the common χ^2 tests". *Biometrics*. 10 (4): 417–451.

Worthington, Roger L.; Whittaker, Tiffany A J. (2006). "Scale development research: A content analysis and recommendations for best practices". *The Counseling Psychologist*. 34 (6): 806–838.

Yıldırım, H. (2015). 2012 yılı seviye belirleme sınavı matematik alt testinin madde yanlılığı açısından incelenmesi (Yüksek Lisans Tezi). Ankara:Gazi Üniversitesi.

Yılmaz, H. (2014). Eğitimde ölçme ve değerlendirme. Konya: Çizgi Kitabevi.

Yılmaz, H. B. (2009). On the development and measurement of spatial ability. *International*

Zumbo, B. D. (1999). A Handbook on the theory and methods of differential item functioning (DIF): logistic regression modeling as a unitary framework for binary and likert-type (Ordinal) item Scores. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

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

Zumbo, B.D. (2007). Three generations of differential item functioning (DIF) analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4, 223–233.

Zwick, R. and Ercikan, K. (1989). Analysis of differential item functioning in the NAEP History Assessment *Journal of Educational Measurement*, Vol. 26, No. 1., pp. 55-66.



EKLER

Ek 1. Pilot uygulama Soruları

1) Yeşilçam ne yazık ki geride kaldı. Haliyle o devirler, naziklik, naiflik, insanlık da o devirlerde kaldı. "Aile Şerefi" adlı filme bakınca her şeyi, sevgiyi, saygıyı, paylaşmayı, dayanışmayı görebiliyoruz. Bu filmde oynamış olan Şevket Altuğ ise görüşlerini şu şekilde ifade etmiştir; "Türk toplumunun değerleri değişti. Türk toplumuna sunulan işlerin içerikleri değişti. Yani ben şu andaki içeriklerle hiçbir dizinin içinde olamam. Eleştiri olarak kabul etsinler, biraz da yaşlılığımı versinler... Bütün yapılan işlerde tabanca, tüfek, millet birbirini öldürüyor. Bütün erkekler sakallı. Bizim zamanımızda sakal rol gerekirse bırakılırdı. Bu ortamda ben olamam. Çünkü biz yaptığımız işlerde topluma sevgiyi, hoşgörüyü, toleransı, birlikte yaşamayı, dayanışmayı öğretmeye çalıştık. Böyle bir senaryo ile karşılaşırsam yaşıma rağmen hâlâ oynayabilirim. Ama karşılaşacağımı da pek zannetmiyorum."

Yukarıdaki paragrafta Şevket Altuğ'un Türk sinemasıyla ilgili yakındığı temel problem nedir?

- A) Şimdiki işler eskisi gibi kaliteli değildir.
- B) Şu anda sanatımızda bir yozlaşma hâkimdir.
- C) Bütün dizi ve filmler şiddet unsuru içermektedir.
- D) Dizi ve filmler eskisi gibi değerleri öğretmiyor.

2) Ebru sanatı renklerin su üzerindeki dansı ve ahengi olarak tanımlanır. Ebru sanatının nasıl ortaya çıktığı tam olarak bilinmese de Osmanlı Dönemi'nde, kitapların iç kısımlarını süsleme amaçlı olarak kullanılmıştır. Özel boya ve fırça malzemeleriyle, su üzerine yapılan desenlerin kâğıda aktarılmasıyla yapılan bir sanat olarak karşımıza çıkmaktadır.

Aşağıdaki resimlerden hangisi yukarıdaki açıklamaya en uygundur?



A)



B)



C)



D)

3) Futbol görünüşte 90 dakika oynanan 11'e 11 rakiple oynanması üzerine geliştirilen bir oyundur. Sarı ve kırmızı kartlar vardır. İki sarı kart 1 kırmızı yapar ve kişi direk de kırmızı kart görebilir. Kırmızı kart gören oyuncu oyundan diskalifiye olur. Galibiyetler 3 puan, beraberlikler 1 puan ile ödüllendirilir. Mağlubiyete ise puan verilmez. Her şeyde olduğu gibi futbolda da özü kaybediyoruz. Futbol bir mantık, zekâ ve tutku oyunudur. Bazen kazanınca deliler gibi sevinirsiniz. Bazen kaybedince küçük çocuklar gibi ağlarsınız. Yine de her spor gibi zevk veren bir oyundur. Ama denildiği gibi sadece bir oyundur. Oyun bilincinin farkında olmak dileğiyle...

Aşağıdaki seçeneklerden hangisi yukarıdaki metinden kesin olarak çıkarılabilir?

- A) Futbol sadece zekâ ve mantık oyunudur.
- B) Kazanınca sadece seviniriz yenilince sadece üzülürüz.
- C) İki sarı kart veya kırmızı kart gören oyuncu oyundan atılır.
- D) Futbol sadece zevk veren bir oyundur.

4) **Aşağıdaki cümlelerin hangisi bir şiirden alınmıştır?**

- A) Sana gitme demeyeceğim
Ama gitme, Lavinia.
- B) Nerede benim çocukluğumun evleri
Haniymiş bizim eski komşular.
- C) Dün yine neredeydin
Seni evde bulamadım.
- D) Öğrenci sıraları yine yerindeydi
Sadece öğretmenler yoktu.

5) Duygularımızı yalnız ölümün uyandırdığına dikkat ettiniz mi? Bizden yeni ayrılmış dostlarımızı ne kadar severiz, değil mi? Ağzları toprakla dolup hiç konuşmaz olmuş hocalarımıza ne kadar hayranızdır! Saygı o zaman çok doğal olarak gelir, belki de tüm yaşamları boyunca bizden bekledikleri o saygı. Ama biliyor musunuz niçin ölümlere karşı hep daha dürüst ve cömertizdir? Nedeni basittir! Onlara karşı bir yükümlülüğümüz yoktur.

Aşağıdakilerden hangisi yukarıdaki metinden kesin olarak çıkarılabilir?

- A) Bizden yeni ayrılmış dostlarımızı hep severiz.
- B) Çok konuşmamalarının sebebi bizden saygı beklemleridir.
- C) Aslında sadece ölümlere karşı dürüst ve cömertizdir.
- D) Konuşmayan, ketum duran hocalarımıza sevgimizden bayılırız.

6) "Oynamak" sözcüğünün sözlükteki anlamlarını defterine yazmak isteyen bir öğrenci, bu sözcüğün mecaz anlamı olarak aşağıdakilerden hangisini yazmalıdır?

- A) Vakit geçirme, eğlenme, oyalanma vb. amaçlarla bir şeyle uğraşmak
- B) Herhangi birine karşı önemsemeyici davranışlarda bulunmak
- C) Eşyanın herhangi bir parçası kımıldamak, hareket etmek
- D) Bir şeyi sürekli evirip çevirmek veya sürekli olarak ona dokunmak.

7)Örnek

Tablo

Ülkeler	Obezite Oranı
Türkiye	%22,3
Almanya	%23,6
Rusya	%19,6
İspanya	%16,7
İsrail	%17,8
İngiltere	%26,9
Fransa	%15,3
OECD Ülkeleri Ortalaması	%19,5

Yukarıdaki tabloya göre aşağıdakilerden hangisinin çıkarımı yapılabilir?

- A) Obezite oranı en yüksek olan satır OECD ülkeleri ortalamasıdır.
- B) Hiçbir iki ülkenin toplamı en yüksek oranlı ülkeden daha düşük değildir.
- C) Türkiye'de obezite oranı giderek artmaktadır.
- D) Hiçbir ülkenin obezite oranı OECD Ülkeleri ortalamasından düşük değildir.

8) Bülbülü altın kafese koymuşlar ille de vatanım demiş.

Bu söz ile anlatılmak istenen kavram aşağıdakilerden hangisidir?

- A) Özgürlük
- B) Bireysellik
- C) Yurt sevgisi
- D) Parasallık

Evlilik kelimesi her toplum ve kültür için farklı bir anlam içerir. (I) Bizim kültürümüzde evliliğin, kolay alınabilecek bir karar olduğu gibi artan eğitim düzeyi ile birlikte zor verilen kararlar arasında yerini almaya başladığını görmekteyiz. (II) Üzerine yazılan binlerce makale de bulunan bir kavramdır. (III) Evlilik olgusu bizim toplumumuzda asırlar öncesinden bu güne gelen gizli kıstaslar içermektedir. (IV) Erkek ve kadın olmanın getirdiği toplumsal roller tarafların içinde bulunduğu sosyal statü, eğitim düzeyi, duyguların niteliği, aile yapıları ve kültürel unsurlar, hangi koşullarda evliliğe karar verildiği gibi önemli faktörler alınan evlilik kararını ciddi oranda etkilemektedir.

9) Yukarıdaki paragrafta kaçınıcı cümleden sonra metnin akışı değişmiştir?

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

10) "Sözde kalan dilek ve tasarıların iş bitirmede hiçbir etkisi olmaz." anlamına gelen atasözü aşağıdakilerden hangisidir?

- A) Hazıra dağlar dayanmaz.
- B) Büyük lokma ye, büyük söz söyleme.
- C) Taşıma su ile değirmen dönmez.
- D) Bal bal demekle ağız tatlanmaz.

11) İnsanoğlu yaradılışından beri hayvanları konuşturmaya onlarla iletişim kurmaya hatta kurabilmeye özlem duymaktadır. Bu da beraberinde Fabl metin türünü ortaya çıkarmıştır. İnsana verilmek istenen ders hayvanlar üzerinden verilmeye çalışılmıştır. Kaplumbağa tavşanı geçer insana çalış geride kalma dersi verilir. Tilki, kargayı kandırır, karga kendini Zeki Müren zanneder. Kimseye aldırma, elin sözüne güvenme mesajı verir. Ama bu mesajları alamamış olacak ki insanlar hala hayatına bunlardan uzak şekilde devam ediyor.

Aşağıdakilerden hangisi fabl türünün ortaya çıkmasının temel nedenidir?

- A) İnsanların kendine dost arayışı
- B) Hayvanlarla iletişim kurma isteği
- C) İnsanlara ders verme amacı
- D) Hayvanların değerli olduğunun bilinmesi isteği

12) Hayvanlar ve temsil ettiği simgeler

Beyaz Güvercin	Barış
Karınca	Çalışkanlık
Keçi	İnatçılık
Tilki	?

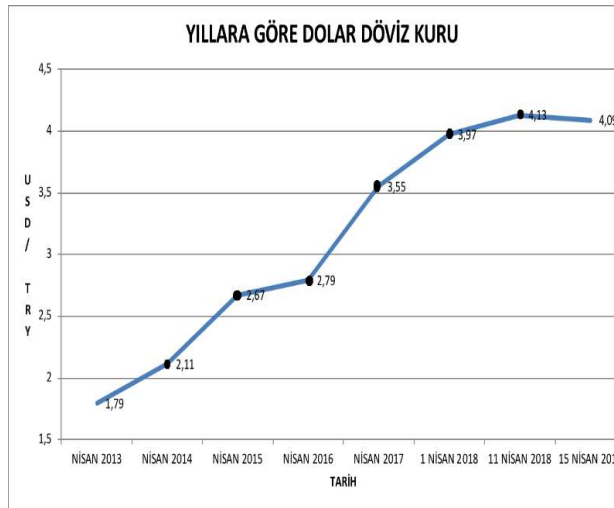
Yukarıdaki tabloya göre ? olan yere aşağıdakilerden hangisi gelmelidir?

- A) Açıkgozlülük
- B) Uyanıklık
- C) Kurnazlık
- D) Açgozlülük

13) Ben bu dünyaya şarkıcı olarak gelmedim. Benim birinci işim bu değil. Ben düşüncelerimi bu dünyaya aktarmak için geldiğime inanıyorum. Bu düşünceler bazen müzik eşliğinde daha güzel, daha şirin, daha hoş algılanıyor. Onun için bana verilen bu nimeti kullandım. Ben her şeyi gayet iyi anlattığımı düşünüyorum. Acaba şunu anlatamadım mı diye hiç kendimi sınırlandırmadım. Yazara göre sanatın temel işlevi aşağıdakilerden hangisidir?

- A) İnsanın kendini sınırlandırmamasıdır.
- B) Duygu ve düşünceleri hoş ve şirin aktarma amacıdır.
- C) Duygu ve düşüncelerin sadece müzik eşliğinde çok iyi anlatılacağıdır.
- D) İnsanlar arasında paylaşımı ve etkileşimi sağlamasıdır.

14) Yıllara Göre Döviz Kuru Grafiği



Yukarıdaki grafiğe göre aşağıdakilerden hangisinin çıkarımı yapılabilir?

- A) Dolar şu an 4 lira 50 kuruş seviyesini aşmıştır.
- B) Sürekli bir yükseliş vardır.
- C) En yüksek değer kazanması 66 kuruş olmuştur.
- D) Grafik belirli bir süreyi kapsamaktadır.

15) Zaman kavramı üzerine sayısız söz söylenmiş, sayısız kitaplar yazılmıştır. Mesela "Ne içindeyim zamanın. Ne de büsbütün dışında" derken kafası karışık mıydı Ahmet Hamdi'nin acaba? İşlerim gereği çok dizi izleyemem ama geçenlerde popüler bir dizide denk gelerek duyduğum repliği paylaşayım sizinle. Oradaki karakter Shakespeare'den esinlenerek "Çok zaman harcadım şimdi zaman beni harcıyor" diyor. Teoman ise zamanın nasıl hızlı geçtiğini şu şekilde ifade ediyor; "Nasıl oluyor; vakit bir türlü geçmezken, yıllar, hayatlar geçiyor?" Zaman günümüzde de özellikle gençler tarafından kıymetinin bilinmediği bir kavram olarak gözüme çarpıyor.

Yukarıdaki paragrafın üslubu ile ilgili aşağıdakilerden hangisine ulaşılabılır?

- A) Konu üzerinde birden fazla örnek verilmesi akışı bozmuştur.
- B) Farklı kişilerden farklı örnekler verilmesi konunun anlaşılmasını güçleştirmektedir.
- C) Metinde bir durumdan yakınma görülmektedir.
- D) Zaman kavramı farklı örneklerle iyice pekiştirilmemiştir.

16) Basketbol son yıllarda ülkemizde hızla yükselişe geçen herkesin ilgisini çeken bir spor dalı olarak görmekteyiz. Basketbol 5'e 5 oynanan 4x10=40 dakikalık bir periyot oyunudur. Basketbolda her galibiyet 2 puan mağlubiyet 1 puan şeklinde puanlanır. Basketbolda berabere kalmak durumu yoktur. Takımlar yenişemezse 5'er dakikalık uzatmalar oynanır. Biten ilk uzatmada kim önde ise o kazanır.

Parçaya göre hangisi basketbol oyununun temel kurallarından biri değildir?

- A) Basketbolda aynı sayıda oyuncunun oynaması şarttır.
- B) Periyot olarak eşit parçalara bölünmüş zaman diliminde oynanır.
- C) Basketbol oyunu berabere bitmez.
- D) Uzatmalar normal periyodun 8/1'i kadardır.

17) Savaş insanoğlunun tarihi kadar eskidir. İnsanoğlu yapısı gereği kendi yaşamı için başka yaşamları yok etmeyi kabullenmiştir. Bu da ölme ve öldürme olgusunu ortaya çıkarmıştır. "Mezopotamya'da kan tarihle eşit" derler bu konu hakkında. İnsanoğlu bu durum için önce teknolojiyi geliştirmiştir. Geliştirdikçe de öldürme usullerini geliştirmiştir. Gariptir ki insanoğlu yaşatmanın bir tek yolunu bulamazken öldürmenin çeşitli yollarını bulmuş ve geliştirmiştir.

Parçaya göre insanoğlunun savaşma stratejileri üretmesi ve bu konuda çeşitli teknolojik silahlar üretmesinin temelinde hangi güdü yer almaktadır?

- A) Yok Etme
- B) Saf Dışı Bırakma
- C) Kıskanma
- D) Zarar Verme

18) *Ey mavi göklerin beyaz ve kıvırl süsü,
Kız kardeşimin gelinliği, şehidimin son örtüsü.*
Yukarıdaki dörtlük hangi resimle ilgili yazılmıştır?



19) Tiyatro, kökeni hala çözülememiş bir sanat olarak ilgiyi hep canlı tutmuştur. Ne zaman, nerede başladığı hakkında kesin ve net bilgiler yoktur. Tiyatro bütün sanat dalları içinde kendine saygın bir yer edinmiştir. Tek unsuru insan olan, sahnede acı, tatlı, hüznü, neşeli, her şeyin anlatıldığı tiyatro eşsiz bir sanattır. Tiyatro'nun ana felsefesi şudur: Tiyatro;

Yukarıdaki paragrafın devamının hangisi ile bitirilmesi en uygun olur?

- A) Her şeyi istisnasız sahnede gösterme sanatıdır.
- B) Gelişen ve gelişecek olan bir sanat olmalıdır.
- C) Geçmişinin çözülmemesi ve geçmişinin merak edilmemesi gereken sanattır.
- D) İnsanı insana insanla insanca anlatma sanatıdır.

20) Bilgisayar oyunları bağımlılıktır ve sıradan bir bağımlılık değildir. Kişinin kısıtlı hayatından yılları götürür, öyle gizliden de yapmaz bunu. Açık açık, gözünüzün içine bakarak, götürür sizden geri dönüşü olmayan zamanı. Özellikle uzun süren ve bir türlü sonu gelmeyen savaş oyunları ve askeri oyunlarla haftalar, aylar harcanmaktadır. Haber sitelerinde ise bilgisayar başında uzun zaman geçirerek ölenlerin sayıları giderek atmaktadır.

Yukarıdaki paragraftaki gibi düşünen bir kişinin aşağıdakilerden hangisini söylemesi beklenemez?

- A) Bilgisayar oyunları insanların zamanlarını alırken onları meşgul eder.
- B) Bilgisayar oyunları her bağımlılık gibi sıradandır.
- C) Bilgisayar bağımlılığı bizi göz göre göre tüketir.
- D) Dünyada bilgisayar başında ölümler görülmektedir.

Ek.2. AYSE Başvuru Formu

HASAN KALYONCU ÜNİVERSİTESİNE

05/10/2020



BAŞVURU NO	20201005453266169
ÜNİVERSİTE ADI	HASAN KALYONCU ÜNİVERSİTESİ
ENSTİTÜ ADI	Sosyal Bilimler Enstitüsü
BÖLÜM ADI	Eğitimde Ölçme ve Değerlendirme
ÜNVAN	Öğrenci
TC KİMLİK NUMARASI	14603137244
KONU	Madde yanlılığının sınav uygulamasından önce bulunması üzerine bir çalışma
ARAŞTIRMA TÜRÜ	Yüksek Lisans Tezi
ÖRNEKLEM GRUBU	Öğrenci,
KAPSAMI	MEB Merkez Teşkilatı, İl Millî Eğitim Müdürlüğü, İlçe Millî Eğitim Müdürlüğü, Okul/Kurum, MEB Yurtdışı Teşkilatı,
İLLER	GAZİANTEP
KURUM TÜRLERİ	Resmî Ortaokul, Resmî İmam - Hatip Ortaokulu,
İLETİŞİM BİLGİLERİ	Adres:Bülbülzade Mahallesi 136012 nolu cadde no:10 Yusuf Apartmanı 1.Kat Daire:4 Dış Kapı Şifresi: 2525- Telefon:(543) 961-7878- Eposta:boztimur91@gmail.com

Ektrak Tarih ve Sayısı :16.11.2020 - E.2011160013 Yazının Ekidir

Yukarıda bilgileri bulunan proje uygulamaları için Millî Eğitim Bakanlığınızdan gerekli izinlerin alınması hususunda gereğini bilgilerinize arz ederim.

Ek listesi

Tez Önerisi

Bu belge, güvenli elektronik imza ile imzalanmıştır.
http://ebys.bku.edu.tr/Doordama/Index?EktrakNo=E.2011160013&ErisimKodu=3a03148c

Ek.3. Sosyal Bilimler Enstitü Müdürlüğü Dilekçe

T.C
HASAN KALYONCU ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü Müdürlüğü'ne

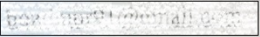
 Konu: Yüksek Lisans Tez Çalışması

08/10/2020

Enstitünüzde Eğitimde Ölçme ve Değerlendirme bölümü yüksek lisans programında Dr. Öğretim üyesi Ufuk AKBAŞ danışmanlığında *"Madde yanlılığının sınav uygulamasından önce bulunması üzerine bir çalışma"* başlıklı yüksek lisans tez çalışmasını Gaziantep ilindeki Milli Eğitim Bakanlığı'na bağlı resmi ortaokul, imam hatip ortaokulu ve bünyelerindeki 8.sınıf öğrencileri ile yapmak istiyorum. Gereğinin yapılmasını arz ederim.

Öğrenci no: 172110014

Cep Tel: 

Mail: 

Behçet ÖZTİMUR
14903137244

Ekler: 1. Sınav uygulama formu (1 adet)
2. Ayşe başvuru formu (1 adet)

Evrak Tarih ve Sayısı : 16.11.2020 - E.2011160013 Yazının Ekidir

Ek.3. Gaziantep Valiliği İl Millî Eğitim Müdürlüğü İzin Yazısı



T.C.
HASAN KALYONCU ÜNİVERSİTESİ
Sosyal Bilimler Enstitüsü Müdürlüğü

Sayı :E-58446768-804.01-2011160013
Konu :Uygulama İzni Hak.

Tarih:16.11.2020

GAZİANTEP VALİLİĞİ
İl Millî Eğitim Müdürlüğüne

Sosyal Bilimler Enstitüsü Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Tezli Yüksek Lisans programı öğrencilerinden 172110014 numara ile kayıtlı Behçet ÖZTİMUR' un; **"Madde Yanlılığının Sınav Uygulanmasından Önce Bulunmuş Üzerine Bir Çalışma"** konulu Yüksek Lisans çalışmasına veri oluşturmak amacıyla, ilinizde bulunan Millî Eğitim Müdürlüğüne bağlı resmi ortaokul, imam hatip ortaokulu ve ilköylerinde 8.sınıf öğrencilerinden ekteki anket ve ölçekleri uygulamak istemektedir.

Adı geçen öğrencimizin yukarıdaki sözü edilen çalışmaya katılabilmesi için gerekli iznin verilebilmesi hususunda gereğini arz ederim.

Doç.Dr Yakup DURMAZ
Müdür

Ekler:
1-Behçet ÖZTİMUR Dilekçe.pdf (Elektronik Ek)
2-Behçet ÖZTİMUR.pdf (Elektronik Ek)

Ek.4. Araştırma İzni Taahhütnamesi

MİLLÎ EĞİTİM BAKANLIĞINA BAĞLI OKUL VE KURUMLARDA GERÇEKLEŞTİRİLECEK ARAŞTIRMA UYGULAMALARINA İLİŞKİN ARAŞTIRMA İZNİ BAŞVURU TAAHHÜTNAMESİ

1. Araştırmam boyunca anayasa/kanun ve yönetmeliklere uygun davranacağımı,
2. Araştırmayı yürüteceğim okulun/kurumun kurallarına uyacağımı,
3. Araştırmam boyunca hiç kimseyi araştırmama/çalışmama katılmaya zorlamayacağımı,
4. Araştırmayı/çalışmayı bana tahsis edilen mekân/sınıf ve zamanda gerçekleştireceğimi,
5. Araştırmanın olası fiziksel/ruhsal zararları konusunda katılımcıları bilgilendireceğimi,
6. Araştırmam/ çalışmam sırasında topladığım kişisel bilgileri koruyacağımı,
7. Araştırmam/çalışmam için gerektiği kadar veri toplayacağımı,
8. Araştırma/çalışma sırasında öğrencilerin derslerinde/çalışmalarında herhangi bir kayıplarının olmayacağını,
9. Araştırmam/çalışmam sırasında herhangi bir ticari faaliyette bulunmayacağımı, katılımcıları herhangi bir ürün/eser/tedaviye yönlendirmeyeceğimi,
10. Araştırma izin evraklarını okul yönetimine teslim edeceğimi,
11. Araştırma/çalışma sırasında izni olan evrakları kullanacağımı,
12. Tıbbi araştırmalarda araştırma/çalışmanın uygulama sırasında etik kurallara uyacağımı,
13. Araştırma/çalışma sırasında topladığım ses ve görüntü kayıtlarını güvenilir ortamlarda saklayacağımı ve araştırma/çalışma sonrasında imha edeceğimi,
14. Genelge hükümlerine aykırı davranmam ve herhangi bir yanlış ifade, beyan ve maddi gerçeği gizleme gibi durumlarda adli ve idari işlemlerin yürütülmesini kabul edeceğimi,
15. İzin alınmış araştırmalarda/projelerde insanlarla ilgili yapılacak anket, görüşme, gözlem, alan araştırması, uygulama ve incelemelerde sağlık, güvenlik, insan hakları, mevcut mevzuat hükümleri, hukukun genel ilkelerini ihlal etmeyeceğimi ve etik ilkelere uyacağımı,
16. Araştırma ile ilgili sonuç raporlarını çalışmanın bitiş tarihinden itibaren 30 gün içinde izin aldığım birime ulaştıracağımı,

Kabul ettiğimi beyan ederim.

Araştırmanın Adı : Madde Yanlılığının Sınav Uygulamasından Önce Bulunması
Üzerine Bir Çalışma

Araştırmacı

Tarih:

İmza:

İsim – Soyisim:

05/03/2021

Behçet ÖZTIMUR

Ek.5. Örneklem Toplanacak Okullar Listesi

“Madde Yanlılığının Sınav Uygulamasından Önce Bulunması Üzerine Bir Çalışma” uygulaması için örneklem alınarak, veri toplanmak istenen okullar listesi aşağıdadır;

Şahinbey İlçesi Okul Listesi

23 Nisan Ortaokulu
Adnan Menderes Ortaokulu
Akkent Ortaokulu
Bülbülzade İmam Hatip Ortaokulu
Feridun Oral Ortaokulu
Şehit Samet Kaymakçı İmam Hatip Ortaokulu
Kazıklı Ortaokulu
Kocatepe Ortaokulu
Nurel Enver Taner Ortaokulu
Turgut Özal Ortaokulu
Vali M. lütfullah Bilgin Ortaokulu
Mehmet Görmez İmam Hatip Ortaokulu
Şehit Serdar Gören Ortaokulu
Sena Büyükkonuk Ortaokulu
Şehit Uğur Kutku Ortaokulu

Şehitkamil İlçesi Okul Listesi

Adnan A.ş. Ortaokulu
Ali Küncülü Ortaokulu
Aşlı Alevli Ortaokulu
Batıkent Ortaokulu
Fadiloğlu İmam Hatip Ortaokulu
Mahmut Hümayun Özhelvacı Ortaokulu
Osmangazi Ortaokulu
Şehit Hüseyin Gümüş İmam Hatip Ortaokulu
Ünler Ortaokulu
Kaşibeyaz Ortaokulu
Şehit Muhammed Emin Çelik İmam Hatip Ortaokulu
Mehmet Tuncay Aykaç Ortaokulu

Ek. 6. İl Milli Eğitim'in Gaziantep Valiliğine Yazdığı İzin Yazısı

	T.C. GAZİANTEP VALİLİĞİ İl Milli Eğitim Müdürlüğü
Sayı : E-34659092-605.01-22559811	17/03/2021
Konu : Araştırma İzin Talebi (Behçet ÖZTİMUR)	
VALİLİK MAKAMINA	
İlgi: Hasan Kalyoncu Üniversitesi Rektörlüğünün 16.02.2021 tarihli ve 2011160013 sayılı yazısı.	
Hasan Kalyoncu Üniversitesi Sosyal Bilimler Enstitüsü Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Tezli Yüksek Lisans programı Öğrencilerinden Behçet ÖZTİMUR'un yürüttüğü "Madde Yanlılığının Sınav Uygulamasından Önce Bulunması Üzerine Bir Çalışma" konulu Anket uygulama isteği kapsamında, İlimiz Şahinbey ve Şehitkamil İlçesinde bulunan ekli listede isimleri belirtilen okullarda öğrenim gören öğrencilere yönelik okul idaresinin gözetiminde ve bilgisi dahilinde araştırma çalışma isteği, ilgi yazıda belirtilmektedir. Bu kapsamda bahsi geçen anket uygulama isteği, Bakanlığımız Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 21.01.2020 tarihli ve 2020/2 sayılı genelgesi kapsamında değerlendirilmiş olup; araştırmacının, araştırmasının bitiminden itibaren 15 gün içerisinde araştırma sonuçlarını 2 kopya halinde CD içerisinde Müdürlüğümüze bildirmesi şartıyla, İlimiz Şahinbey ve Şehitkamil İlçesinde bulunan ekli listede isimleri belirtilen okullarda öğrenim gören öğrencilere yönelik okul idaresinin gözetiminde ve bilgisi dahilinde anket uygulama isteğinin, eğitim öğretimi aksatmayacak şekilde gönüllülük esasına göre uygulanması Müdürlüğümüz Ar-Ge bürosu bünyesinde oluşturulan komisyonun uygunluk raporu doğrultusunda uygun mütalaa edilmektedir. Makamınızca da uygun görüldüğü takdirde; Olurlarınıza arz ederim.	
Yasin TEPE İl Milli Eğitim Müdürü	
OLUR Rızzan EROĞLU Vali a. Vali Yardımcısı	

Ek.7. Sınav Yönergesi

	Okul Adı:	
	Sınıfı:	
	Okul No:	
	Dosya No:	
	Cinsiyet: Kız ☐ Erkek: ☐	

Sevgili öğrenciler bu form Türkçe alanında oluşturulan bir okuduğunu anlama testidir. Lütfen testi dikkatlice okuyup inceleyiniz. Hiçbir bilgi hiçbir şekilde kimseyle paylaşılmayacaktır. Herhangi bir puanlama not verme olmayacaktır. Aşağıdaki sınav kurallarını lütfen dikkatlice okuyunuz. Sınavla alakalı bütün soru görüş ve önerileriniz için en altta yer alan telefon numarasını arayabilirsiniz. Yine aşağıda yer alan mail adresime mail atabilirsiniz. Hepinize şimdiden teşekkür ediyorum. Bütün öğrencilere başarılar dilerim. 😊😊😊

Sınav Kuralları

Sınav formu 20 sorudan oluşmaktadır.

Sınav süresi 40 dakikadır.

Her soru 4 şıktan oluşmaktadır.

Her sorunun bir doğru cevabı vardır.

Aynı soruya birden fazla cevap vermeniz durumunda cevabınız geçersiz sayılacaktır.

Lütfen bütün soruları cevaplayınız.

Yanlışlar doğruyu götürmemektedir.

Lütfen cevaplarınızı hem kitapçığa hem de optik forma kopyalayınız.

Sorularda puanlama, not verme hiçbir şekilde olmayacaktır.

Sınav esnasında kalem, silgi vs alışverişi yasaktır.

Behçet ÖZTİMUR
Cep Telefonu:
Mail Adresi:

Evrak Tarih ve Sayısı :16.11.2020 - E.2011160013 Yazının Ekidir

Bu belge, güvenli elektronik imza ile imzalanmıştır.
http://ebys.bku.edu.tr/Doculama/Index?EvrakNo=E.2011160013&ErsimKodu=3a03148c