

**YARATICILIĞIN VE MADDE YAZARLIĞI EĞİTİMİNİN
ÖLÇME ARACININ PSİKOMETRİK ÖZELLİKLERİNE
ETKİSİ**

**THE EFFECT OF CREATIVITY AND EDUCATION IN ITEM
WRITING ON MEASUREMENT TOOL'S PSYCHOMETRIC
PROPERTIES**

Kenan Burak YÜKSEL

Hacettepe Üniversitesi

Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

Yüksek Lisans Tezi

olarak hazırlanmıştır.

2017

KABUL ve ONAY

Eđitim Bilimleri Enstitüsü M¼d¼rl¼ę¼'ne,

Kenan Burak Y¼KSEL'in hazırladıđı "Yaratıcılık ve Madde Yazarlıđı Eđitiminin ¼lęme Aracının Psikometrik ¼zelliklerine Etkisi" bařlıklı bu ęalıřma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde ¼lęme ve Deđerlendirme Bilim Dalı'nda Y¼ksek Lisans Tezi** olarak kabul edilmiřtir.

Bařkan Yrd. Doę. Dr.,Hamide Deniz G¼LLEROĐLU

¼ye
(Danıřman) Prof. Dr.,Nuri DOĐAN

¼ye Yrd. Doę. Dr.,K¼bra ATALAY KABASAKAL

ONAY

Bu tez Hacettepe ¼niversitesi Lisans¼st¼ Eđitim-¼đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından **13.06.2017** tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca/...../..... tarihinde kabul edilmiřtir.

Prof. Dr. Ali Ekber řAHIN
Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Tezimin tamamı dünya çapında erişime açılabilir ve bir kısmı veya tamamının fotokopisi alınabilir.

(Bu seçenекle teziniz arama motorlarında indekslenebilecek, daha sonra tezinizin erişim statüsünün değiştirilmesini talep etmeniz ve kütüphane bu talebinizi yerine getirirse bile, teziniz arama motorlarının önbelleklerinde kalmaya devam edebilecektir)

Tezimin 01/06/2018 tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç Kapak, Özet, İçindekiler ve Kaynakça hariç) istemiyorum.

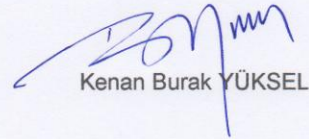
(Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisi alınabilir).

Tezimin 01/06/2018 tarihine kadar erişime açılmasını istemiyorum ancak kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisinin alınmasını onaylıyorum.

Serbest Seçenek/Yazarın Seçimi:

.....

20/06/2017

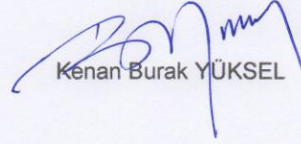

Kenan Burak YÜKSEL

ETİK BEYANNAMESİ

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.


Kenan Burak YÜKSEL

TEŞEKKÜR

Bana, yüksek lisans eğitimi görme ve bir bilimsel eser yazma onuruna erişebilme imkânı sunan, mensubu olmaktan daima şeref ve gurur duyduğum Kara Kuvvetleri Komutanlığı'nın manevi şahsına,

Öğrenirken eğlenebilmeyi, sayılarla anlaşabilmeyi ve ölçmeyi öğreten Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Eğitimde Ölçme ve Değerlendirme Programı akademisyenleri, değerli hocalarım, başta Prof.Dr. Selahattin GELBAL olmak üzere Prof.Dr. Hülya KELECİOĞLU, Doç.Dr. Duygu ANIL, Doç.Dr. Cem Oktay GÜZELLER ve Yrd.Doç.Dr. Derya ÇOBANOĞLU AKTAN'a,

Mesai, zaman ve konum mefhumu gözetmeksizin, bu tezin hazırlanmasında yardım, destek, ilgi, motivasyon ve daha sayamayacağım birçok katkıları bulunan değerli danışmanım Prof.Dr. Nuri DOĞAN'a,

Araştırmalarımnda büyük yardımlarını gördüğüm başta Gözde ALAGÖZ hocama, yüce dost Süleyman KESER'e, kardeşim Burhan YÜKSEL, sevgili eşi Sinem ve yavruları Mercan Ela'ya, altın kızlar Emine-Kevser-Neslihan AKTAN'a; büyük Fenerbahçeli Bahadır, değerli eşi Hacer ve araştırmamın gizli kahramanı Elif Ece ASLAN'a, tek kız kardeşim Burcu KILINÇ'a,

Bu araştırmamın bilimsel boyutunda katkılarını aldığım başta Prof.Dr. Esra ALAN olmak üzere isimlerini burada belirlemediğim, alan uzmanları, öğretmenler ve değerli öğrencilere,

Dünüm, varlık sebebim, destekleri ve ilgileri ile büyüdüğüm, bana, özellikle o güzel yemekleriyle güç veren kıymetli annem Kıymet YÜKSEL'e ve nasihatları ile yol gösteren değerli babam Talat YÜKSEL'e,

Bugünüm ve her günüm, hayat arkadaşım, can yoldaşım ve gönüldaşım sevgili eşim Funda YÜKSEL'e,

Geleceğim, biricik evladım Arda Emre YÜKSEL'e,

Son olarak henüz kendisi ile tanışmadığım, öpüp koklamak için sabırsızlandığım, bana abisi ile birlikte bir şeyler yapabilme gücü veren, halen anne karnında hayata hazırlanan isimsiz oğlum'a

Teşekkür ederim.

YARATICILIK VE MADDE YAZARLIĞI EĞİTİMİNİN ÖLÇME ARACININ PSİKOMETRİK ÖZELLİKLERİNE ETKİSİ

Kenan Burak YÜKSEL

ÖZ

Bu araştırmanın amacı madde yazarının yaratıcılık düzeyi ve madde yazarlığı eğitiminin, testlerin psikometrik özelliklerine etkisini belirlemektir. Bu maksatla Ankara İl Millî Eğitim Müdürlüğü devlet okullarında 8. sınıf öğrencilerine eğitim veren, seçkisiz belirlenmiş 40 matematik ve 40 Türkçe öğretmenine Torrance Yaratıcı Düşünme Testi uygulanarak yaratıcılık indeksleri belirlenmiştir. İki branşta yaratıcılık indekslerine göre yaratıcılığı yüksek ve düşük gruplar oluşturulmuştur. 2016-2017 eğitim ve öğretim yılı 8. sınıf merkezî ortak asıl sınavının matematik ve Türkçe alt testlerinin alan uzmanları görüşlerine dayalı belirtke tablosu oluşturulmuştur. Belirtke tablosundaki her bir kazanım yaratıcılığı yüksek ve düşük gruptaki öğretmenlere seçkisiz atanmış ve birer madde geliştirmeleri istenmiştir. Sonra iki branşta aynı öğretmen grubundan seçkisiz olarak deney ve kontrol grupları oluşturulmuştur. Deney grubundaki 20'şer öğretmene madde yazarlığı eğitimi verilmiş ve araştırmanın başlangıcındaki kazanımlardan birer madde daha geliştirmeleri istenmiştir. İki alanda geliştirilen maddelerle üçer form oluşturulmuş ve 8. sınıf öğrencilerine (n=1313) uygulanmıştır. Öğrenci yanıtlarından elde edilen madde ve test puanlarından madde ve test psikometrik özellikleri kestirilmiştir. Verilerin analizi neticesinde yaratıcılık ile madde ve testlerin psikometrik özellikleri arasında anlamlı ilişki bulunmazken, madde yazarlığı eğitiminin madde ve testlerin psikometrik özellikleri ile ilişkili ve bu özellikler üzerine etkili olduğu sonucuna ulaşan bulgular elde edilmiştir.

Anahtar sözcükler: Torrance yaratıcı düşünme testi, yaratıcılık, madde yazarlığı eğitimi, madde yazım kuralları, madde ve testin psikometrik özellikleri.

Danışman: Prof.Dr. Nuri DOĞAN, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

THE EFFECT OF CREATIVITY AND EDUCATION IN ITEM WRITING ON MEASUREMENT TOOL'S PSYCHOMETRIC PROPERTIES

Kenan Burak YÜKSEL

ABSTRACT

The aim of this research is to determine the effect of the item writer's creativity and item writing education on the psychometric properties of the items developed, and the tests made up with these items. For this purpose, the creativity indices were determined by applying the Torrance Tests of Creative Thinking to randomly selected 40 mathematics and 40 Turkish lesson teachers who gave education to the 8th grade students in the Ankara Provincial Directorate of National Education public schools. According to creativity indices in each branch, high and low creativity groups were established. Blueprints are defined from 2016-2017 education and training year's 8th grade central common main exam's mathematics and Turkish subtests based on field experts' opinions. Each goal in the blueprints was assigned to teachers from high and low creativity groups, and item for these goals requested. Subsequently, experimental and control groups were established in each branch randomly. Twenty teachers in the experimental groups were given item writing education and they were asked to develop one more item from initial goals. Three test forms were gained with the developed items and applied to 1313 8th grade students. Item and test psychometric properties were estimated from item and test scores obtained from student responses. As a result of analysis of the data, there was no significant relationship between creativity and the item and test psychometric properties, but findings were obtained that item writing education was effective on item and test psychometric properties.

Keywords: Torrance Tests of Creative Thinking, creativity, item writer education, item writing guidelines, psychometric properties of item and test.

Advisor: Prof. Dr. Nuri DOĞAN, Hacettepe University, Department of Educational Sciences, Division of Educational Measurement and Evaluation

İÇİNDEKİLER

KABUL ve ONAY	ii
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI.....	iii
ETİK BEYANNAMESİ	iv
TEŞEKKÜR.....	v
ÖZ.....	vi
ABSTRACT.....	vii
İÇİNDEKİLER	viii
TABLolar DİZİNİ	xi
ŞEKİLLER DİZİNİ	xvi
KISALTMALAR VE SİMGELER DİZİNİ	xvii
1. GİRİŞ.....	1
1.1. Problem Durumu	1
1.2. Araştırmanın Amacı ve Önemi.....	6
1.3. Problem Cümlesi	8
1.3.1. Alt Problemler	8
1.4. Sayıtlar	9
1.5. Sınırlılıklar	10
1.6. Tanımlar	10
1.7. Araştırmanın Kuramsal Temeli	10
1.7.1. Psikometri	11
1.7.1.1. Psikometrik Özellikler	12
1.7.1.1.1. Geçerlik.....	12
1.7.1.1.2. Güvenirlik	13
1.7.1.1.3. Hakkaniyet	13
1.7.1.1.4. Madde ve Test Psikometrik Özellikleri.....	14
1.7.1.1.5. Madde Çeldiriciliği ve İndeksi	19
1.7.2. Test Geliştirme	26
1.7.3. Madde Yazarlığı Eğitimi	32
1.7.4. Madde Yazım Kuralları.....	34
1.7.5. Etki Büyüklüğü.....	38
2. İLGİLİ ARAŞTIRMALAR.....	39
3. YÖNTEM.....	44
3.1. Araştırmanın Yöntemi	44
3.2. Çalışma Grubu	47
3.2.1. Çalışma Grubunun Özellikleri.....	48
3.2.2. Çalışma Grubunu Oluşturan Öğretmenlere İlişkin Demografik Bilgileri.....	48
3.2.2.1. Çalışma Grubunu Oluşturan Öğretmenlerin Cinsiyete Göre Dağılımları.....	48
3.2.2.2. Çalışma Grubunu Oluşturan Öğretmenlerin Yaşa Göre Dağılımları.....	49

3.2.2.3. Çalışma Grubunu Oluşturan Öğretmenlerin Meslekî Kıdeme Göre Dağılımları	49
3.3. Veri Toplama Araçları	50
3.3.1. Torrance Yaratıcı Düşünme Testi.....	50
3.3.2. Merkezî Sınav Eş Formaları	56
3.3.2.1. Belirtke Tablosunun Oluşturulması	57
3.3.2.2. Maddelerin Yazılması ve Eş Formların Elde Edilmesi	58
3.3.3. Madde Uygunluk Formu	59
3.4. Veri Toplama Araçlarının Uygulanışı	59
3.4.1. Torrance Yaratıcı Düşünme Testinin Uygulanışı	60
3.4.2. Merkezî Sınav Eş Formlarının Uygulanışı	60
3.4.3. Madde Uygunluk Formlarının Uygulanışı.....	61
3.5. Verilerin İşlenmesi ve Çözümlemesi	61
3.5.1. TYDT Uygulamasından Elde Edilen Verilerin İşlenmesi ve Analizi	61
3.5.1.1. Matematik Öğretmenleri TYDT Puanlarına Ait Güvenirlik Çalışması	62
3.5.1.2. Türkçe Öğretmenleri TYDT Puanlarına Ait Güvenirlik Çalışması. .	68
3.5.1.3. TYDT Puanlarına Yönelik Yapılan Yapı Geçerliği Çalışması.....	73
3.5.1.4. Yaratıcılık Düzeylerine Göre Gruplandırma	78
3.5.1.5. Deney ve Kontrol Gruplarının Oluşturulması	79
3.5.2. MSEF Uygulamasından Elde Edilen Verilerin İşlenmesi ve Analizi.....	81
3.5.2.1. Madde Psikometrik Özelliklerinin Hesaplanması.....	81
3.5.2.2. Test Psikometrik Özelliklerinin Hesaplanması.....	83
3.5.2.3. Öğrenci MOAS Puanlarının İşlenmesi ve Analizi	84
3.5.3. MUF Uygulamasından Elde Edilen Verilerin İşlenmesi ve Analizi.....	86
3.5.4. Alt Problemlerin Yanıtlanmasına Yönelik Analizler	86
3.6. Araştırmanın İç ve Dış Geçerliği	92
3.6.1. Araştırmanın İç Geçerliği.....	92
3.6.2. Araştırmanın Dış Geçerliği	93
4. BULGULAR VE TARTIŞMA	94
4.1. Yaratıcılık ile Madde ve Test Psikometrik Özellikleri İlişmesine Yönelik Bulgular	94
4.1.1. Yaratıcılık İndeksleri ile Madde Psikometrik Özellikleri	94
4.1.2. Yaratıcılık Düzeyleri ile Test Psikometrik Özellikleri	95
4.1.3. Yaratıcılık Düzeyleri ile Yapı Geçerlikleri	111
4.2. Madde Yazarlığı Eğitiminin Madde ve Test Psikometrik Özelliklerine Etkisine Yönelik Bulgular	122
4.2.1. Eğitimin Madde Psikometrik Özelliklerine Etkisi.....	122
4.2.2. Eğitimin Testin Psikometrik Özelliklerine Etkisi.....	131
4.2.3. Eğitimin Yapı Geçerliğine Etkisi.....	139
4.3. Yaratıcılık ve Madde Yazarlığı Eğitiminin Madde Psikometrik Özelliklerini Yordama Gücüne Yönelik Bulgular	149
4.3.1. Madde Güçlüğü.....	149
4.3.2. Madde Ayırıcılık Gücü	151
5. SONUÇ ve ÖNERİLER	154
5.1. Sonuçlar	154
5.2. Öneriler.....	155
5.2.1. Araştırmaya Dönük Öneriler	155

5.2.2. Uygulamaya Dönük Öneriler	157
KAYNAKÇA	158
EKLER DİZİNİ.....	172
EK 1. ETİK KOMİSYONU ONAY BİLDİRİMİ	173
EK 2. ARAŞTIRMA İZİN BELGESİ.....	177
EK 3. ORJİNALLİK RAPORU.....	174
EK 4. BİLGİLENDİRİLMİŞ GÖNÜLLÜ ONAM FORMU	176
EK 5. TYDT UYGULAMA İZİN BELGESİ	178
EK 6. 2016-2017 EĞİTİM VE ÖĞRETİM YILI 1.DÖNEM 8.SINIF MERKEZİ ORTAK ASIL SINAV MATEMATİK ALT TESTİ A KİTAPÇIĞI ALAN UZMANLARININ MADDE-KAZANIM EŞLEŞTİRMESİ FREKANS VE YÜZDELERİ	179
EK 7. 2016-2017 EĞİTİM VE ÖĞRETİM YILI 1.DÖNEM 8.SINIF MERKEZİ ORTAK ASIL SINAV TÜRKÇE ALT TESTİ A KİTAPÇIĞI ALAN UZMANLARININ MADDE-KAZANIM EŞLEŞTİRMESİ FREKANS VE YÜZDELERİ.....	180
EK 8. ARAŞTIRMADA MATEMATİK ÖĞRETMENLERİNİN GELİŞTİRDİĞİ MADDELERDEN BİR KISMI	181
EK 9. ARAŞTIRMADA TÜRKÇE ÖĞRETMENLERİNİN GELİŞTİRDİĞİ MADDELERDEN BİR KISMI	182
ÖZGEÇMİŞ	183

TABLolar DİZİNİ

Tablo 1.1	: 2011-2016 Yılları Arasında İptal Edilen, Doğru Yanıtı Değiştirilen Madde Sayıları.....	1
Tablo 1.2	: UTM ve WoS Verilerine Dayalı Eğitimde Ölçme ve Değerlendirme Bilim Dalında Bilimsel Eserlerin Konu Başlıkları Frekans ve Yüzdeleri.....	7
Tablo 1.3	: Madde ve Test Psikometrik Özellikleri.....	17
Tablo 1.4	: Örnek Madde Çeldiricilik İndeksi Hesaplaması İçin Seçenek Frekansları.....	25
Tablo 1.5	: Örnek Belirtke Tablosu.....	28
Tablo 1.6	: Madde Geliştirme Formu.....	30
Tablo 1.7	: Madde Yazarlığı Eğitim Programı.....	33
Tablo 1.8	: Çoktan Seçmeli Madde Yazım Kuralları.....	35
Tablo 1.9	: Etki Büyüklüklerinin Yorumlanması.....	38
Tablo 3.1	: Çalışma Grubunu Oluşturan Öğretmenlerin Cinsiyete Göre Dağılımları.....	49
Tablo 3.2	: Çalışma Grubunu Oluşturan Öğretmenlerin Yaş Aralıklarına Göre Dağılımları.....	49
Tablo 3.3	: Çalışma Grubunu Oluşturan Öğretmenlerin Mesleki Kıdem Aralıklarına Göre Dağılımları.....	50
Tablo 3.4	: TYDT Türkçe Uyarılama Çalışma Sonuçları.....	56
Tablo 3.5	: Alan Uzmanlarının Mezuniyet Durumlarına Göre Dağılımları.....	57
Tablo 3.6	: Kappa İstatistiği Yorum Ölçütleri.....	58
Tablo 3.7	: Madde Uygunluk Formu Örneği.....	59
Tablo 3.8	: Geliştirilen Formların Uygulandığı Öğrenci Sayıları.....	60
Tablo 3.9	: Matematik Öğretmenleri TYDT Ham Puanlarına Ait Betimsel İstatistikler.....	62
Tablo 3.10	: Matematik Öğretmenlerine Ait TYDT Puan Dağılımları Normallik Testi Sonuçları.....	64
Tablo 3.11	: Üç Puanlayıcının Matematik Öğretmenlerine Ait Ölçüt Puanları Arasındaki Korelasyon Katsayıları.....	65
Tablo 3.12	: Üç Puanlayıcının Matematik Öğretmenlerine Ait Ölçüt Puanları ANOVA Sonuçları.....	66
Tablo 3.13	: G Çalışması ile Matematik Öğretmenleri Yaratıcılık Ölçmesinin Kestirilen Varyansları ve Toplam Varyansı Açıklama Oranları.....	67
Tablo 3.14	: Türkçe Öğretmenleri TYDT Ham Puanlarına Ait Betimsel İstatistikler.....	68
Tablo 3.15	: Türkçe Öğretmenlerine Ait TYDT Puan Dağılımları Normallik Testi Sonuçları.....	69
Tablo 3.16	: Üç Puanlayıcının Türkçe Öğretmenlerine Ait Ölçüt Puanları Arasındaki Korelasyon Katsayıları.....	70
Tablo 3.17	: Üç Puanlayıcının Türkçe Öğretmenlerine Ait Ölçüt Puanları ANOVA Sonuçları.....	71
Tablo 3.18	: G Çalışması ile Türkçe Öğretmenleri Yaratıcılık Ölçmesinin Kestirilen Varyansları ve Toplam Varyansı Açıklama Oranları.....	72
Tablo 3.19	: KMO Değerlerinin Yorumlanma Ölçütleri.....	73

Tablo 3.20	:	Matematik Öğretmenlerini TYDT Ortalama Ölçüt Puanlarına Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri.....	74
Tablo 3.21	:	Öğretmenlerin Yaratıcılık İndeksleri Betimsel İstatistikleri ve Normallik Testi Sonuçları.....	74
Tablo 3.22	:	Öğretmenlerin Yaratıcılık İndekslerinin Yaratıcılık Düzey Gruplarına Göre Bağımsız Gruplar t Testi Sonuçları.....	78
Tablo 3.23	:	Öğretmenlerin Meslekteki Yıllarının Yaratıcılık Düzey Gruplarına Göre Normallik Testi ve Bağımsız Gruplar t Testi Sonuçları.....	76
Tablo 3.24	:	Deney ve Kontrol Gruplarına Bağımsız Gruplar t Testi Sonuçları.....	80
Tablo 3.25	:	Öğretmenlerin Meslekteki Yıllarının Yaratıcılık Düzey Gruplarına Göre Bağımsız Gruplar t Testi Sonuçları.....	80
Tablo 3.26	:	Matematik Eş Formları Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri.....	82
Tablo 3.27	:	Türkçe Eş Formları Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri.....	83
Tablo 3.28	:	Öğrenci MOAS Puanları Normallik ve Homojenlik Testi Sonuçları.....	85
Tablo 3.29	:	Öğrenci MOAS Puanlarının Uygulanan Eş Formlara Göre ANOVA Sonuçları.....	85
Tablo 3.30	:	Alt Problemlerin Yanıtlanmasında Kullanılan Analizler.....	87
Tablo 4.1	:	Matematik Öğretmenlerine Ait Yaratıcılık İndeksleri ile Madde Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki Korelasyon Analizi Sonuçları.....	94
Tablo 4.2	:	Farklı Yaratıcılık Düzeylerinde Aynı Kazanıma Yönelik Madde Geliştiren Matematik Öğretmenlerinin Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki z Testi Sonuçları.....	95
Tablo 4.3	:	Türkçe Öğretmenlerine Ait Yaratıcılık İndeksleri ile Madde Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki Korelasyon Analizi Sonuçları.....	98
Tablo 4.4	:	Farklı Yaratıcılık Düzeylerinde Aynı Kazanıma Yönelik Madde Geliştiren Türkçe Öğretmenlerinin Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki z Testi Sonuçları.....	99
Tablo 4.5	:	Farklı Yaratıcılık Düzeyindeki Matematik Öğretmenlerinin Geliştirdiği Maddelerle Oluşturulan Testlerin İstatistiklerinin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları.....	102
Tablo 4.6	:	Farklı Yaratıcılık Düzeyindeki Matematik Öğretmenlerinin Geliştirdiği Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları.....	104
Tablo 4.7	:	Öğrencilerin Merkezî Sınav Matematik Alt Testi Puanları ile Eş Formlardan Aldıkları Matematik Puanlarının Korelasyon Analizi Sonuçları.....	105

Tablo 4.8	: Farklı Yaratıcılık Düzeyindeki Türkçe Öğretmenlerinin Geliştirdiği Maddelerle Oluşturulan Testlerin Psikometrik Özelliklerinin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları.....	106
Tablo 4.9	: Farklı Yaratıcılık Düzeyindeki Türkçe Öğretmenlerinin Geliştirdiği Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları.....	108
Tablo 4.10	: Öğrencilerin Merkezî Sınav Matematik Alt Testi Puanları ile Eş Formlardan Aldıkları Türkçe Puanlarının Korelasyon Analizi Sonuçları.....	109
Tablo 4.11	: Matematik Form-I'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri	111
Tablo 4.12	: Matematik Form-II'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri	113
Tablo 4.13	: Matematik Form I ve Form II'e Ait AFA Sonucu Elde Edilen Aynı Kazanımdaki Maddelerin Döndürülmemiş Faktör Yüklerinin z Testi Sonuçları.....	116
Tablo 4.14	: Türkçe Form-I'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri.....	117
Tablo 4.15	: Türkçe Form-II'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri.....	119
Tablo 4.16	: Türkçe Form I ve Form II'e Ait AFA Sonucu Elde Edilen Aynı Kazanımdaki Maddelerin Döndürülmemiş Faktör Yüklerinin z Testi Sonuçları.....	121
Tablo 4.17	: Matematik Öğretmenlerinin Madde Yazarlığı Eğitimi Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki z Testi Sonuçları.....	123
Tablo 4.18	: Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları.....	124
Tablo 4.19	: Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Ayırıcılık Gücü İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları.....	125
Tablo 4.20	: Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Çeldiricilik İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları.....	126
Tablo 4.21	: Türkçe Öğretmenlerinin Madde Yazarlığı Eğitimi Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki z Testi Sonuçları.....	127
Tablo 4.22	: Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları.....	128

Tablo 4.23	: Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Ayırıcılık Gücü İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları.....	129
Tablo 4.24	: Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Çeldiricilik İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları.....	130
Tablo 4.25	: Matematik Öğretmenlerinin Eğitim Sonrası ve Öncesi Geliştirdiği Maddelerle Oluşturulan Testlerin İstatistiklerinin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları.....	131
Tablo 4.26	: Madde Yazarlığı Eğitimi Alan Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdiği Maddelerle Oluşturulan Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları.....	133
Tablo 4.27	: Öğrencilerin Merkezî Sınav Matematik Alt Testi Puanları ile Eş Formlardan Aldıkları Matematik Puanlarının Korelasyon Analizi Sonuçları.....	134
Tablo 4.28	: Türkçe Öğretmenlerinin Eğitim Sonrası ve Öncesi Geliştirdiği Maddelerle Oluşturulan Testlerin İstatistiklerinin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları.....	135
Tablo 4.29	: Madde Yazarlığı Eğitimi Alan Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdiği Maddelerle Oluşturulan Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları.....	137
Tablo 4.30	: Öğrencilerin Merkezî Sınav Matematik Alt Testi Puanları ile Eş Formlardan Aldıkları Türkçe Puanlarının Korelasyon Analizi Sonuçları.....	138
Tablo 4.31	: Matematik Form-III'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri.....	140
Tablo 4.32	: Matematik Form-V'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri.....	142
Tablo 4.33	: Türkçe Form-III'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri.....	144
Tablo 4.34	: Türkçe Form-V'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri.....	146
Tablo 4.35	: Türkçe Form III ve Form V'e Ait AFA Sonucu Elde Edilen Aynı Kazanımdaki Maddelerin Döndürülmemiş Faktör Yüklerinin z Testi Sonuçları.....	148
Tablo 4.36	: Matematik Maddelerinin Güçlüğünün Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları.....	149

Tablo 4.37	:	Türkçe Maddelerinin Güçlüğünün Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları	150
Tablo 4.38	:	Matematik Maddelerinin Ayırıcılık Gücünün Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları.....	151
Tablo 4.39	:	Türkçe Maddelerinin Ayırıcılık Gücünün Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları.....	152
Tablo 5.1	:	Araştırma Bulguları Özeti.....	154

ŞEKİLLER DİZİNİ

Şekil 1.1	:	ÇSM'nin Yapısı.....	16
Şekil 1.2	:	Üç Çeldiricili Maddenin Çeldirici Frekanslarının Üç Boyutlu Düzlemde Gösterimi.....	22
Şekil 3.1	:	Araştırma Süreci.....	46
Şekil 3.2	:	Matematik Öğretmenleri TYDT Ortalama Ölçüt Puanları AFA Sonucu Elde Edilen Yamaç Grafiği.....	75
Şekil 3.3	:	TYDT'nin Faktör Yapısına İlişkin Tanımlanan Tek Faktörlü Ölçme Modeli.....	76
Şekil 4.1	:	Matematik Form-I'e Ait Yamaç Grafiği.....	112
Şekil 4.2	:	Matematik Form-II'e Ait Yamaç Grafiği.....	114
Şekil 4.3	:	Türkçe Form-I'e Ait Yamaç Grafiği.....	118
Şekil 4.4	:	Türkçe Form-II'e Ait Yamaç Grafiği.....	120
Şekil 4.5	:	Matematik Form-III'e Ait Yamaç Grafiği.....	141
Şekil 4.6	:	Matematik Form-V'e Ait Yamaç Grafiği.....	143
Şekil 4.7	:	Türkçe Form-III'e Ait Yamaç Grafiği.....	145
Şekil 4.8	:	Türkçe Form-V'e Ait Yamaç Grafiği.....	147

KISALTMALAR VE SİMGELER DİZİNİ

ACRP	:	The Association of Clinical Research Professionals (Klinik Araştırma Uzmanları Derneği)
AFA	:	Açımlayıcı Faktör Analizi
ÇSM	:	Çoktan Seçmeli Madde
DFA	:	Doğrulayıcı Faktör Analizi
EBSCO	:	Elton B Stephens Company (Bilimsel yayınların bulunduğu veri tabanı)
ERIC	:	Education Resource Information Center (Tez, makale, konferans bildirileri ve araştırma raporlarını barındıran veri tabanı)
ETS	:	Educational Testing Service (Eğitimsel Test Merkezi)
IETLS	:	The International English Language Testing System (Uluslararası İngilizce Dil Test Sistemi)
ITEA	:	International Test and Evaluation Association (Uluslararası Test ve Değerlendirme Kuruluşu)
KGİ	:	Kapsam Geçerlik İndeksi
KGO	:	Kapsam Geçerlik Oranı
KPSS	:	Kamu Personeli Seçme Sınavı
KTK	:	Klasik Test Kuramı
K-S	:	Kolmogorov-Smirnov Normallik Testi
LYS	:	Lisans Yerleştirme Sınavı
MEB	:	Millî Eğitim Bakanlığı
MOS	:	Merkezî Ortak Sınavı
MOAS	:	Merkezî Ortak Asıl Sınavı
MSEF	:	Merkezî Sınav Eş Formu
MUF	:	Madde Uygunluk Formu
ÖSYM	:	Ölçme, Seçme ve Yerleştirme Merkezi Başkanlığı
PAY	:	Paralel Analizler Yöntemi
PMÇK	:	Pearson Momentler Çarpımı Korelasyonu
SBSFK	:	Spearman Brown Sıra Farkları Korelasyonu
S-W	:	Shapiro-Wilk Normallik Testi
TBA	:	Temel Bileşenler Analizi
TUS	:	Tıpta Uzmanlık Sınavı
TYDT	:	Torrance Yaratıcı Düşünme Testi
UTM	:	Ulusal Tez Merkezi
vb.	:	ve benzeri
vd.	:	ve diğerleri
WoS	:	Web of Science (Veri tabanı)
YHAT	:	Yokluk Hipotezi Anlamlılık Testi
YGS	:	Yüksek Öğretime Geçiş Sınavı

p_i	:	Madde Güçlük İndeksi	p	:	İstatistiksel Anlamlılık Değeri
r_{jx}	:	Madde Ayırıcılık Gücü İndeksi	S_x	:	Standart Sapma
ζ	:	Madde Çeldiricilik İndeksi	S_x^2	:	Varyans
ζ_{min}	:	Kabuledilebilir En Düşük ζ	r	:	PMÇK Değeri
t	:	t Testi Değeri	R^2	:	Determinasyon Katsayısı
U	:	Mann Whitney U Testi Değeri	ρ	:	SBSFK Değeri

1. GİRİŞ

Bu bölümde problem durumu, araştırmanın amacı ve önemi, problem cümlesi, alt problemler, sayılılar, sınırlılıklar, tanımlar ve araştırmanın kuramsal temellerine yer verilmiştir.

1.1. Problem Durumu

Türkiye’de, devlet kadrolarında istihdam edilen nüfusun 290 bin 919’u, 18/3/2002 tarih ve 2002/3975 sayılı Bakanlar Kurulu Kararı ile 2002 yılından beri yürürlükte olan KPSS sonuçları ile 13 milyon 17 bin 959 kişi arasından seçilerek işe alınmıştır (DPB, 2017). KPSS dışında bir seçme yöntemi ile işe kabul edilmiş kamu personeli ile özel sektörde iş hayatını sürdüren bireyler de ÖSYM ve/veya MEB tarafından uygulanan sınavlardan (test) aldıkları puanlar karşılığında diploma, mezuniyet ve işe kabul gibi belgeler almaktadırlar.

Sonucu bu denli mühim olan testlerin amaçlanan işlevlerini yerine getirebilmesi yani, amaçlanan özelliği ölçebilmesi, bunun için de bir takım özelliklere (psikometrik) sahip olması beklenir. Fakat son yıllarda ülkemizde geniş ölçekli testlerde, birçok madde iptali veya doğru cevap şıkkı (doğru yanıt) değiştirme durumları ile karşılaşmıştır (Tablo 1.1). Bu da ölçme sistemine güvensizlik ve hak kayıplarına neden olabilmekte, organizasyonların (kurum, şirket vb.) belirli görevleri yapabilecek uygun özelliklere sahip bireylere ulaşmasına engel teşkil edebilmektedir.

Tablo 1.1: 2011-2016 Yılları Arasında İptal Edilen, Doğru Yanıtı Değiştirilen Madde Sayıları*

<i>Sınav Türü</i>	<i>Sınav Yılı</i>	<i>İptal Edilen Madde Sayısı</i>	<i>Cevap Şıkkı Değiştirilen Madde Sayısı</i>
LYS	2016	2	2
YGS	2016	2	1
LYS	2015	3	
YGS	2015	1	
KPSS	2015	12	3
KPSS	2014	1	
KPSS	2013	7	
TUS	2013	11	
KPSS	2012	2	
KPSS	2011	3	

* Kamuya açık kaynaklardan derlenmiştir.

Kamuoyunda genel olarak sınav, anket, test, soru, ölçek gibi isimlerle adlandırılan ölçme araçları belirli basamaklar izlenerek meydana getirilmektedir. Bu işlemler bütünü test geliştirmedir (Baykul, 2015, s.269). Test geliştirme sürecinin tüm işlemlerinin doğru yapıldığı koşullarda, maddenin iptalinin ya da doğru yanıtın değişmesinin kaynağının madde yazarı olduğu açıktır. Madde yazarının, ölçülmek istenen yapıyı ortaya çıkaracak uygun maddeyi geliştirememesi, süreçteki eksiklik ve hatalardan kaynaklanabileceği gibi, madde yazarının yetersizliğinden de ortaya çıkabilmektedir. Deneme uygulaması, ölçme uzmanlarının maddeleri kontrolü (Baykul, 2015, s.318; Haladyna & Rodriguez, 2013, s:26; Tarrant, Knierim, Hayes & Ware, 2006) gibi tedbirlerle Tablo 1.1'deki gibi durumların önüne geçilebilir. Bunlar ve benzeri önlemler alınsa dahi, madde, yaratıcısı olan yazardan bağımsız düşünülemez.

Basit bir akıl yürütmeye; yetersiz madde yazarı kötü madde yazabilecek; kötü maddeyi cevaplayacakların muhtemelen akli karışacak; akli karışmış cevaplayıcılardan istedik düzeyde olanlar başarısız, istenmedik düzeyde olanlar da başarılı görülebilecek; başarısız kişiler, başarılı kişilerle yürütülmesi gereken işlere kabul edilebilecek ve sonunda iş yetersiz bireylere emanet edilebilecektir. "Bilen, soru kötü olsa da cevaplar" iddiasını çürütebilecek birçok araştırma mevcuttur (Downing, 2005; Pais vd., 2016; Tarrant, Knierim, Hayes & Ware, 2006; Tarrant & Ware, 2008). Söz konusu araştırmalar kusurlu maddelerden oluşan testlerin güvenilirlik ve geçerliğinin, kusurlu olmayan maddelere sahip testlere göre daha düşük düzeyde olduğu bulgusuna ulaşmışlardır. Bireylere ait başarı ve yetenek gibi özelliklerin yanısıra, sonucunda müebbet hapis cezasına varan kararların alınmasına etki eden cezaî ehliyetin tespitinde bile psikolojik testler kullanılabilir. Bu nedenle belirli niteliklere sahip olması beklenen test maddelerini geliştirecek olan madde yazarlarının da belirli nitelikleri taşıması gerekir.

Nasıl ki iyi bir roman yazmanın veya güzel bir resim çizmenin bir formülü yoksa uygun maddeler geliştirmenin de bir tarifi bulunmamaktadır. Ebel (1951, s.185) madde yazmanın bir sanat olduğunu, özel yeteneklerin nadir bir birleşimine ihtiyaç duyduğunu, ancak geniş ve eleştirel denetimli uygulamayla gelişebileceğini, ustalık gerektirdiğini belirtmiş; madde yazmanın esasen yaratıcı bir işlem olduğunu vurgulamıştır. İyi madde yazabilmek için madde yazarı;

1. Üzerine madde yazdığı konu alanına tam anlamıyla hâkim olmalı,
2. Akılcı ve iyi gelişmiş eğitimsel değerlere sahip olmalı,
3. Testin uygulanacağı grubun psikolojik yapılarının ve eğitimsel seviyelerinin ne düzeyde olduğunu bilmeli,
4. Sözel iletişimde uzman olmalı,
5. Madde yazma teknikleri konusunda uzman olmalıdır (Ebel, 1951, ss.186-188).

Ebel'e (1951), yaratıcılığın, madde yazmada belirlenmiş ilkeleri esnekçe yorumlayabilmek için ihtiyaç olduğunu belirten Wesman da (1971) katılmıştır. Psikometrinin önde gelen isimleri yaratıcılık gibi doğuştan gelen özelliklere bel bağlamak yerine madde yazmayı bilimsel altyapılı bir sistematığe dönüştürmeye uğraşmışlardır (Haladyna, 2004:127; Haladyna & Downing, 1989a; 1989b; Haladyna & Shindoll, 1989; Irvine & Kyllonen, 2002). Bu uğraşılara rağmen, madde yazma konusunda bir takım gelişmeler sağlansa da, madde yazmanın esas olarak yaratıcılık gerektirdiği; madde yazma kuralları, madde üretim sistematikleri, madde kabukları (item shells) gibi yöntemlerin kısmen veya tamamen tereddütlü oldukları düşünülmektedir (Haladyna, Downing & Rodriguez, 2002). Çoktan seçmeli madde (ÇSM) yazmanın bilimden ziyade bir sanat olarak kabul görülmesi tezini (Ebel, 1951, s.185; Haladyna, 2004, s.ix; Hambleton & Eignor, 1979; Wood, 1977) çürütecek güvenilir ve geçerli kanıtlar bugüne değin elde edilememiştir.

Madde yazarlığı üzerine yapılan bazı araştırmalar, tecrübeli madde yazarlarının geliştirdikleri maddelerin psikometrik özelliklerinin, tecrübesiz yazarlarınkine nazaran daha iyi düzeyde olduğu bulgusuna ulaşmışlardır (Fulkerson, Nichols & Mittelholtz, 2010; Rodriguez, 1997; Salisbury, 2005). Madde yazarlığı için gerekli diğer bir özellik de madde yazma konusunda eğitilmiş olmasıdır (Downing & Haladyna, 1997; Iramaneerat, 2012; Jozefowicz vd., 2002).

Baykul'a (2015, s.292) göre madde yazarı;

1. Hedef, davranış, davranışların sınıflandırılması, aşamalı sınıflamanın ve bu sınıflamada her sınıfın özellikleri hususunda yetişmiş,
2. Çoktan seçmeli test, ÇSM ve özellikleri konusunda yetişmiş,
3. Madde yazmada dikkate alınacak hususlarda yetişmiş,
4. Maddenin yazılacağı dile hâkim,
5. Özelliği ölçülecek grubun yaptığı yanlışlara aşina olmalıdır.

Şu ana kadar değinilen alan uzmanları görüşleri ve araştırma sonuçları birlikte düşünüldüğünde; madde yazarında bulunması gereken nitelikler 6 ana başlıkta toplanabilir.

1. Alana hâkimiyet
2. Hedef grubu tanıma
3. Dile hâkimiyet
4. Madde yazarlığında tecrübe
5. Yaratıcılık
6. Madde yazarlığı eğitimi

Başlıklar incelendiğinde ilk 4 maddenin kısa sürede kazanılabilecek özellikler olmadığı, 5'incinin ise eğitimle geliştirilebilse de doğuştan sahip olunan bir özellik olduğu görülmektedir (Leikin & Sriraman, 2016, s.1). 6 özellik içerisinde kısa sürede kazanılabilecek tek özelliğin madde yazarlığında eğitimli olmak olduğu söylenebilir. Söz konusu 6 özelliğin birlikte, geliştirilen maddeleri ne derecede açıkladığının ve eğer açıklamakta yetersiz iseler, varsa başka hangi değişkenlerin madde psikometrik özelliklerine etki ettiğinin tespit edilmesi; arzulanan düzeyde madde geliştirebilecek madde yazarlarının tespiti ve yetiştirilmesi için gereklidir denebilir.

Aksiyomatik olarak madde yazarının yaratıcı olması gerektiği benimsense de, madde yazarının yaratıcılık düzeyi ile geliştirdiği maddenin psikometrik özellikleri arasındaki ilişkiyi sınavan araştırmalar da bulunmamaktadır. İddia edildiği gibi madde yazma yaratıcılık gerektiren bir eylem midir? Eğer öyle ise madde yazarlığında ne derecede etkilidir? Yoksa yalnızca üst düzey bilişsel süreçleri

yoklamak için mi gereklidir? sorularına alanın tatminkâr yanıtlar veremediği görülmüştür.

Öğrencileri, belirli bir alanda yetiştiren öğretmenlerin, 6 başlıkta toparlanan madde yazarlığı özelliklerinin ilk 3'üne farklı düzeylerde olsa da, sahip oldukları söylenebilir. Öğretmenlerin yeterlilikleri arasında bulunan (MEB, 2008) mesleki gelişim ile program ve içerik bilgisi birinci başlık; öğrenciyi tanıma ile de ikinci başlık için uygun oldukları düşünülmüştür. Ayrıca öğrenmeyi, gelişimi izleme ve değerlendirme yeterlilikleri, sınıf içi uygulamalarda öğrencilere farklı ölçme araçları kullanarak ölçme yapmaları ve sıklıkla da ÇSM'li testleri kullanmaları (Anıl ve Acar, 2008; Çakan, 2004) sebebiyle 4'üncü özelliğe de sahip oldukları değerlendirilmiştir. Fakat her bir öğretmenin sahip oldukları yaratıcılık düzeylerinin farklı olması doğaldır. Ayrıca öğretmenlerin bir kısmının mezun oldukları üniversitelerce ve göreve başladıktan sonra MEB tarafından ölçme ve değerlendirmeye dönük eğitimler verilse de yeteri kadar nitelikli olmadığına dair kimi araştırmalar da bulunmaktadır (Birgin ve Gürbüz, 2008; Erdemir, 2007; Erdoğan ve Kurt, 2012; Okur, 2008). Güvenilir ve geçerli ölçme sonuçları veren araçların geliştirilebilmesi için bu araçları geliştiren madde yazarlarının yaratıcı olup olmamaları ve madde yazarlığı konusunda eğitim sahibi olup olmamalarının belirlenmesinin yararlı olacağı değerlendirilmektedir.

Öğrencileri ileri eğitim kademelerine ve gelecek yaşantılarına hazırlamak için oluşturulmuş eğitim programları içerisinde matematik ve Türkçe derslerinin diğer derslere nazaran daha büyük öneme sahip olduğu söylenebilir. Çünkü bu iki ders hayat boyu bireylerin iş hayatlarını ve günlük yaşamlarını doğrudan etkileyebilmektedir. Alışverişte bir şey alırken ödenmesi gereken paranın hesabından, saatin kaç olduğuna ve hatta karşılaşılan problemleri algılama, bu sorunlara çözüm bulma, göze alınacak riskler ya da olasılıkların hesabı da yaşam içindeki matematiktir (NCTM, 2000, s.73). Yazılı bir metni okuyabilmek ve anlayabilmek, metinle öne sürülen bilgiye sahip olabilmenin ön şartlarıdır. Bu nedenle ülkemizde uygulanan zorunlu eğitimin resmî dili olan Türkçe'nin sözlü iletişim, okuma ve yazma becerilerini kazandırmayı amaçlayan Türkçe dersinin yeri ayrıdır (MEB, 2015a). Güleç ve Alkış'ın (2003) yaptıkları araştırma sonucu elde ettikleri bulgular; Türkçe dersinin matematik, hayat bilgisi, sosyal bilgiler ve fen bilgisi dersleri ile yakından ilişkili olduğuna işaret etmektedir ($r=0,83$). Eğitim

sürecinde önemleri büyük olan bu iki dersin kazanımlarına öğrencilerin ne düzeyde sahip olduklarını ölçmek de değerli olacaktır. Bu nedenle, çalışma kapsamında, sayısal yetenek ve okuduğunu anlama gibi temel eğitim alanları olan matematik ve Türkçe dersleri öğretmenlerinin, yaratıcılık düzeyleri ölçülerek ve aynı öğretmenlerin bir grubuna madde yazarlığı eğitimi verilerek, geliştirdikleri maddelerin ve bu maddelerle oluşturulan testlerin psikometrik özelliklerine yansımaları araştırılmıştır.

1.2. Araştırmanın Amacı ve Önemi

Bu araştırmanın genel amacı; madde yazarında bulunması gerektiği iddia edilen özelliklerden yaratıcılık ve madde yazarlığında eğitimlilik değişkenlerinin, bu özelliklere farklı düzeylerde sahip olan madde yazarı öğretmenlerin geliştirdiği maddeler ve bu maddelerle oluşturulan testlerin psikometrik özelliklerine etkisini belirlemektir.

Testler bireyler hakkında önemli kararların alınmasında kullanılmak üzere geliştirilmektedir. Testler ve temel yapı taşları olan maddeler, alınacak kararların isabetli olabilmesi için, ölçmeyi amaçladıkları özellikleri ölçebilecek nitelikte olmalıdırlar. Maddelerin insana ait bilgi, tutum ve becerileri; özelliği ölçülecek bireylerin bilişsel, duyuşsal ve devinişsel özelliklerini ölçebilmesi gerekir. Farklı özellikleri farklı düzeylerde ve uygun şekilde ölçebilecek maddeler ancak belirli niteliklere sahip madde yazarlarınca geliştirilebilir. Bu özelliklerden yaratıcılık ve madde yazarlığı eğitiminin madde ve testin psikometrik özelliklerine etki edip etmediğinin belirlenmesi, başta ÖSYM ve MEB gibi geniş ölçekli test uygulamaları yapan kurumlarla, test geliştirme ve uygulaması yapan şirketlere, madde yazarlığı konusunda çalışma yapan araştırmacılara, ölçme ve değerlendirme alanında eğitim veren bölümlere madde yazarı seçimi ve yetiştirilmesinde katkı sağlayıcı bulgular sunabilecektir. Araştırma bulguları, sonucu madde iptali veya doğru yanıtın değiştirilmesine varan durumlarla karşılaşmamak için alınacak önlemlerin belirlenmesinde kullanılacaktır. Ayrıca geniş ölçekli testler için geliştirilen bir maddenin ortalama bedelinin 1800 ila 2000 \$ olduğu (Ebel, 1951, s.186; Koon & Kamata, 2013) düşünülüğünde madde geliştirme üzerine araştırma yapmanın ekonomik değeri de ortadadır.

Google Akademik (17.02.2017), EBSCO Host (03.03.2017), ERIC (25.03.2017), ProQuest Digital Dissertations (06.04.2017) veritabanlarında; “item writer”, “item writing”, “item psychometry” ve Ulusal Tez Merkezi (UTM) (21.03.2017) veritabanında “madde yazarı”, “madde psikometrik özellikleri”, “soru yazarı” sözcük grupları ile yapılan taramalar sonucunda; bir bütün olarak madde yazarında bulunması gereken tüm özellikleri açıklamaya yönelik yapılmış bir araştırmaya rastlanılmamıştır. Özellikle Türkiye özelinde madde yazarının sahip olduğu özelliğin, geliştirdiği maddeleri nasıl etkilediğine yönelik deneysel araştırma eksikliği gözlemlenmiştir. Eksikliği görülen bu durumu açıklamaya dönük bir araştırmanın, ölçme ve değerlendirme alanına katkı sağlayacağı düşünülmüştür.

UTM çevrimiçi sayfasında eğitimde ölçme ve değerlendirme bilim dalı kapsamına giren yüksek lisans ve doktora tezleri (n=227) tarandığında (21.03.2017 tarihi itibari ile) üzerine çalışılan konu başlıkları ile Ukşul (2016) tarafından yapılan bibliyometrik yöntem çalışmasında 2006-2015 yılları arasında eğitimde ölçme ve değerlendirme alanında yayımlanmış WoS veri tabanında taranan 205 makale ve 51 tam metin bildirinin konu başlıkları frekans ve yüzdeleri Tablo 1.2’de sunulmuştur.

Tablo 1.2: UTM ve WoS Verilerine Dayalı Eğitimde Ölçme ve Değerlendirme Bilim Dalında Bilimsel Eserlerin Konu Başlıkları Frekans ve Yüzdeleri

<i>Konu Başlığı</i>	<i>Y.Lisans Tezi</i>		<i>Doktora Tezi</i>		<i>Bilimsel Yayın</i>		<i>Toplam</i>	
	f	%	f	%	f	%	f	%
Ölçme Aracının Yapısal Nitelikleri	45	27,3	8	12,9	68	26,6	121	25,1
İstatistiksel Yöntem	27	16,4	22	35,5	70	27,3	119	24,6
Değerlendirme	56	33,9	5	8,1	51	19,9	112	23,2
Kuram (KTK, MTK, Genellenebilirlik)	11	6,7	14	22,6	24	9,4	49	10,1
Değişen Madde Fonksiyonu	6	3,6	6	9,7	17	6,6	29	6,0
Test Eşitleme	8	4,8	5	8,1	6	2,3	19	3,9
Test/Ölçek Geliştirme	5	3,0	1	1,6	11	4,3	17	3,5
Test/Ölçek Uyarlama	1	0,6	0	0,0	9	3,5	10	2,1
Standart Belirleme	5	3,0	1	1,6	0	0,0	6	1,2
Şans Başarısı	1	0,6	0	0,0	0	0,0	1	0,2
Toplam	165	100,0	62	100,0	256	100,0	483	100,0

Tablo 1.2 incelendiğinde üzerine en çok araştırma yapılan konunun “Ölçme Aracının Yapısal Nitelikleri” olduğu görülmektedir. Bunu “İstatistiksel Yöntemler” ve ölçme sonuçlarını belirli ölçütlerle kıyaslayan “Değerlendirme” takip etmektedir.

Türkiye’de yapılan bu çalışmalarda test/ölçek geliştirmenin oransal olarak az olduğu görülmektedir. Bu araştırmalar içinde de amacı test ve ölçeklere madde geliştiren madde yazarını inceleyen ise bulunmamaktadır. Bu bulgu Şenyurt ve Özkan (2017)’in bulguları ile de uyumaktadır. Alanın ölçme aracının psikometrik özelliklerine yoğunlaşması doğal olmakla birlikte buzdağının görünmeyen kısmındaki madde yazarının da araştırma konusu olması gerekliliği yadsınamaz bir gerçektir. Bu sebeple araştırmanın alanyazındaki bu eksikliği kapatmaya yönelik bir katkı sağlayacağı da değerlendirilmektedir.

Bu araştırma ile, öğrencilerin eğitim süreçleri içerisinde ve sonunda hedeflenen başarı düzeylerine ulaşip ulaşmadıklarını belirlemede kullanılan maddeleri geliştirecek olan madde yazarlarının sahip olmaları gereken özelliklerden yaratıcılık ve almaları gereken madde yazarlığı eğitiminin, maddelerin psikometrik özelliklerine etkisini belirleyen bulgular elde edilebilecektir.

1.3. Problem Cümlesi:

Madde yazarının, yaratıcılık ve madde yazarlığı eğitim durumu, geliştirdiği ÇSM ve bu maddelerde oluşturulan testin psikometrik özelliklerine ne derecede etki etmektedir?

1.3.1. Alt Problemler:

1. Öğretmenlerin yaratıcılık düzeyleri ile geliştirdikleri ÇSM ve testlerin psikometrik özellikleri arasında anlamlı bir ilişki var mıdır?

a. Matematik ve Türkçe öğretmenlerine ait yaratıcılık indeksleri ile geliştirdikleri maddelerin güçlük, ayırıcılık gücü, çeldiricilik indeksleri arasında anlamlı bir ilişki var mıdır?

b. Matematik ve Türkçe öğretmenlerinin geliştirdikleri maddelerle oluşturulan testlerin güvenilirlikleri, varyansları, aritmetik ortalamaları, kapsam geçerlikleri, uyum geçerlikleri yaratıcılık düzeylerine göre anlamlı bir şekilde farklılaşmakta mıdır?

c. Yaratıcılığı yüksek ve düşük matematik ve Türkçe öğretmenlerinin geliştirdikleri maddelerle oluşturulan testlerin faktör analizi sonuçları nasıldır?

2. Madde yazarlığı eğitimi alan öğretmenlerin, eğitim öncesi geliştirdikleri ÇSM ve testlerin psikometrik özellikleri, eğitim sonrası anlamlı bir şekilde farklılaşmakta mıdır?

a. Matematik ve Türkçe öğretmenlerinin geliştirdikleri maddelerin güçlük, ayıricılık gücü, çeldiricilikleri eğitim sonrası ve öncesine göre anlamlı bir şekilde farklılaşmakta mıdır?

b. Matematik ve Türkçe öğretmenlerinin geliştirdikleri maddelerle oluşturulan testlerin güvenilirlikleri, varyansları, aritmetik ortalamaları, kapsam geçerlikleri, uyum geçerlikleri eğitim sonrası ve öncesine göre anlamlı bir şekilde farklılaşmakta mıdır?

c. Madde yazarlığı eğitimi alan matematik ve Türkçe öğretmenlerinin eğitim öncesi ve sonrası geliştirdikleri maddelerle oluşturulan testlerin faktör analizi sonuçları nasıldır?

3. Yaratıcılık ve madde yazarlığı eğitimi ÇSM psikometrik özelliklerinin anlamlı birer yordayıcısı mıdır?

a. Matematik ve Türkçe öğretmenlerinin yaratıcılık indeksleri ve madde yazarlığı eğitim durumları geliştirdikleri maddelerin güçlük indekslerinin anlamlı birer yordayıcısı mıdır?

b. Matematik ve Türkçe öğretmenlerinin yaratıcılık indeksleri ve madde yazarlığı eğitim durumları geliştirdikleri maddelerin ayıricılık gücü indekslerinin anlamlı birer yordayıcısı mıdır?

1.4. Sayılılar:

Bu araştırma kapsamında;

a. Görüşlerine başvuru alan uzmanlarının Madde Uygunluk Formlarını (MUF),

b. Yaratıcılık düzeyleri ölçülen öğretmenlerin Torrance Yaratıcı Düşünme Testini (TYDT),

c. Madde yazarlığı eğitimi alan öğretmenlerin madde yazarlığı eğitimini,

d. TYDT'ni puanlayan uzmanların puanlamayı,

e. Geliştirilen eş formları yanıtlayan öğrencilerin test maddelerini, doğru anlayıp, içtenlikle ve tam performansla yanıtladıkları ve değerlendirdikleri varsayılmaktadır. Ayrıca madde yazarı öğretmenlerin Türkçe dilinde yazma becerilerinin ve alanlarına hâkimiyet düzeylerinin benzer seviyede olduğu varsayılmıştır.

1.5. Sınırlılıklar:

Bu araştırma KTK, matematik ile Türkçe ders alanları ve ÇSM'lerle sınırlıdır.

1.6. Tanımlar:

- Madde Geliştirme** : Belirli ölçütlerle, özelliklerin bireylerde bulunma derecesini ortaya çıkaracak ölçme aracını oluşturma işlemidir.
- Madde Yazarı** : Ölçme aracı olan maddeyi geliştirerek ortaya koyan uzmandır.
- Madde Yazarlığı Eğitimi** : Madde yazarına, uygun nitelikte madde geliştirebilme becerisini kazandırmayı hedefleyen süreçtir.

1.7. Araştırmanın Kuramsal Temeli

Özelliklerin var olup olmadığı ve eğer var ise ne derecede bulunduğu ölçme ve değerlendirmenin çıktıları ile ispatlanabilmektedir. Sağlıklı kararlar, bilim dallarına geçerli ve güvenilir bilgiler sunan ölçme ve değerlendirme ile alınabilir. Karar, kuram, genelleme, teori ve kanunlarla işlevini yerine getiren bilim de ölçmenin genel geçer verilerine ihtiyaç duymaktadır. Erkuş (2003, s.4), “genel anlamda denilebilir ki ölçmesiz bilim olmaz” demekle ölçmenin bilim için ne kadar hayati olduğunu vurgulamıştır. Stevens (1946), ölçmeyi, belirli kurallara göre nesnelere ve olaylara sayı atanması olarak tanımlamış; Lord & Novick (1968) ölçmenin, nesnelere değil de, nesnelere özelliklerine uygulandığını belirterek, tanımı daha doğru bir hale taşımıştır (s.13). Turgut (1977, s.11) “bir niteliğin gözlenip gözlem sonuçlarının sayılarla veya sembollerle gösterilmesidir” diyerek tanımı operasyonel şekilde belirginleştirmiştir. Farklı tanımlarından istifade ile ölçme, *niteliklerin nicelik evrenlerinde buldukları konumlarını tespit etmektir* şeklinde tanımlanabilir.

Ölçme işlemiyle elde edilen ölçme sonuçlarının anlamlandırılması için değerlendirme yapılmalıdır. Değerlendirme, ölçme sonuçlarının bir ölçütle kıyaslanarak bir karara varılması işi olarak tanımlanmaktadır (Turgut, 1977, s.225).

Değerlendirme, ölçme işlemi ile elde edilen *ölçme sonuçları*, beklenen seviyeyi temsil eden *ölçüt* ve ölçme sonuçları ile ölçütü kıyaslayarak varılan *karar* öğelerinden meydana gelmektedir. Alınacak kararlar uygun şekilde belirlenmiş ölçüt ve bu ölçütle kıyaslanacak güvenilir ve geçerli ölçme sonuçlarını gerektirir. Ölçme sonuçları ise psikolojik test olarak adlandırılan, insan özelliklerini ölçmek için geliştirilmiş araçlarla elde edilir.

Psikolojik test; bilişsel, duyuşsal veya kişiler arası özelliklerle ilgili davranış örnekleri elde etmek ve bu örnekleri standartlara göre puanlamak ve değerlendirmek için sistematik bir süreçtir (Urbina, 2014, s.1). Psikolojik test, bir davranış örneklemdir, davranış örneklemini sistematik yolla toplanmalıdır, amacı iki veya daha fazla bireye ait davranışın karşılaştırılmasıdır. Farklı ihtiyaçlara göre değişik alanlarda başarı testleri, yetenek testleri, zekâ testleri, kişilik testleri, tutum ölçekleri gibi birçok psikolojik test çeşidi bulunmaktadır. Ölçülmeye çalışılan özellikler fiziksel oldukları sürece, ölçme işlemi kesin veya kesine yakın sonuçlar üretebilirken, psikolojik özellikler söz konusu olduğunda kesinlikten uzaklaşan sonuçlara ulaşılmaktadır. Bu uzaklığın belirlenmesi ölçme aracının psikometrik özelliklerinin kestirilmesiyle mümkündür.

1.7.1. Psikometri

Psikometri, Yunanca ruh anlamında psyche ($\psi\chi\eta$) ile ölçme anlamında metron ($\mu\acute{\epsilon}\tau\rho\nu$) kelimelerinin bir araya gelmesi ile oluşmuştur. İlk kez Galton (1879) tarafından “zihinsel işlemlere ölçü ve sayı koyma sanatı” olarak tanımlanan psikometrinin, fizyoloji profesörü Buchanan (1885) tarafından parapsikolojideki uygulamalarla alanı genişletilmiştir. Spence (2003) terimin parapsikolojideki kullanımını; bir kişiye ait eşyaya dokunarak, o kişinin fiziksel, zihinsel ve ahlaki özelliklerinin belirlenmesi olarak açıklamıştır. Parapsikolojideki kullanımı sürmekle birlikte, psikometri, asıl olarak psikoloji biliminde kullanılarak gelişmiştir. Stuart-Hamilton (2007, s.214) “psikolojik özelliklerin ve becerilerin ölçülmesidir” şeklinde tanımladıkları terimin, genellikle, bireysel farklılıkların eşanlamlılığı olarak görüldüğünü belirtmişlerdir. Psikometrik ise insan özelliklerinin ölçülmesinde kullanılan ölçme araçlarının özellikleri ile ilgilenen alandır. Psikometriğin ilgi alanına ölçme araçlarının üç özelliği girmektedir. Birincisi, psikolojik testlerle elde edilen bilgi çeşidi (çoğunlukla test puanları); ikincisi psikolojik testlerden elde

edilen verilerin güvenilirliği; üçüncüsü ise psikolojik testlerden elde edilen verilerin geçerliği ile ilgili konulardır (Furr & Bacharach, 2013, s.7).

1.7.1.1. Psikometrik Özellikler

Psikometrik özellikler madde/test puanları, puanların güvenilirliği ve geçerliğine dair yapılan analizlerle ölçme aracının ölçme gücüne kanıt sunan ve betimleyen istatistiksel kavramlardır. AERA, APA & NCME (2014, s.11), eğitim ve psikolojideki testler için standartlar kitabında testlere ait geçerlik, güvenilirlik/kesinlik ve ölçmenin hatası, test etmede hakkaniyet olarak üç esastan söz etmektedir. Test psikometrik özellikleri olan bu kavramların, var/yok, az/çok gibi nitelendirilmesinden daha çok; elde edilen verilerle haklarında kanıt toplanarak yeterlilikleri ile ilgili yorum yapılır.

1.7.1.1.1. Geçerlik

Geçerlik, ölçme aracının, ölçmek üzere geliştirildiği özelliği ne derecede ölçebildiği ile ilgili değerlendirmedir (Urbina, 2014, s.168). Ölçmenin doğal amacı da özelliği mümkün en kesin şekilde kestirmektir. Bu sebeple geçerlik, psikometrik özelliklerden en önemli olanıdır. Tüm diğer özelliklere sahip bir ölçme aracı, geçerlik sınavında başarılı olamıyorsa; ilgili özelliği istedik düzeyde ölçebildiği iddia edilemez. Bir ölçme aracının geçerlik ya da AERA vd.'nin (2014) belirttiği gibi geçerliğine sunulacak kanıt arama çalışması, elde edilen puanların kıyaslanabileceği uygun ölçüte gereksinim duyar (s.12). Ölçülmek istenen özelliğin karşılığı değer olarak bir ölçüt değeri belirlenir ve bu değere nazaran doğru bir kestirimde bulunan ölçme aracının geçerli olduğu kabul edilir (Kane, 2001). Literatürde geçerlik çeşitleri olarak adlandırılan fakat doğru bir ifade ile geçerlik kanıtı kaynakları olan; kapsam, bir ölçüte dayalı ve yapı kaynaklı geçerlik kanıtı türleri olduğu söylenebilir.

Bir testin, bu testle ölçülmek istenen davranışları ne derecede kapsadığına dair sunulacak kanıtlar *Kapsam Geçerliğine* sunulacak kanıtlardır. *Ölçüt Dayanaklı Geçerlik* kanıtı ise bir testten elde edilen puanların, bu puanlara yordayıcı ya da uyumlu kabul edilen ölçütle olan ilişkilendirmedir. İnsanlarda var olduğu iddia edilen ve hipotetik kavramlar olan yapının ölçme aracı ile ortaya konulma derecesi de *Yapı Dayanaklı Geçerlik* kanıtıdır.

1.7.1.1.2. Güvenirlik

Güvenirlik ölçme araçlarından elde edilen puanların ölçüm hatalarından arınlık derecesini belirtir. Güvenirlik kavramı; birbirinin ardına gelen tekrarlı ölçmelerin aynı sonucu vermesi olarak tanımlanır (Wainer & Dorans, 2000, s.159).

Güvenirlik tahmini (kestirimi) ölçmenin kesinliği ve kesinlikten uzaklığı (ölçmenin hatası) ile ilgilidir. Güvenirlik, test maddelerinden elde edilen en az iki puanın tutarlığı veya benzerlik derecesi ile zamana karşı değişmezliğidir. Güvenirlik kestiriminin farklı yöntemleri, gerçek ve hata puanlarıyla ilgili özel varsayımlara dayanır ve bu nedenle farklı hata kaynakları belirtir. Varsayımlarla kabul ettikleri hata kaynaklarına göre KTK (Gerçek Puan Kuramı), Genellenebilirlik Kuramı ve Madde Tepki Kuramı (Örtük Özellikler Kuramı) olmak üzere psikometride üç temel kuram bulunmaktadır. Alanyazında genellenebilirlik kuramı, hata kaynaklarını bir arada ele almakla birlikte, KTK'nın bir uzantısı olarak kabul görmektedir (Brennan, 2001, s.2). Güvenirlik KTK'da gerçek ve hata puanları üzerine temellendirilmiştir (Crocker & Algina, 2008, s.122). KTK'ye göre gözlenen (testle elde edilmiş) puan, hata puanı ve gerçek puanın birleşiminden oluşmaktadır. Puanlara dayalı elde edilen güvenilirlik değeri (indeksi) ise gerçek puanların standart sapmasının, gözlenen puanların standart sapmasına oranıdır (Gulliksen, 1950, s.34). Gerçek puanlar bilinmeyeceğinden dolayı standart sapması da hesaplanamaz, bu sebeple güvenilirlik indeksi teoriktir. Güvenirlik kestirimi, değişkenliğin, hatalı ölçmeler ve gerçek puanların hangisinden ve ne miktarda kaynaklandığını ortaya koymaktadır (Levy & Mislevy, 2016, s.154). Testlerin güvenilirliğinin kestirilebilmesi için; test-tekrar test; eşdeğer formlar; eşdeğer yarılar; iç tutarlık yöntemleri kullanılmaktadır.

1.7.1.1.3. Hakkaniyet (Fairness)

Hakkaniyet, ölçme aracının, bireylerin asıl ölçülmek istenen özelliğini, ölçme dışındaki özellikleri ile karıştırmadan ölçebilmesidir (AERA vd., 2014, s.49). Görme engelli bir öğrenci, görme kabiliyeti ile ilgisi bulunmamasına rağmen diğer görme sorunu olmayan öğrencilerle aynı koşullarda teste alındığında, maddeleri göremeyeceğinden haliyle de tepkide bulunamayacaktır. Bu şekilde düşünüldüğünde ölçme aracı, bireyin başarısını ölçmekten ziyade görme becerisini ölçmüş olacaktır. İlk bakışta hakkaniyetin ihlali geçerliğin de ihlali gibi görünmektedir. Çünkü ölçmek istediğimiz özelliğin ölçülmesine engel teşkil eden

özel bir durum testin yapmak istediğini yapamamasına sebebiyet vermektedir. Fakat yalnız geçerlik değil test puanlarının güvenilirliğine de etkisi nedeniyle ayrı bir psikometrik esas olarak ele alınmaktadır (AERA vd., 2014, ss.49-50).

Hakkaniyette grup farklılıkları genellikle; etnik köken, cinsiyet, kültür, dil, yaş, fiziksel engellilik, ekonomik farklılaşma gibi özelliklerle ifade edilmektedir. Hakkaniyetle ilgili erişilebilirlik ve evrensel tasarım olmak üzere iki temel kavramdan söz edilebilir. Erişilebilirlik insanların eksikliklerinden kaynaklanabilecek (dil bilmeme, görememe, duyamama, yürüyememe) durumların ölçme işleminde dikkate alınarak, amaçlananın yani ölçülmek istenen özelliğin gerçekleştirilmesi ile ilgilidir. Evrensel tasarım ise teste tüm bireylerin erişebileceği en optimum düzeyli test tasarımının oluşturulması ile ilgilidir. Test geliştiriciler ölçülecek yapıyı açıklık ve kesinlikle tanımlayıp ortaya koyacak, bu yapıyı ölçmeye engel teşkil edebilecek durumları tespit ederek, tasarımı bu tespitlere göre yapacaklardır. Bu tespitler ilave test uyarlamaları olabileceği gibi (Braille alfabesi ile test formu geliştirmek veya daha büyük fontlarla hazırlamak), ilave süre vermek, yardımcı veya refakatçi eşliğinde test uygulamak gibi başka tedbirler de olabilir. Hakkaniyet kavramının yalnız ölçme aracının yapısal özelliği ile ilgili olmadığı aynı zamanda uygulanması ile de ilgili olduğu görülmektedir. Ölçme ve değerlendirmede hakkaniyete yönelik kanıtlar başta DMF (Değişen Madde Fonksiyonu) olmak üzere çeşitli istatistiksel yöntemlerle ortaya konabilmektedir.

1.7.1.1.4. Madde ve Test Psikometrik Özellikleri

Madde ve test kavramları günlük yaşamda ve bilim dallarında çokça ve birçok anlamda kullanılmaktadır. TDK'na (2017) göre soru "Bir şey öğrenmek için birine yöneltile ve karşılık gerektiren söz veya yazı, soru" dir. Kazanım ya da yapının tanımlı davranışını ölçmede kullanılan madde, daima soru kipinde yazılmaz. "Değerlendiriniz, irdelleyiniz vb." ifadelerle de madde geliştirilebilir. Benzer bir durum sınav ve test kavramları için de geçerlidir. Test bir ölçme aracı, sınav ise test etme eylemi ve sürecidir. Bu nedenle terimsel olarak ölçme araçları ifade edilirken; *soru yerine madde; sınav yerine test* kullanımının ölçme bilimi açısından uygunluğu ortadadır (Erkuş, 2010).

Madde bir testteki yapı taşıdır. Tek başına ölçme yapabilir ve puanlanabilir. Bu sebeple puanları üzerine analizler yapılabilir. KTK'nda bir madde cevaplayıcılara

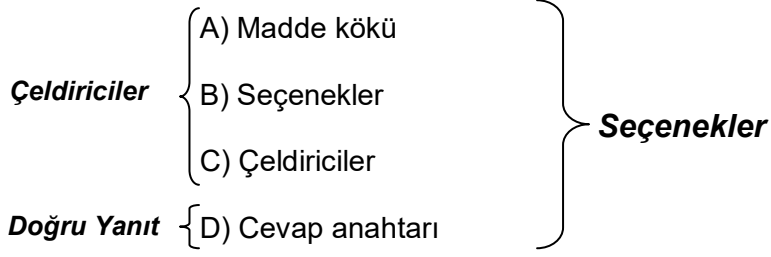
uygulandığında, cevaplayıcıların verdiği yanıtlara göre puanlanır. Bu hem cevaplayıcı puanı hem de madde puanı olarak yorumlanabilir. Bu sebeple KTK'ndaki madde analizi ile elde edilen bulgular elde edildikleri gruba bağlıdır. Madde yazarının, ölçme yapılacak grubun özelliklerine vakıf olmasına ve maddeyi geliştirmesine yardımcı olması sebebiyle, madde puanlarını analiz etmeye ve analiz sonuçlarını yorumlamaya hâkim olması gerekir (Mullis & Martin, 2013, s.9).

Maddenin biçimsel olarak tasarımında; yanıtlayanların bir cevabı seçmesi mi (seçmeli cevap) yoksa ortaya koyması mı (yapılandırılmış cevap) beklendiği esas alınmaktadır. Seçmeli maddeler tip bakımından; ikili tip, çoklu tip, Likert tipi, kategorik tip, kontrol formlu tip olmak üzere 5 türüdür (Kaplan ve Saccuzzo, 2012, s.159). Yapılandırılmış cevaplı maddeler ise; açık uçlu, boşluk doldurma, kısa cevaplı, metin/paragraf yazma, şekil/diyagram/grafik doldurma gibi yanıtlayıcıların seçerek değil de yazarak yanıtladıkları madde tipleridir.

Seçmeli maddeler içerisinde ÇSM en çok kabul gören ve en çok kullanılan nesnel ölçme aracıdır (Haladyna, 2004, s.10; Kaplan & Saccuzzo, 2012, s.160; Urbina, 2014, s.242;). Bu madde tipi, yüksek ve düşük seviyede yetenek veya başarı sergileyen kişilerin ayrımını yapmak için çok yönlü ve etkilidir (Haladyna, 2004, s.67). ÇSM'lerin yanıtlanması kısa sürdüğünden, yapılandırılmış maddelere nazaran tek oturumda daha fazla kullanılabilirler. Bu özelliği sayesinde yapılandırılmış cevaplı maddelere göre kapsam geçerliğine daha iyi kanıt sunmaktadırlar (Newstead & Dennis, 1994). Ayrıca madde sayısının artması güvenilirliği arttıran bir faktör (Angoff, 1953) olduğundan, güvenilirlik konusunda da diğer tip maddelere göre daha iyi düzeydedirler. Bu tip maddelerin uygulama ve puanlanması kolaydır. Puanlaması elle veya bir kısım cihazlar kullanılarak yapılabildiğinden çok sayıda cevaplayıcıya uygulanabilirler. Bir cevap anahtarına dayalı olduğundan puanlayıcı yargılarından bağımsızdır ve nesnedir. Ayrıca köküdeki soru, ifade veya hüküm ile ilgili doğru yanıt ile yanlış yanıtlar, seçeneklerde verildiğinden, öğretici bir yönü de bulunmaktadır. ÇSM'lere yapılan en büyük eleştiri üst düzey bilişsel süreçlerde ölçme yapamadığı yönündedir fakat yeterli çalışma, araştırma ve tecrübeye dayalı madde yazma kurallarını takip ederek geliştirildiklerinde, yapılandırılmış cevaplı maddelerle aynı türde beceri ve bilgiyi ölçebilirler (Hancock, 1994; Klein, 1998). ÇSM, şekil 1.1'de görüldüğü gibi biri kökü ve diğeri seçenekleri olmak üzere iki kısımdan oluşmaktadır.

Çoktan seçmeli test maddesinde bulunmayan unsur hangisidir?

} **Madde**
} **Kökü**



Şekil 1.1: ÇSM'nin Yapısı

Alanyazında seçeneklerin farklı puanlanması üzerine yapılan araştırmalar (Akkuş ve Baykul, 2000; Frary, 1982), bu tip uygulamaların, ölçme aracının geçerlik ve güvenilirliğine anlamlı katkı sağlamadığını belirtmektedirler. Bu sebeple günümüzde, özellikle geniş ölçekli testlerde, sıklıkla 1-0 puanlaması kullanılmaktadır (LYS, YGS, ALES, KPSS vb.). Puanlama sonucu yanıtlayanın kazanıma sahip olup olmadığı belirlendiği gibi maddenin ölçmeye dönük özellikleri olan madde psikometrik özellikleri de belirlenebilir. Araştırmada incelenen madde ve test psikometrik özellikleri Tablo 1.3'de özetlenerek verilmiştir.

Tablo 1.3: Madde ve Test Psikometrik Özellikleri

Psikometrik Özellik	Tanım	Hesaplama Yöntemi	Ölçüt
Madde Güçlüğü	Bir maddenin doğru yanıtlanma oranı ya da yüzdesidir. Kolaylık zorluk belirtir (Cohen & Swerdlik, 2009, s.159). Maddenin doğru yanıtlanma olasılığı olarak da görülebilir (Baykul, 2015, s.219).	Madde Güçlük İndeksi (p_i) $p_i = \frac{N_D}{N}$ (Crocker & Algina, 2008, s.311)	$0,20 \geq p_i \geq 0,00$ Çok zor $0,40 \geq p_i \geq 0,21$ Zor $0,60 \geq p_i \geq 0,41$ Orta güçlükte $0,80 \geq p_i \geq 0,61$ Kolay $1,00 \geq p_i \geq 0,81$ Çok kolay (Sözbilir, 2010)
Madde Ayıricılık Gücü	Madde ile ölçülmek istenilen özelliğe sahip olanlarla olmayanları ayırt edebilmesine madde ayıricılık gücü; maddenin bu özelliği, ölçmenin amacını ifade ettiğinden elde edilen indekse madde geçerlik katsayısı da denilmektedir (Crocker & Algina, 2008, s.313).	Madde Ayıricılık Gücü İndeksi (r_{jx})* $r_{j(x-j)} = \frac{r_{jx} \cdot S_x - S_j}{\sqrt{S_j^2 + S_x^2 - 2 \cdot r_{jx} \cdot S_x \cdot S_j}}$ (Baykul, 2015, s.242)	$r_{jx} \geq 0,30$ Madde uygun $0,29 \geq r_{jx} \geq 0,20$ Madde düzeltilmeli $0,19 \geq r_{jx}$ Madde uygun değil (Ebel & Frisbie, 1991, s.232)
Madde Çeldiriciliği**	Maddenin çeldiricilerinin birlikte ne derecede dengeli işlediğinin ifadesidir.	Madde Çeldiricilik İndeksi (ζ) $\zeta = 1 - \frac{\sqrt{\sum_{i=0}^n (f_i - \delta)^2}}{\sqrt{(q \cdot n - \delta)^2 + (a - 1)(0 - \delta)^2}}$	Her bir çeldirici frekansının en az %5 yanıtlayıcı tarafından seçilmiş olması gerekir (Haladyna, 1989b).
Test Varyansı	Bireylerin testten elde ettikleri puanların, grubu temsil eden bir değer olan aritmetik ortalamanın etrafında nasıl dağıldığını belirten istatistiktir (Baykul, 2015, s.250).	Test Varyansı (S_x^2) $S_x^2 = \sum_{j=1}^K S_j^2 + 2 \cdot \sum_{j < k} S_{jk}$ (Baykul, 2015, s.257)	Test geliştirmede çoğu kez test varyansının mümkün olduğu kadar büyük olması istenir (Baykul, 2015, s.259). İki testin varyansı kıyaslanırken, anlamlı şekilde daha büyük varyans değerine sahip testin öğrencileri ölçülen özellik bakımından daha fazla ayırttığı ve daha iyi ölçme yapabildiği yorumu yapılmıştır.
Test Güvenirliği	Maddelerin tek boyutlu özelliği birbirleriyle ne derece tutarlı ölçtüğüne işaret kabul edilebilecek bir istatistiktir (Schmitt, 1996; Tan 2009a).	İç Tutarlık Güvenirliği (KR-20) $KR - 20 = \left(\frac{K}{K - 1} \right) \cdot \left[\frac{\sum_{j=1}^K p_j \cdot (1 - p_j)}{S^2(X)} \right]$ (Kuder & Richardson, 1937)	KR-20 Yorum 0,95 Ölçme hatası yok denecek kadar etkisiz 0,85 Standart başarı testleri için yüksek güvenirlilik 0,75 Sınıfçı testler için orta-düşük güvenirlilik 0,50 Gerçek puanlar ve hata miktarı eşit düzeyde (Davidshofer & Murphy, 2005, s.150)

Test Kapsam Geçerliği	Ölçme aracının ölçmek istediği davranış evrenini ne derece iyi temsil ettiğinin belirlenmesidir (Mehrens & Lehmann, 1991, s.266).	KGO $KGO_i = \frac{N_G - \frac{N}{2}}{\frac{N}{2}}$ (Lawshe, 1975)	Uzman Sayısı Kritik KGO Değeri 12 uzman 0,667 14 uzman 0,571 (Ayre & Scally, 2014) Kritik ve üzerinde KGO değerine sahip maddelerin ortalaması KGİ değerini oluşturur. KGİ değerinin 0,67'den büyük olması beklenir (Alpar, 2014, s.509).
Test Uyum Geçerliği	Bir testin geçerliğine kanıt sunmada önceden gerçekleştirilmiş geçerli ölçmelerle, test puanları arasındaki ilişkiyi ortaya koymaktır (Ebel & Frisbie, 1991, s.106).	PMÇK Katsayısı (r) $r(X, Y) = \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - (\sum_{i=1}^n x_i) \cdot (\sum_{i=1}^n y_i)}{\sqrt{[n \cdot \sum_{i=1}^n x_i - (\sum_{i=1}^n x_i)^2] \cdot [n \cdot \sum_{i=1}^n y_i - (\sum_{i=1}^n y_i)^2]}}$ (Pearson, 1895)	Uyum geçerliği çalışmalarının sonucunda elde edilen korelasyon katsayısının 0,80 civarında olması önerilmektedir (Ebel & Frisbie, 1991, s.106).
Test Yapı Geçerliği	Bir araçla ölçülmek istenen yapının o araçla ortaya konulma derecesidir (Lord & Novick, 1968, s.278).	AFA (TBA, PAY)*** Madde Özdeğerleri Özdeğer Açıklama Yüzdeleri Madde Faktör Yükleri (Tabachnick & Fidell, 2015, s.612; Büyüköztürk, 2015, s.133)	Boyut/Faktör belirlemede: PAY ile elde edilen faktör sayısı; öz değeri 1'den büyük faktörler; toplam varyansın 2/3'ünü kapsayan faktör sayısı; tek faktörlü testlerde açıklanan varyansın %20 ve daha fazla olması (Reckase, 1979); faktör öz değerine dayalı çizilen çizgi grafiğinin incelenerek yüksek ivmeli, hızlı düşüşlerin yaşandığı faktörler esas alınır. Madde değerlendirilmede: Madde faktör yüklerinin 0,30 ve üzerinde olması; tek bir faktöre yüksek yük diğer faktörlere düşük yük vermesi beklenir (en yüksek ile bir sonraki faktöre verilen yük değeri arasındaki farkın en az 0,10 olması). (Büyüköztürk, 2015, ss.134-135)

* Sürekli bir değişkenle (test puanı) iki kategorili bir değişken (madde puanı) arasındaki ilişkinin en uygun ölçüsü nokta çift serili korelasyon katsayısıdır (Baykul, 2015, s.241). Nokta çift serili korelasyon formülünde madde ayıricılığı hesaplanan maddenin puanının da test puanı içerisinde yer alması sebebiyle, madde sayısının 25'den az olduğu hallerde olduğundan daha büyük çıkmaktadır. Bu nedenle bu araştırmada maddenin test içerisindeki etkisinin arındırılmasında düzeltilmiş nokta çift serili korelasyon katsayısı hesaplanmıştır.

** Müteakip başlıkta açıklanmıştır.

*** Faktör çıkarma teknikleri içerisinde en sık kullanılan ve görel olarak yorumlanması kolay olanı TBA (Tabachnick & Fidell, 2015, s.637), TBA'nde faktör sayısını belirlemede kullanılan en yaygın yöntem ise PAY'dir (Timmerman & Lorenzo, 2011).

1.7.1.1.5. Madde Çeldiriciliği ve İndeksi

Birçok test geliştirme uzmanına göre, madde analizinin hareket noktası madde güçlüğü, ayırtediciliği ve cevapların seçeneklere dağılımıdır. Bunlar, bir maddenin işleyişi ile ilgili bilinmesi zorunlu göstergelerdir (Doğan ve Tezbaşaran, 2003).

Problem durumu

Çeldiriciler, madde kökünde yazılı uyarana verilecek yanlış tepkiyi temsil ederler. Çeldiricilerin, kazanıma sahip olmayan cevaplayıcılar tarafından tercih edilmesi beklenir. Bu sebeple öğrencilerin performansı çeldiricilerin tasarlanması ile yakından ilişkilidir (Dufresne, Leonard & Gerace, 2002). Çeldiricileri iyi tasarlanmış bir maddede, maddeye doğru cevap verenlerin dışında kalanların yani maddeye yanlış cevap verenlerin (bir öğrenme yanışı ya da kavramsal hata yoksa) çeldiricilere eşit dağılması beklenir. Çeldiricilerin kazanıma sahip olmayanlar ile kısmen sahip olanlara cazip gelerek seçilmesi, kazanıma tam anlamıyla sahip olanların ortaya çıkmasına böylece de maddenin özelliğe sahip olanlarla olmayanları ayırmasına yani geçerliğine neden olması beklenir. Doğru yanıt verenlerin tüm yanıt verenlere oranının istatistiksel olarak sağlıklı olabilmesi diğer bir deyişle madde analizi sonucu elde edilen madde güçlük indeksinin güvenilir bir değer alması; tesadüfen olmamasına, seçeneklerin eşit derecede güçlü olmasına bağlıdır (Horst, 1933, s.231).

Çeldiricilerdeki işlerlik, maddenin iyi inşa edilmiş olup olmadığına veya amacını gerçekleştirmede başarılı olup olmadığına delil olarak sunulabilir (Hingorjo & Jaleel, 2012). Öğrencilerin % 5'inden azı tarafından seçilen herhangi bir çeldirici, işlevsiz bir çeldirici olarak kabul edilir (Haladyna, 1989b). İşlevsiz çeldiriciler, özellikle ÇSM cevaplama konusunda tecrübeli olan öğrencilerin şansla doğru yanıt bulma olasılığını arttıracaktır. Bu sebeple madde analizinde güçlük ve ayırtıcılık gücü hesaplanır iken, çeldiricilere dönük analizlerin yapılması önem arz eder.

Çeldiricilerin madde güçlüğü ile ilişkisi

Çeldiricileri tercih eden öğrenci sayısı arttıkça, madde güçlük indeksi daha küçük değer alacak ve madde daha zor olacaktır. Aynı şekilde, çeldiricilere yönelen öğrenci sayısı azaldıkça, madde güçlük indeksi de artacak, madde kolaylaşacaktır.

Bu sebeple madde güçlüğü çeldiricilerin tercih edilmesiyle yani çeldiricilerin işlerliğiyle yakından ilgilidir.

Çeldiricilerin ayırıcılık gücü ile ilişkisi

Teorik olarak bakıldığında ayırıcılığı yüksek maddelerin çeldiricilerinin de iyi işlemesi gerekmektedir (Haladyna & Downing, 1993). Hingorjo ve Jaleel (2012) orta düzeyde güçlükte ve yüksek ayırıcılık gücüne sahip maddelerin çeldiricilerinin ayırıcılık gücü düşük maddelere göre daha işler olduğunu belirtmektedir. Bu nedenle bir maddenin ayırıcılık gücü çeldiricilerin işlerliği ile yakından ilişkilidir ve kökü ile doğru yanıt aynı kalmak kaydıyla çeldiricileri değiştirildiğinde madde güçlük ve ayırıcılık gücünün de değişeceği söylenebilir (Haladyna, 2004, s.273). Ayrıca maddenin daha fazla öğrenci tarafından seçilen çeldiricileri ile daha az öğrenci tarafından seçilen çeldiricileri üzerinde yapılacak değişikliklerden sonra, maddenin ayırıcılık gücünün biraz daha yükselmesi beklenir (Baykul, 2015, s.327). Bu sebeple madde ayırıcılık gücü indeksi değerinin ortaya çıkmasında çeldiricilerin ne derecede işlediği etkilidir. Aynı ayırıcılık gücü indeksi değerine sahip iki maddeden çeldiricileri daha dengeli işleyen maddenin ayırıcılık gücü; çeldiricileri daha dengesiz işleyen maddeye göre istatistiksel olarak daha anlamlı olabilecektir. Çünkü çeldiriciler eksik bilgiye sahip ve bu bilgiye hiç sahip olmayan öğrencilerce tercih edilebilecek uygunlukta geliştirilerek; doğru yanıtın şans başarısı ile seçilmesi olasılığı düşürülmüş olacaktır. Bu gerekçelerle madde ayırıcılık gücünün, çeldiricilerin işlerliği ile ilişkili olduğu iddia edilebilir. Bu noktada üzerinde durulacak husus; ayırıcılık gücü düşük bir maddenin de çeldiricilerinin dengeli işleyebileceğidir. Çeldiricilerin dengeli işlemesi, madde ayırıcılığını arttıran bir etkiden ziyade, ortaya çıkan ayırıcılık gücü indeksinin istatistiksel olarak sağlıklı değer aldığına işaret olarak kabul edilebilecek bir durumdur.

Seçenek analizi

Madde analizinde; maddeyi doğru yanıtlayanların sayısının maddeyi alanların sayısına oranından madde güçlük indeksi; maddenin kazanıma sahip olanlarla olmayanları ayırabilmesine işaret olacak istatistiklerle (alt-üst grup, nokta-çift serili korelasyon vb.) de madde ayırıcılık gücü indeksi hesaplanır. Alanyazında 60'tan fazla madde analizi yöntemi bulunduğu bilinmektedir (Baykul, 2015, s.320). Bu analizlerde (Henrysson yöntemi, alt-üst gruplar, çeldirici tercih yüzdeleri, test

puanları ile korelasyon vb.) ortak olan nokta çeldiricilerin analizlerinin çeldirici bazında yapılması, tüm çeldiricilere tahvil edilebilecek tek bir indeksin ise bulunmamasıdır. Deneme uygulaması sonucu çeldiricilerden daha az işleyenlerin değiştirilerek daha iyi düzeyde ölçme yapabilecek maddeler geliştirilebilir ve bu şekilde çeldirici analizi fayda sağlayabilir. Nihai test uygulaması sonucu hangi çeldiricilerin işlediğinin tespit edilmesine dönük bir analizin öğrencilerin yanlış öğrenmeleri, eğitim sürecinin eksikleri, eğitici performanslarının ortaya konmasında yararlı aynı zamanda madde yazarının ne derecede işler çeldirici üretebildiğine de işaret olabilecektir. Madde kökü yazmak ve bu madde köküne uygun doğru yanıtı belirlemek, çeldirici üretmeye nazaran daha kolaydır (Gregory, 2011, s.140). Özellikle 4 veya 5 seçenekli ÇSM'lere çeldirici üretmek, madde yazarının en çok zorlandığı husus olabilmektedir (Wesman, 1971, s.99). Bu sebeple bir maddenin çeldiricilerinin ne derecede dengeli işlediğini ortaya koyan bir indeks kullanılarak, madde yazarının madde geliştirme becerisi ile ilgili yorumlanmasının uygun olacağı iddia edilebilir.

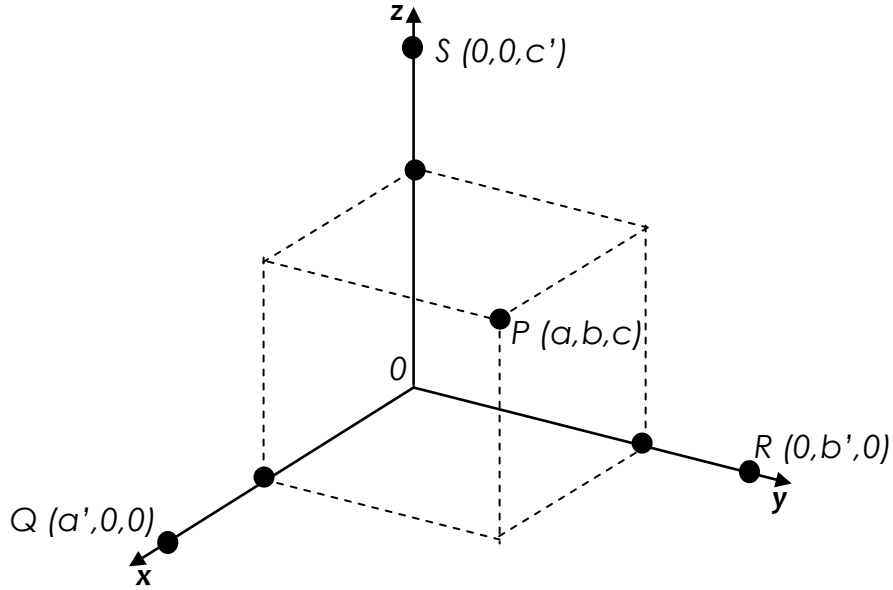
Çeldiricilere yönelik indeksler

Alanyazında kategorik verilerin dağılımlarını özetleyen bazı istatistikler mevcuttur (Dickinson, 2015; Wilcox, 1973). Genel olarak niteliksel dağılım indeksi (Index of Qualitative Variation) olarak adlandırılan bu istatistikler incelendiğinde Dickinson'ın (2014) da belirttiği gibi yüksek düzeyde negatif çarpıklığa sahiptirler. Çünkü bu istatistikler her bir çeldiricinin alması gereken değeri dikkate almadan çeldirici dağılımlarını ortaya koymaktadır. Bu sebeple araştırmacı tarafından, çeldiricilerin kuramsal olarak alması gereken değerden hareketle bir istatistik belirlenmiştir ve *madde çeldiricilik indeksi (item distraction index)(Ç)* olarak adlandırılmıştır.

Çeldiricilik indeksi varsayımları ve formülün elde edilmiş yöntemi

Mükemmel ÇSM'lerin iki karakteristiği vardır. Birincisi, maddenin ölçtüğü kazanıma sahip olanlar her zaman doğru yanıtı seçerler, ikincisi ise kazanıma sahip olmayanlar çeldiricilere rastgele dağılırlar (Davidshofer & Murphy, 2005, s.204; Gregory, 2011, s.145). Diğer bir deyişle, ÇSM'nin çeldiricilerinin eşit sayıda yanlış cevaplayıcı tarafından seçilmesi arzu edilir. Çeldiricilere dağılımın en arzu edilen durumu; tüm çeldiricilere yönelen yanlış cevaplayıcı sayısının eşit, en istenmedik durum ise çeldiricilerden yalnız birine tüm yanlış cevaplayıcıların toplanmasıdır.

En istendik ve en istenmedik durumların çeldirici frekansları, örneğin 4 seçenekli, 3 çeldiricili bir madde için, üç boyutlu koordinat sisteminde gösterilirse;



Şekil 1.2: Üç Çeldiricili Maddenin Çeldirici Frekanslarının Üç Boyutlu Düzlemde Gösterimi

$a = b = c$ olduğu P noktası en istendik durum, Q, R ve S noktaları ise en istenmedik durumun koordinat düzlemindeki yerleridir. Çeldiricilerin frekanslarının düzlemdeki kesişim noktası P noktasına yaklaştıkça çeldiricilik güçlenecek; Q, R, S'ye yaklaştıkça zayıflayacaktır. Herhangi bir çeldirici frekansları kesişim noktasını \bar{P} ile gösterirsek, \bar{P} noktasının P'ye olan uzaklığı;

$$|\mathbf{P}\bar{\mathbf{P}}| = \sqrt{(a - a')^2 + (b - b')^2 + (c - c')^2} \quad (\text{eşitlik 1})$$

(1) eşitliğinden hesaplanabilir. Doğru parçasının uzunluğu minimum değerini, \bar{P} koordinatının P koordinatına eşit olduğu durumda alarak "0" olacaktır. Bu uzunluğun maksimum değeri ise aşağıda birbirine eşit üç denklemden biri ile hesaplanabilecektir;

$$\begin{aligned}
|\mathbf{P}\bar{\mathbf{P}}| &= \sqrt{(\mathbf{a} - \mathbf{a}')^2 + (\mathbf{b} - \mathbf{0})^2 + (\mathbf{c} - \mathbf{0})^2} && \text{(eşitlik 2)} \\
&= \sqrt{(\mathbf{a} - \mathbf{0})^2 + (\mathbf{b} - \mathbf{b}')^2 + (\mathbf{c} - \mathbf{0})^2} \\
&= \sqrt{(\mathbf{a} - \mathbf{0})^2 + (\mathbf{b} - \mathbf{0})^2 + (\mathbf{c} - \mathbf{c}')^2}
\end{aligned}$$

Eşitlik (2) ile elde edilen uzunluğun, eşitlik (1)'in ne kadarına eşit olduğu ise;

$$\frac{\sqrt{(\mathbf{a} - \mathbf{a}')^2 + (\mathbf{b} - \mathbf{b}')^2 + (\mathbf{c} - \mathbf{c}')^2}}{\sqrt{(\mathbf{a} - \mathbf{a}')^2 + (\mathbf{b} - \mathbf{0})^2 + (\mathbf{c} - \mathbf{0})^2}} \quad \text{(eşitlik 3)}$$

(3) oranı ile bulunabilir. Bu elde edilen uzunluğun mümkün maksimum uzunluğa oranını vermektedir. Bir başka ifade ile maddenin çeldirici frekanslarının, mümkün en istenmedik çeldirici frekansına (doğru yanıtlamayanların tek bir çeldiriciye yönleneceği) oranını vermektedir. Bu oranı çeldiricilerle aşağıdaki şekilde ifade edebiliriz.

$$\frac{\sqrt{\sum_{i=1}^n (f_i - \delta)^2}}{\sqrt{(q \cdot n - \delta)^2 + (a - 1)(0 - \delta)^2}} \quad \text{(eşitlik 4)}$$

f_i : i'nci çeldiricinin frekansı.

δ : çeldirici başına beklenen frekans

$$\delta = q \cdot n / a$$

a : çeldirici sayısı.

$q \cdot n$: maddeyi doğru yanıtlamayanların frekansı.

q : maddenin doğru yanıtlanmama olasılığı ($q = 1 - p$).

n : maddeyi alan cevaplayıcı sayısı (yanıtlamayanlar da dahildir).

(4) oranının 1 değerini alması, madde çeldirici frekanslarının en istenmedik dağılıma sahip olduğu yani bir çeldiricinin tüm doğru yanıtlamayan cevaplayıcılarca seçilmesi anlamına gelmektedir. Tüm çeldiricilerin eşit sayıda

cevaplayıcı tarafından seçilmesi durumunda ise pay 0 değerini alacağından bu oran 0 olacaktır.

Buradan Madde Çeldirici İndeksi (Ç) ise;

$$\text{Ç} = 1 - \frac{\sqrt{\sum_{i=0}^n (f_i - \delta)^2}}{\sqrt{(q \cdot n - \delta)^2 + (a - 1)(0 - \delta)^2}} \quad (\text{eşitlik 5})$$

(5) eşitliği ile hesaplanabilir. Burada Ç, çeldirici frekans dağılımının, mümkün olan en istenmedik dağılıma sahip olmama olasılığı ya da en istendik dağılıma sahip olma olasılığıdır. Eşitlik (5)'den anlaşılacağı üzere, Ç; 0 ile 1 arasında değer alır. Değeri 1'e yaklaştıkça çeldiricilerin dengeli şekilde işlediği, 0'a yaklaştıkça da çeldiricilerin birlikte dengeli işlemediği anlamına gelir.

Kabul edilebilir en düşük çeldiricilik indeksinin hesaplanması

Madde çeldiricilerinin dengeli işlemesi gerektiği iddiasına dayanak olan varsayım çeldiricilere yönelen cevaplayıcı sayılarının eşit olmasıdır. Çeldiricilerle ilgili diğer beklenti ise; bir çeldiricinin maddeyi alan tüm öğrencilerin sayısının en az % 5'i kadarı tarafından tercih edilmesidir (Haladyna, 1989b). Bu nedenle çeldiricilik indeksinin yorumlanmasında ölçüt bir değer olarak kabul edilebilir en düşük madde çeldiricilik indeks (Ç_{\min}) değerinin hesaplanması gerekmektedir. Ç_{\min} , üç çeldiricili bir maddede, alt sınır bir değer olarak tüm çeldiricilerin maddeyi alanların sayısının % 5'i kadar olacak çeldirici frekansları esas alınarak aşağıdaki şekilde hesaplanır.

$$\text{Ç}_{\min} = 1 - \frac{\sqrt{a \cdot (n \cdot 0,05 - \delta)^2}}{\sqrt{(q \cdot n - \delta)^2 + (a - 1)(0 - \delta)^2}} \quad (\text{eşitlik 6})$$

Elde edilen çeldiricilik indeksi, eşitlik (6) ile verilen kabul edilebilir en düşük çeldiricilik indeksine eşit ve büyük ise, maddenin çeldiricilerinin dengeli işlediği; bu değerden küçük ise dengeli işlemediği yorumu yapılabilir. Ayrıca aynı ayırıcılık gücü indeksine sahip iki maddeden, çeldiricilik gücü indeksi daha büyük değere sahip olanın, ölçme gücünün daha iyi düzeyde olduğu öne sürülebilir.

Örnek çeldiricilik indeksi hesaplaması

Örnek bir veri seti esas alınarak çeldiricilik indeksinin hesaplanmasının uygun olacağı değerlendirilmiştir. TAP 16.11.13 programıyla, 20 maddeli, 4 seçenekli, 100 cevaplayıcılı ve ortalama madde güçlüğü 0,50 olan bir teste ait veri seti üretilmiş ve Tablo 1.4'de sunulmuştur.

Tablo 1.4: Örnek Madde Çeldiricilik İndeksi Hesaplaması İçin Seçenek Frekansları

Madde No	A Seçeneği	B Seçeneği	C Seçeneği	D Seçeneği	p_i	r_{jx}	ζ
1	9	17	43*	31	0,43	0,29	0,66
2	4	52*	8	36	0,52	0,39	0,37**
3	9	39*	18	34	0,39	0,28	0,64
4	3	10	8	79*	0,79	0,17	0,70**
5	12	18	38	32*	0,32	0,36	0,65
6	5	60*	12	23	0,60	0,21	0,61
7	7	19	33	41*	0,41	0,35	0,62
8	20	17*	21	42	0,17	0,26	0,74
9	75*	2	6	17	0,75	0,40	0,46**
10	6	25	25*	44	0,25	0,16	0,56
11	14	24	50	12*	0,12	0,15	0,63
12	44*	4	30	22	0,44	0,25	0,59**
13	71*	3	9	17	0,71	-0,02	0,58**
14	5	8	78*	9	0,78	0,23	0,84
15	60*	5	16	19	0,60	0,32	0,68
16	8	45*	15	32	0,45	0,35	0,61
17	10	50*	14	26	0,50	0,38	0,71
18	15	19	36*	30	0,36	0,31	0,79
19	8	15	58*	19	0,58	0,37	0,77
20	18	17	20	45*	0,45	0,35	0,95

* Doğru Yanıt

** $\zeta < \zeta_{min}$

Birinci maddeyi 100 öğrenci yanıtlamış, 43 öğrenci doğru yanıtta, 57 öğrenci ise çeldiricilere yönelmiştir. Çeldirici sayısı 3, çeldirici başına beklenen frekans, $\delta = 57/3 = 19$ 'dur. Bu değerler ile madde çeldiricilik indeksi hesaplanırsa;

$$\zeta = 1 - \frac{\sqrt{(9 - 19)^2 + (17 - 19)^2 + (31 - 19)^2}}{\sqrt{(57 - 19)^2 + (2)(0 - 19)^2}}$$

= 0,66 değeri elde edilir.

0,43 güçlükteki 57 çeldirici tercihli madde için kabul edilebilir en düşük çeldiricilik indeksi ise;

$$\zeta_{\min} = 1 - \frac{\sqrt{3 \cdot (100.0,05 - 19)^2}}{\sqrt{(57 - 19)^2 + 2(0 - 19)^2}}$$

= 0,27 değeri elde edilir.

Birinci maddeye ait ζ değeri, ζ_{\min} 'den büyük olduğu için, çeldiricilerinin dengeli işlediği söylenebilir.

Tablo 1.4'de 16'ncı ve 20'nci maddeler dikkate alındığında aynı madde güçlük ve ayıricılık gücü indeksine sahip oldukları görülmektedir. Fakat 20'nci maddenin çeldiricilik indeksi (0,95), 16'nci maddenin çeldiricilik indeksinden (0,61) daha büyük bir değere sahip olduğundan, ölçme gücünün daha iyi düzeyde olduğu söylenebilir. Eğer bu iki madde aynı kazanımı aynı bilişsel basamakta ölçmek üzere geliştirilmiş iseler, 20'nci maddenin öğrencilerin ilgili özelliğe sahip olup olmadığının yoklanmasında 16'nci maddeye tercih edilmesinin uygun olacağı söylenebilir.

Tablo 1.4'deki 2, 4, 9, 12 ve 13'üncü maddelerin ζ_{\min} değerinden daha düşük düzeyde oldukları görülmektedir çünkü asgari 5 tercihten (100.0,05) daha düşük frekansta çeldiricilere sahiptirler.

Maddenin madde çeldiricilik indeksine, madde güçlük ve ayıricılık gücü indekslerinin düzeylerinin oluşmasının altında yatan bir istatistik gözüyle bakılabilir. Araştırmada analizleri yapılan maddelerin çeldiricileri için bu istatistik hesaplanmış ve çeldiricilerin ne derecede dengeli işler olduğunun yorumlanmasında kullanılmıştır.

1.7.2. Test Geliştirme

Psikometrik özelliklerine değinilen testler bir takım işlemlerle elde edilmektedir. Özellikleri önceden belli bir yaklaşıklıkla kestirilebilen test hazırlama işi olan test geliştirme (Baykul, 2015, s.269), farklı başlıklar altında sıralanmakla (Baykul, 2015, s.269; Crocker & Algina, 2008, s.66; Downing & Haladyna, 2006, s.5; Kline, 2015a, s.147; Urbina, 2014, s.237) birlikte icra edilen faaliyetler bakımından benzerdir. Genel kabul gören ve bu araştırmada esas alınan test geliştirme adımları aşağıdaki gibidir.

1. Testin Amacının, Ölçeceği Özelliğın, Kapsamının Tespiti, Belirtke Tablosunun Elde Edilmesi,

2. Madde Geliştirme,
3. Dış Kontrol,
4. Test Formunun Oluşturulması,
5. Testin Uygulanması,
6. Puanlama, Analiz, Teknik İşlemler.

Her test belirli bir amaç için geliştirilip, kullanılır. Bu amaçlar; kişilerin seçimi (seçme, yerleştirme, işe kabul vb.), bireylerin sınıflanması (ruh, akıl hastalığı, yeteneğe göre kabul vb.), uygulanan yöntemlerin değerlendirilmesi (psikoterapi, ilaç kullanımı sonuçları vb.), araştırmalarda bilimsel denencelerin kontrol edilmesidir (Özgüven, 2014, s.25). Amacın belirlenmesinden sonra bu amaca ulaşmak için yoklanacak özelliklerin tespit edilmesi gerekir. Örneğin üniversiteye öğrenci seçimi amacıyla geliştirilecek bir testin, öğrencilerin hangi özelliklerini yoklayacağı belirlenmelidir (sayısal yetenek, okuduğunu anlama vb.). Bu özellikler psikolojik yapılar olabileceği gibi, öğrencinin öğrenme dönemi sonunda ne bilmesinin, ne anlamasının ve/veya yapabilmesinin belirlendiği ifadeler olan kazanımlar (Donnelly & Fitzmaurice, 2005) da olabilir. Kazanımlar ya da diğer bir ifade ile davranışlar bilişsel (bilgiye dönük), duyuşsal (duyguya dönük) ve devinişsel (bedeni faaliyetlere dönük) olarak sınıflandırılır. Bilişsel süreçlerin sınıflandırılması (taksonomisi) için birçok yaklaşım bulunmakla (Yüksel, 2007) birlikte Bloom Taksonomisinin (Anderson, Krathwohl & Bloom, 2001) sıklıkla kullanıldığı söylenebilir (Lee, Kim, Jin, Yoon & Matsubara, 2017, s.6). Bilişsel süreçlere ilişkin taksonomiler insan zihninin işlemsel fonksiyonunun basitten karmaşığa sıralamasıdır. Anderson vd.'ne (2001) göre bu sıralama; hatırlama, anlama, uygulama, analiz, değerlendirme ve yaratmadır. Hangi kazanımların yoklanacağını belirlenmesi ile de testin kapsamı belirlenmiş olacaktır. Öğrencileri üniversiteye seçmek amacıyla kullanılacak ve sayısal yetenek özelliklerini yoklayacak bir testin hangi kapsamda ölçme yapacağını belirlenmesi ile test geliştirmenin birinci adımı tamamlanır ve bir belirtke tablosu oluşturulur. Tablo 1.5'de örnek bir belirtke tablosu gösterilmiştir.

Tablo 1.5: Örnek Belirtke Tablosu

Öğrenme Alanı	Alt Öğrenme Alanı	Hatırlama	Anlama	Uygulama
Sayılar ve İşlemler	Çarpanlar ve Katlar		-	8.1.1.2. İki doğal sayının en büyük ortak bölenini ve en küçük ortak katını hesaplar. (*)
	Üslü İfadeler		8.1.2.4. Sayıları 10'un farklı tam sayı kuvvetlerini kullanarak ifade eder. (*)	-

(*) Madde sayısı

Belirtke tablosu ile öğrenme alanı, kazanımlar, kazanımların bilişsel basamakları ve madde sayıları ortaya konarak, testin kapsamı belirlenmiş olacaktır. Belirtke tablosu bu haliyle, testin kapsam geçerliğine yönelik sunulacak bir kanıt ayrıca madde geliştirmede kullanılacak bir rehber durumundadır. Belirtke tablosunun oluşturulmasından sonra maddelerin geliştirilmesi adımına geçilir.

Madde geliştirme adımı madde yazarınca işletilen bir süreçtir. Test geliştirici veya geliştiriciler madde yazarına aşağıdaki bilgileri içeren bir bilgi formu vermelidir.

1. Madde yazarınca madde geliştirilmesi istenen kazanımlar ve madde miktarları,
2. Madde tipi ve seçenek adedi,
3. Maddenin uygulanacağı grubun özellikleri (seviye, eğitim sürecinin hangi aşamasında oldukları, genelden farklı zihinsel, duygusal, bedeni durumları),
4. Test puanları ile verilecek karar (seçme, seviye tespit).

Bu bilgiler, madde yazarına geliştireceği maddenin sınırları bakımından gereklidir. Test puanlarının ne amaçla kullanılacağına madde yazarına bildirilmesi, maddenin yaklaşık güçlüğüne belirlenmesine ışık tutacaktır. Burada altı çizilmesi gereken bir husus da; seçenek sayısının cevaplayıcıların eğitim düzeylerine uygun olması ve tüm maddelerdeki seçenek sayısının istatistikî analizlerin doğrulukla yapılabilmesi için aynı olması gerektiğidir (Atılğan, Kan ve Doğan, 2009, s.249; Baykul, 2015, s.301). Bu, test geliştirici tarafından alınması gereken bir karardır. Madde yazarının bu ölçütlerle kazanımların bireylerde bulunma derecesini ortaya çıkaracak ölçme aracını oluşturması işlemine *madde geliştirme* denir. Madde geliştirme de test geliştirme gibi bir süreçtir. Planlı ve sistematik çalışabilme adına

bir çizelge, form veya şablon kullanmak faydalı olabilecektir. Tablo 1.6'da arařtırmacı tarafından geliřtirilen, bu arařtırmada madde yazarlarına kullanmaları yönünde tavsiye edilen madde geliřtirme formu sunulmuřtur. Formda madde, madde kökü, seenekler, dođru yanıt ve eldiricilerin Tablo 1.8'de verilen madde yazım kurallarına uygunlukları, madde yazarınca kontrol edilmesi için kutucuklar bulunmaktadır. Bu kontroller madde yazarınca yapılacak kontrollerdir ve *i kontrol* olarak adlandırılabilirler. Madde geliřtirme iřlemi i kontrol tamamlanarak test geliřtiriciye aday maddenin teslim edilmesi ile son bulur.

Tablo 1.6: Madde Geliştirme Formu

KAZANIM	8.1.1.3. Verilen iki doğal sayının aralarında asal olup olmadığını belirler.																		
BİLİŞSEL BASAMAK	Anlama																		
ALAN	Matematik																		
ÖĞRENME ALANI	Sayılar ve İşlemler																		
ALT ÖĞRENME ALANI	Çarpanlar ve Katlar																		
KAZANIMIN ÖN KOŞULLARI	1) Görme kusuru olmama.																		
	2) Türkçe okuyabilme.																		
	3) Okuduğunu anlayabilme.																		
	4) Kalem tutabilme ve kullanabilme.																		
	5) Doğal sayıları tanıma ve bilme.																		
	6) Aralarında asal kavramını bilme.																		
YOKLANACAK GRUP	MEB 8'inci sınıf düzeyi öğrencileri																		
GRUBUN EĞİTİM DURUMU	2016-2017 eğitim ve öğretim yılı 1. dönem sonu kazanımlarına sahip durumdadır.																		
MADDE TİPİ	Çoktan Seçmeli Madde																		
SEÇENEK ADEDİ	4																		
MADDENİN AMACI	Kazanımın hedef grupta bulunup bulunmadığını yoklamak.																		
PUANLAMA ÖLÇÜTÜ	Tek bir işaretleme yapılması ve bu işaretlemenin de doğru yanıt olması durumunda "1" puan; hiç işaretlememe, birden çok işaretleme, tek bir işaretleme yapılması ve bu işaretlemenin de çeldiricilerden birisi olması "0" puan verilir.																		
CEVAPLAMA SÜRESİ	90 saniye																		
MADDE DAYANAĞI	Ortaokul Matematik Öğretim Programı MEB (2013); Aydın (2016)																		
ŞEKİL, GRAFİK, RESİM, ÇİZİM (varsa)																			
MADDE KÖKÜ																			
KURALA UYGUNLUK	20	21	22	23	24														
DOĞRU YANIT																			
KURALA UYGUNLUK	38									39									
ÇELDIRİCİ-1																			
KURALA UYGUNLUK	34	35	36	37															
ÇELDIRİCİ-2																			
KURALA UYGUNLUK	34	35	36	37															
ÇELDIRİCİ-3																			
KURALA UYGUNLUK	34	35	36	37															
SEÇENEKLERİN KURALA UYGUNLUĞU	25	26	27	28	29	30	31	32	33										
MADDENİN BÜTÜN OLARAK KURALLARA UYGUNLUĞU	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

Madde yazarı aday maddeyi geliřtirdikten sonra, maddenin bazı yönlerden kontrol edilmesi gerekir (Baykul, 2015, s.311). *Dış kontrol* olarak adlandırılabilen bu işlemler aşağıdaki gibidir.

1. Ölçme uzmanlarınca geçerlik ve uygunluğun teknik kontrolü,
2. Alan uzmanlarınca bilimsel kontrolü,
3. Dilbilgisi uzmanlarınca anlaşırılık ve dilbilgisi, noktalama işaretleri, yazım kuralları ile imlâ kurallarına uygunluğun kontrolü,
4. Eğitim psikoloğunca ve/veya sosyoloğunca cevaplayıcı grubunun psikolojik/sosyolojik özelliklerine uygunluğun kontrolü,
5. Benzer bir gruba uygulanarak elde edilen puanlarca psikometrik özelliklerin uygunluğunun kontrolü (Deneme Uygulaması).

Dış kontrol sonucu tespit edilen aksaklıklar madde yazarına dönüt olarak verilir. Düzeltmesi maddenin geçerliğini etkilemeyecek olanlar madde yazarınca yapılabilir. Eğer tespit edilen husus maddenin geçerliğini etkileyecek bir düzeltme gerekiyorsa, maddenin tekrar geliştirilmesi daha uygun olacaktır. Dış kontroller sonucu uygun aday maddeler nihai madde olarak teste alınır.

Test maddeleri elde edildikten sonra test geliřtirici/geliřtiriciler tarafından test formu hazırlanır. Bu safhada maddelerin form içerisine yerleřtirilmesi ve yanıtlayıcıların hareket tarzlarını belirleyen madde ve/veya test bazında yönerge hazırlanır. Maddeler forma yerleřtirilirken, deneme uygulamasında tespit edilen güçlükleri dikkate alınarak kolaydan zora yerleřtirilmeli (Tan, 2009b), aynı köklü maddeler birlikte verilmeli, cevaplayıcı grubunun düzeyine uygun büyüklükte (punto) olmalıdır (Baykul, 2015, s.316).

Test formu elde edildikten sonra özelliđi ölçülecek gruba uygulanır. Uygulamada eđer standart belirlenmiş bir süre sınırlaması yok ise madde başına; matematik testlerinde 80-90 saniye, sosyal bilimler testlerinde 50-60 saniye, Türkçe testlerinde 45-55 saniye, yabancı dil testlerinde 45-55 saniye zaman verilebilir (Baykul, 2015, s.318). Madde yazarı madde geliřtirirken ortalama başarı düzeyindeki bir öğrencinin maddeyi cevaplayabileceđi süre olarak bu deđerleri esas alabilir.

Test uygulamasının ardından cevaplayıcıların yanıtları esas alınarak puanlama yapılır. Madde ve test psikometrik özellikleri belirlenir. Bu özelliklerle birlikte belirli

kodlamalar ile soru bankasına maddeler girilir. Turgut'un (2011) belirttiği gibi maddelere yap-kullan-at gözüyle bakmamak gerekmektedir. Sonraki dönemlerde benzer özelliklerin ölçülmesinde kullanmak ve kullandıkça güncellemek için soru bankasında ihtiyaç duyulan özelliklere sahip maddelerin kolaylıkla ulaşılabileceği düzenlemeler yapılmalıdır. Son olarak test geliştirmenin tüm adımlarını içeren bir test teknik kılavuzu hazırlanır ve madde/test ile birlikte basılı ya da bilgisayar ortamında arşivlenir.

1.7.3. Madde Yazarlığı Eğitimi

Dünyada ve ülkemizde, öğrenci başarılarını ölçme ve değerlendirme ile görevli olan başta öğretmenler ve madde yazarlarının bir kısmı, geliştirdikleri maddelerin yeterince iyi olmadığını düşünmektedirler (Baş ve Beyhan, 2016; Çakan, 2004; Stiggins & Bridgeford, 1985; Şenel, 2008). Bu görüşün aksine kendilerini yeterli görenler öğretmen ve madde yazarları da bulunmaktadır (Oescher & Kirby, 1990; Wise, Lukin & Roos, 1991). Fakat bu iki grubun da üzerinde anlaştıkları husus güçlü ölçme becerilerine sahip olmak istemeleri ve test geliştirme ile madde yazım kuralları konusunda eğitilmeleri gerektiğine inanmalarıdır (Aşık, 2009; Iramaneerat, 2012; Yapıcı ve Demirdelen 2007; Yıldız, 2011; Wise, Lukin & Roos, 1991).

Uluslararası test merkezleri ve alanyazında yapılan araştırmaların madde yazarlığı eğitim programları incelendiğinde Tablo 1.7'deki konuların esas alındığı söylenebilir (ACRP, 2017; ETS 2017; Haladyna & Rodrigues, 2013; Hingorjo & Jaleel, 2012; Ingham, 2008; Iramaneerat, 2012; ITEA, 2012). Tablo 1.7'de eğitiminin konu başlıkları, konu kapsamı, süreleri ve başarı ölçütleri sunulmuştur.

Tablo 1.7: Madde Yazarlığı Eğitim Programı

Konu Başlığı	Konu Kapsamı	Ortalama Süre (Dakika)	Başarı Ölçütleri
Kapsamın Tanımlanması	- Madde ile yoklanacak kazanımın sınırlarının belirlenmesi* - Kazanımın ön koşullarının belirlenmesi - Kazanıma uygun uyarının belirlenmesi	20	- Kazanımın sınırlarını belirler. - Kazanımın ön koşullarını belirler. - Kapsamı genişletmeden ve daraltmadan yoklayacak uyarıcıyı (madde kökünü) belirler.
Madde Tipleri	- Yapılandırılmış cevaplı maddeler - Seçmeli maddeler	20	- Madde tiplerini bilir. - Kazanıma uygun madde tipini belirleyebilir.
ÇSM'nin Yapısı, Avantaj ve Dezavantajları	- Çoktan seçmeli maddenin yapısı, öğeleri - Çoktan seçmeli maddelerin diğer madde tiplerine göre üstün ve zayıf yönleri	20	Çoktan seçmeli maddenin öğelerini, üstün ve zayıf yönlerini bilir.
Madde Yazım Kuralları	Tablo 1.8	90	Tablo 1.8'deki kuralları bilir.
Madde Geliştirme	Tablo 1.6	30	Tablo 1.6'i kullanarak çoktan seçmeli madde geliştirme sürecini işletebilir ve madde geliştirebilir.
Yazılmış Maddelerin İncelenmesi	Türkiye'de uygulanan geniş ölçekli testlerden alınan örnek maddelerin, madde yazım kurallarına uygunluğunun kontrolü	30	Bir maddenin madde yazım kurallarına uygun olup olmadığını belirleyebilir.
Madde Puanları Üzerine Analizler	- Madde güçlük indeksinin hesaplanması ve yorumlanması - Madde ayırıcılık gücü indeksinin hesaplanması ve yorumlanması (%27'lik gruplar üzerinden) - Seçenek analizi ve yorumlanması (%27'lik gruplar üzerinden)	30	- Madde güçlük ve ayırıcılık gücü indeksini hesaplayabilir ve ne anlama geldiğini yorumlayabilir. - Seçeneklere yönelen öğrenci sayılarından çeldiricilerin işlerliğini yorumlayabilir.

* Öğretmenlere seçkisiz olarak atanan kazanım esas alınmıştır.

1.7.4. Madde Yazım Kuralları

Alanyazında ÇSM yazılırken dikkat edilmesi gerekenlerin neler olması gerektiğine dair bugüne değin yapılan arařtırmalar ve yayınlar incelenmiřtir. Haladyna ve Downing'in (1989b) 46 referans kaynaktan 96 teorik ve deneysel arařtırmaya dayalı hazırladıđı 43 kural ile Frey, Petersen, Edwards, Pedrotti & Peyton'ın (2005) 20 sınıfıçı deđerlendirme kaynađına dayalı 40 kural esas alınarak; bu arařtırmada öğretmenlere verilen madde yazarlıđı eđitimi esnasında kullanılan madde yazım kuralları belirlenmiřtir. Ayrıca diđer arařtırma ve yayınların (Atılđan vd., 2009; Baykul, 2015; ETS, 2017; Haladyna, 2004; Haladyna & Rodriguez, 2013; Ingham, 2008; ITEA, 2012; Mullis & Martin, 2013; Nathan & Jennifer, 2015) öne sürdüđü kurallarla farklılıklar/benzerlikler irdelenmiř ve üzerine mutabık olunanlar tespit edilmiřtir. Bu kuralların bir kısmına çeřitli arařtırma bulguları ile kanıt sunulmuř, bir kısmı ise hâlihazırda teoriktir. Sınıf içi öğretmen yapımı testlerden, geniş ölçekli testlere kadar uygulanabilecek kurallar, Tablo 1.8'de kapsamlı, biçimle, madde köküyle, seřiteneklerle, çeldiricilerle, dođru yanıtla ilgili olmak üzere ayrı ayrı belirtilmiřtir. Tablo 1.8'de kuralın etkilendiđi psikometrik özellik kaynaklara dayanılarak; kuralın teorik mi yoksa arařtırmaya mı dayalı olduđu ve kuralın kaynađı ile birlikte sunulmuřtur.

Tablo 1.8: Çoktan Seçmeli Madde Yazım Kuralları

<i>Kural</i>	<i>Etkilenen Psikometrik Özellik</i>	<i>Dayanak</i>	<i>Kaynak</i>
<i>Kapsamla İlgili Kurallar</i>			
1. Madde tek bir kapsamı (kazanımı) ve bilişsel süreci yoklamalıdır.	Geçerlik	Teorik	Atılğan vd. (2009); ; Baykul (2015); Haladyna (2004); ITEA (2012)
2. Maddenin yokladığı konunun önemli olması gerekir.	Geçerlik, Güvenirlilik	Teorik	Atılğan vd. (2009); Frey vd. (2005); Haladyna (2004)
3. Üst düzey bilişsel süreçleri yoklayabilmek için yeni veya bilinmedik kalıp ve tasarımlar kullanılmalıdır.	Geçerlik	Teorik	Haladyna (2004); ITEA (2012)
4. Maddenin içeriği testteki diğer maddelerin içeriğinden bağımsız olmalıdır.	Geçerlik, Güvenirlilik	Teorik	Atılğan vd. (2009); Haladyna (2004)
5. Aşırı spesifik ve aşırı genel kapsamdan kaçınılmalıdır.	Geçerlik	Teorik	Baykul (2015); Haladyna (2004)
6. Yorumla açık maddelerden sakınılmalıdır.	Hakkaniyet	Teorik	Haladyna (2004); Nathan & Jennifer (2015)
7. Madde hileli olmamalıdır.	Geçerlik, Güvenirlilik	Araştırma	Haladyna (2004); ITEA (2012); Roberts (1993)
8. Maddenin bilimsel dayanağı bulunmalıdır.	Geçerlik	Teorik	ITEA (2012)
<i>Biçimle İlgili Kurallar</i>			
9. Madde mümkün olduğunca basit ifadelerle oluşturulmalıdır.	Geçerlik	Araştırma	Abedi, Lord, Hofstetter & Baker (2000); Haladyna (2004); ITEA (2012)
10. Maddenin okunma süresi minimize edilmelidir.	Geçerlik, Güvenirlilik	Teorik	Haladyna (2004)
11. Madde; dilbilgisi, noktalama işaretleri ve imlâ kuralları açısından uygun olmalıdır.	Geçerlik, Güvenirlilik	Araştırma	Atılğan vd. (2009); Frey vd. (2005); Haladyna (2004); Hansen & Dexter (1997); ITEA (2012)
12. Madde netlik için düzenlenmiş olmalıdır.	Güvenirlilik	Teorik	Haladyna (2004); ITEA (2012); Nathan & Jennifer (2015)
13. Madde teknik hatalara karşı denetlenmiş olmalıdır.	Geçerlik, Güvenirlilik	Teorik	Haladyna (2004); Nathan & Jennifer (2015)

Kural	Etkilenen Psikometrik Özellik	Dayanak	Kaynak
14. Madde test formuna dikey olarak yerleştirilmeli, cevaplayıcılar maddeyi aşağıya doğru takip edebilmelidir.	Güvenirlilik	Teorik	Haladyna (2004)
15. Tüm maddeler numaralandırılmalı, seçenekler harflendirilmelidir.	Güvenirlilik	Teorik	Frey vd. (2005)
16. Aynı formatlı maddeler birlikte bulunmalıdır.	Güvenirlilik	Teorik	Atılğan vd. (2009); Frey vd. (2005)
17. Maddenin nasıl yanıtlanacağına dair yönerge bulunmalıdır.	Geçerlik, Güvenirlilik	Teorik	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005)
18. Maddenin tüm parçaları aynı sayfada görünmelidir.	Güvenirlilik	Teorik	Atılğan vd. (2009); Frey vd. (2005)
19. Form temiz ve okunabilir olmalı, el yazısıyla hazırlanmamalıdır.	Güvenirlilik	Teorik	Atılğan vd. (2009); Frey vd. (2005)
Madde Köküyle İlgili Kurallar			
20. Madde kökündeki ifade mümkün olduğunca açık yazılmalıdır.	Güvenirlilik	Teorik	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005); Haladyna (2004); ITEA (2012); Hansen ve Dexter (1997)
21. Madde kökü mümkün mertebe kısa yazılmalıdır.	Geçerlik, Güvenirlilik	Teorik	Baykul (2015); Frey vd. (2005); Haladyna (2004); Hansen ve Dexter (1997); Nathan & Jennifer (2015)
22. Madde kökü olumlu ifade edilmelidir.	Güvenirlilik	Araştırma	Frey vd. (2005); Haladyna (2004); Hansen ve Dexter (1997); ITEA (2012); Rodriguez (2002); Tamir (1993)
23. Ana düşünce seçeneklerden ziyade madde kökünde bulunmalıdır.	Güvenirlilik	Teorik	Haladyna (2004); ITEA (2012)
24. Madde kökünde ilgisiz konulara yer verilmemelidir.	Geçerlik, Güvenirlilik	Teorik	Atılğan vd. (2009); Baykul (2015); Haladyna (2004); Nathan & Jennifer (2015)
Seçeneklerle İlgili Kurallar			
25. Çeldiriciler ve doğru yanıt maddeye rastgele ve değişen sıralarla yerleştirilmelidir.	Geçerlik, Güvenirlilik	Araştırma	Attali ve Bar-Hillel (2003); Frey vd. (2005); Haladyna (2004); Hohensinn ve Baghaei (2017); Nathan & Jennifer (2015)
26. Seçenekler, mümkünse mantıkî veya sayısal sıraya sokulmalıdır.	Geçerlik	Teorik	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005); Haladyna (2004); Nathan & Jennifer (2015)
27. Seçenekler birbirlerinden bağımsız olmalıdır.	Geçerlik	Teorik	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005); Haladyna (2004); ITEA (2012); Nathan & Jennifer (2015)

Kural	Etkilenen Psikometrik Özellik	Dayanak	Kaynak
28. İçerik ve dilbilgisi yönlerinden seçenekler madde köküyle ve birbirleriyle benzer olmalıdır.	Geçerlik, Güvenirlik	Araştırma	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005); Fuhrman (1996); Haladyna (2004); Hansen ve Dexter (1997)
29. Seçenekler uzunluk olarak birbirlerine yakın olmalı, madde kökünden uzun olmamalıdır.	Geçerlik	Teorik	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005); Haladyna (2004)
30. Hiçbiri seçeneği dikkatli kullanılmalıdır.*	Geçerlik	Araştırma	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005); Haladyna (2004); Hansen & Dexter (1997); ITEA (2012); Knowles & Welch (1992)
31. Hepsi seçeneği dikkatli kullanılmalıdır.*	Geçerlik	Araştırma	Atılğan vd. (2009); Baykul (2015); Frey vd. (2005); Haladyna (2004); Haladyna ve Downing (1989a); Hansen & Dexter (1997); ITEA (2012)
32. Değil, hariç gibi olumsuzluk belirten ifadeler kullanılmamalıdır.	Güvenirlik	Teorik	Atılğan vd. (2009); Haladyna (2004); ITEA (2012); Nathan & Jennifer (2015)
33. Seçeneklerde mizahi ifadeler kullanılmamalıdır.	Geçerlik	Araştırma	Haladyna (2004); McMorris, Boothroyd & Pietrangelo (1997)
Çeldiricilerle İlgili Kurallar			
34. Mümkün olan tüm çeldiriciler listelenip yalnızca çeldiricilikleri güçlü olanlar seçilmelidir.	Geçerlik	Araştırma	Baykul (2015); Frey vd. (2005); Haladyna (2004); ITEA (2012); Lord (1977)
35. Çeldiricilerde, doğru yanıtı ipucu olabilecek ifadeler kullanılmamalıdır.	Geçerlik, Güvenirlik	Teorik	Frey vd. (2005); Haladyna (2004); Hansen & Dexter (1997); ITEA (2012); Nathan & Jennifer (2015)
36. Çeldiriciler işlevsel olmalıdır.	Geçerlik	Teorik	Frey vd. (2005); Haladyna (2004); Hansen & Dexter (1997); ITEA (2012)
37. Çeldiriciler geliştirilirken öğrencilerin tipik hataları kullanılmalıdır.	Geçerlik	Teorik	Baykul (2015); Haladyna (2004)
Doğru Yanıtla İlgili Kurallar			
38. Doğru yanıtın doğruluğu kesin olmalıdır.	Geçerlik	Teorik	Baykul (2015); Hansen & Dexter (1997); ITEA (2012)
39. Tek bir doğru yanıt olmalıdır.	Geçerlik	Teorik	Frey vd. (2005); ITEA (2012)

* Bu kural konusunda uzmanlar arasında görüş birliği sağlanamamıştır. Dikkatli olmak kaydı ile kullanılabileceğini öne sürenler (Atılğan vd., 2009; Baykul, 2015; Knowles & Welch, 1992;) olduğu gibi kesinlikle kullanılmaması gerektiğini iddia edenler de mevcuttur (Frey vd., 2005; Haladyna ve Rodriguez, 2013).

1.7.5. Etki Büyüklüğü

Bilimsel arařtırmalarda nicel verilerin analizi için kullanılan istatistiksel testlerin temelinde bulunan ve bir karar verme sistematığı olan yokluk hipotezi anlamlılık testi (YHAT), kurulan yokluk hipotezinin kabul ya da reddedilmesinin karar vericisidir. YHAT söz konusu kararı verirken I.tip hata ve gözlenen bulgunun elde edilme olasılığını dikkate alır. Fakat arařtırmalarda esas amaç, arařtırma sorusundaki deęişkenler arasında bulunan ilişkiyi ve bu ilişkinin gücünü ortaya koymaktır. YHAT deęişkenler arası ilişkinin gücü hakkında bilgi vermemekle yetersiz kalmakta ve birçok arařtırmacı tarafından eleřtirilmektedir (Ahlbom, 2017; Cohen, 1992; Nickerson, 2000).

Deęişkenler arası ilişkinin ne derece güçlü olduğunu standart şekilde ortaya koymada, YHAT'ne ek olarak kullanılabilcek bir istatistik olan etki büyüklüğü, YHAT ile tanımlanan beklentiden sapma düzeyi ve anlamlı farklılıkların deęişkenler arası etkisini belirleyen standart bir deęerdir (Cohen, 1992). Alanyazında bulunan etki büyüklüğü hesaplama yöntemlerinden birisi de Cohen (1988) tarafından önerilen yöntemdir. Bu yöntemde farklı testler için farklı etki büyüklüğü hesaplamaları yapılır.

İstatistiksel sınamalarda etki büyüklüklerinin verilmesi, yapılacak yorumların daha güçlü olmasını sağlayacaktır (Işık, 2014). Bu sebeple arařtırmada anlamlı farklılık elde edilen test sonuçları için etki büyüklükleri hesaplanarak sunulmuştur. Ortalama güçlük, ayırıcılık gücü ve çeldiricilikler arasındaki etki büyüklüğü Cohen's h ; KR-20 deęerleri arasındaki etki büyüklüğü Cohen's q ; varyanslar ve aritmetik ortalamalar arasındaki etki büyüklüğü Cohen's d ile hesaplanmıştır. Etki büyüklüklerinin hesaplanmasında, R programına <https://github.com/heliosdrm/pwr> bağlantısında bulunan kodlar girilmiştir. Etki büyüklüklerinin mutlak deęerlerinin yorumlanmasında Tablo 1.9'daki ölçütler esas alınmıştır (Cohen, 1988;1992).

Tablo 1.9: Etki Büyüklüklerinin Yorumlanması

d, h	q	r	Yorum
0,00-0,19	0,00-0,09	0,00-0,09	Etki Yok
0,20-0,49	0,10-0,23	0,10-0,29	Küçük Etki
0,50-0,79	0,24-0,36	0,30-0,49	Orta Etki
0,80 ve daha büyük	0,37 ve daha büyük	0,50 ve daha büyük	Büyük Etki

2. İLGİLİ ARAŞTIRMALAR

Alanyazında madde yazarının yaratıcılığının geliştirdiği maddelerin psikometrik özelliklerine etkisine dair herhangi bir araştırmaya rastlanılmadığından, bu bölümde madde yazarlığı, eğitimi ve madde yazım kuralları ile ilgili araştırmalara yer verilmiştir.

Hansen & Dexter (1997) madde yazarlığı eğitiminin ne derecede yazım kurallarına uygun madde geliştirebilmeyi sağladığını inceledikleri araştırmalarında, madde yazarlığı eğitimi CPA (yeminli mali müşavirlik sınavı) yazarlarının geliştirdiği maddeler ile bu eğitimi almamış yazarların geliştirdikleri 10 soru bankasındaki maddelerden oluşan toplam 440 ÇSM'yi 17 madde yazım kuralı açısından incelemişlerdir. İnceleme sonucu toplam 490 madde yazım kuralı ihlali tespit etmişlerdir. İhlallerin eğitimi ve eğitimsiz madde yazarlarına göre farklılaşmasına yönelik yapılan ki-kare testi sonucu hatasız madde sayısının anlamlı bir şekilde madde yazarlığı konusunda eğitimi alan yazarlar lehine olduğu bulgusuna ulaşılmıştır ($\chi^2=4,67$; $p<0,05$). Bu bulguya dayanarak madde yazarlığı eğitimi alan yazarların eğitim almayanlara nazaran madde yazım kurallarına daha uygun madde geliştirdiklerini öne sürmüşlerdir.

Tarrant & Ware (2008) madde yazım kurallarına uygun geliştirilmemiş maddelerin öğrenci başarısına etkisini araştırmışlardır. Hemşirelik okulu öğrencilerine uygulanan 10 testten toplam 664 madde, 4 uzman tarafından madde yazım kurallarına uygunluk açısından incelenmiştir. En az bir madde yazım kuralını ihlal eden maddeler "kusurlu", hiçbir ihlal tespit edilmeyen maddeler ise "uygun" şeklinde tasnif edilmiştir. Maddelerin 826 öğrenciye uygulanması sonucu yapılan madde analizi neticesinde kusurlu maddelerin güçlükleri 0,60-0,76; uygun maddelerin güçlükleri ise 0,52-0,73 arasında hesaplanmıştır. Kusurlu maddelerin ayırıcılık güçleri 0,18-0,27; uygun maddelerinki ise 0,19-0,33 aralığında hesaplanmıştır. Öğrencilerden yüksek puan alanlar ile sınır puanda bulunanların durumları incelendiğinde; başarılı olanların, sınır başarısındaki öğrencilere nazaran kusurlu maddelerden daha çok etkilendiği bulgusuna ulaşılmıştır. Bulgulara dayanarak madde yazım kurallarına uygun geliştirilmemiş maddelerin öğrenci başarısını ölçmede olumsuz etkide bulunduğunu, bu nedenle maddeyi geliştirecek olan uzmanların madde yazım kuralları konusunda eğitilmelerini önermişlerdir.

Fulkerson, Mittelholtz & Nichols (2009;2010) tecrübeli ve tecrübesiz madde yazarlarının madde geliřtirmede kullandıkları biliřsel süreçleri incelemiřlerdir. Madde geliřtirmeyi bir problem olarak tanımlamıřlar ve bu problemin üstesinden gelmek için madde yazarlarının nasıl bir yol izlediklerini ve tecrübenin madde geliřtirme sürecine etkisini tespit etmek amacıyla yaptıkları peřin sıra iki arařtırmada, sözel raporlar ve protokol analiz tekniklerini kullanarak madde yazarlarının madde geliřtirirken biliřsel süreç iřlemlerini analiz etmiřlerdir. Yapılan analizler sonucu, tecrübeli madde yazarlarının madde geliřtirmede temsil, keřif ve çözüm olmak üzere üç ařamalı problem çözüm becerilerini başarıyla sergiledikleri bulgusuna ulařmıřlardır. Ayrıca bir duruma dayalı madde geliřtirmek üzere farklı tecrübe düzeylerindeki 3 madde yazarının madde geliřtirme süreçleri aynı tekniklerle analiz edilmiřtir. Sonuç olarak madde yazarlarının maddeyi geliřtirirken karřılařtıkları sorunları geçmiř tecrübeleri ve sahip oldukları özel bilgilerle ařabildiklerini tespit etmiřlerdir. Arařtırma bulgularına dayanarak tecrübeli madde yazarlarının tecrübesizlere nazaran üst biliřsel süreçlerde ölçme yapabilecek uygun nitelikte maddeler geliřtirebildiklerini, tecrübesiz madde yazarlarının tecrübeli madde yazarları ile birlikte çalıřmaları ve madde yazarlığı eđitimi almaları gerektiđini öne sürmüřlerdir.

Hawkey & Green (2012) tecrübeli ve tecrübesiz madde yazarlarının, kazanım esaslı bir okuma testi için geliřtirdikleri maddeleri ve bu maddeleri geliřtirme süreçlerini incelemiřlerdir. Madde yazarlarının bireysel ve grup halinde madde geliřtirdikleri arařtırmada; akıř řemaları, mülâkat ve elde edilen maddelerin uzmanlarca nitel karřılařtırmaları yapılmıřtır. Arařtırma sonunda tecrübeli madde yazarlarının IELTS üst basamak biliřsel süreçleri yoklayan maddeleri geliřtirme süreçlerinin daha belirgin ve geliřtirdikleri maddelerin daha iyi seviyede olduđu bulgusuna ulařmıřlardır. Tecrübesiz madde yazarlarının, tecrübeli yazarlarca eđitilmeleri gerektiđini, birlikte madde geliřtirmelerinin tecrübe kazanmalarına ve daha iyi düzeyde madde geliřtirmelerine katkı sađlayacađını öne sürmüřlerdir.

Iramaneerat'in (2012) madde yazarlığı eđitiminin madde güçlüğü ve ayırıcılık gücüne etkisini arařtırdığı çalıřmasında, öđretmenlere uygulanan madde yazarlığı eđitimi sonucunda, eđitime katılanlar ile katılmayanların, eđitim öncesi ve sonrası geliřtirdikleri maddelerin özellikleri karřılařtırılmıřtır. İlk eđitimde üç saat süre ile madde geliřtirme prensipleri ve yaygın hatalar, sonraki eđitimde iki saat süre ile

madde analizi konularında katılımcılarla çalışma yapılmıştır. Eğitime gönüllü olarak katılan öğretmenlerin eğitim öncesi geliştirdikleri maddelerin güçlüklerinin anlamlı derecede istenilmeyen düzeyde ya çok yüksek ya da çok düşük olduğu, eğitime katılmayan öğretmenlerin geliştirdikleri maddelerle kıyaslandığında ise benzer madde ayırıcılık gücünde oldukları görülmüştür. Eğitim sonunda eğitime katılan öğretmenlerin geliştirdikleri maddelerin güçlüklerinin, eğitime katılmayanların geliştirdikleri maddelerle aynı düzeyde olduğu, madde ayırıcılık güçlerinin ise eğitime katılmayanlara göre anlamlı derecede daha yüksek düzeyde olduğu tespit edilmiştir. Elde edilen bulgular değerlendirildiğinde, madde yazarlığı eğitiminin maddelerin güçlük ve ayırıcılık güçlerine olumlu şekilde etki ettiği yorumu yapılmıştır.

Pate & Caldwell (2014) madde yazım kurallarına uygun geliştirilmiş maddelerin psikometrik özelliklerini inceledikleri araştırmalarında, 187 ÇSM'yi Haladyna, Downing & Rodriguez (2002)'nin belirlediği madde yazım kurallarına uygunluk açısından tasnif etmişler ve madde psikometrik özellikleri bakımından kıyaslamışlardır. Maddelerde 31 madde yazım kuralından 17'sinin ihlal edildiğini tespit etmişlerdir. İhlal tespit edilen 97 madde ile madde yazım kurallarına uygun geliştirilmiş 90 maddenin ortalama madde güçlük indeksleri ile ortalama madde ayırıcılık gücü indeksleri karşılaştırılmıştır. Analiz sonucunda madde yazım kurallarına uygun geliştirilmiş maddelerin ortalama güçlüğü ($\bar{p}_l = 83,7$), ihlalli maddelerin ortalama güçlüğüne ($\bar{p}_l = 76,3$) göre madde yazım kurallarına uygun geliştirilmiş maddeler lehine anlamlı derecede ($p < 0,05$) farklı bulunmuştur. Madde yazım kurallarına uygun geliştirilen maddelerin ortalama madde ayırıcılık gücü ($\bar{r}_{jx} = 0,242$) ile ihlalli maddelerin ortalama ayırıcılık gücü ($\bar{r}_{jx} = 0,255$) arasında anlamlı farklılık görülmemiştir ($p > 0,05$). Bulgulara göre, maddelerin madde yazım kurallarına uygun geliştirilmesi maddelerin madde güçlüklerine olumlu etki yaparken, madde ayırıcılık güçlerine anlamlı bir etkisi bulunmamıştır.

Abdulghani vd. (2015) tıp fakültesi öğretim üyelerine ÇSM yazma konusunda 2 gün süreli eğitim vermişler ve bu eğitimin, öğretim üyelerinin geliştirdiği maddelerin psikometrik özelliklerine etkilerini, eğitim öncesi ve sonrası geliştirilen maddelerin istatistiklerini karşılaştırarak tespit etmişlerdir. Eğitim sonrası geliştirilen maddelerin güçlüklerinin anlamlı şekilde düştüğü ($p = 0,003$); ayırıcılık güçlerinin anlamlı olmayan şekilde ($p = 0,132$); üst bilişsel süreçlere yönelik ölçme yapmak

amacıyla geliştirilen madde sayılarının ($p < 0,0001$); madde yazım kurallarına uygun madde sayısının ve işler çeldirici (%5 ve fazla) sayılarının artış ($p < 0,02$) gösterdiği bulgularına ulaşmışlardır. Bu bulgular ışığında madde geliştirmekle görevli öğretim üyelerine madde yazarlığı eğitimi verilmesinin uygun olacağı önerisinde bulunmuşlardır.

Vinay, Sunil, Praveen & Yuvaraj (2016) madde yazarlığı eğitiminin ÇSM geliştirmeye etkisini inceledikleri araştırmalarında, önceden madde yazarlığı eğitimi almamış 8 öğretim üyesinden biyokimya dersi için madde geliştirmeleri istenmiştir. 3 alan uzmanınca madde yazarlığı eğitimi verilmesini müteakip aynı grup tekrar madde geliştirmiştir. Eğitim öncesi ve sonrası geliştirilen maddeler, alan uzmanlarınca belirlenmiş 60 maddelik madde yazım kuralları kontrol formu ile değerlendirilmiştir. Eğitim öncesi ve sonrası geliştirilen maddelerin kontrol formuna uygunluk durumlarına yönelik yapılan ki-kare testi sonucu, maddelerin ÇSM tasarımına uygunluğu ve kalitesi ($p < 0,05$) eğitim sonrası lehine istatistiksel olarak anlamlı çıkmıştır. Araştırma bulgularına göre eğitim programından sonra ÇSM kalitesi konusunda önemli bir iyileşme kaydedildiği ve buna dayanarak kariyerinin başlarında öğretim üyelerine madde yazarlığı eğitiminin verilmesi gerektiği önerilmiştir.

Tenzin, Dorji & Tenzin (2017) madde yazarlığı eğitiminin madde psikometrik özelliklerine etkisini inceledikleri araştırmalarında, önceden madde yazarlığı eğitimi almamış farklı alanlardan 4 eğitim fakültesi öğretim elemanlarından oluşan gruptan 8 ÇSM geliştirmelerini istemişlerdir. Geliştirilen maddeler rastgele seçilen 16 kişilik stajyer doktor grubuna uygulanmıştır. Madde yazarlarına 2 saat süreyle madde geliştirme konusunda verilen eğitim sonrası aynı ÇSM'ler madde geliştirme kurallarına uygun hale getirilerek aynı gruba tekrar uygulanmıştır. İki uygulamadan elde edilen puanlara yapılan analizler sonucu; başarı yüzdesinin % 69,8 (11)'den % 81,3 (13)'e, ortalama madde güçlüğü'nün 0,51'den 0,53'e, ortalama madde ayırıcılık gücünün 0,47'den 0,59'a ve %5 ve daha üstü tercih edilen çeldirici oranının %74,9'dan %83,3'e çıktığı bulgusuna ulaşmışlardır. Bulgular ışığında madde yazarlığı eğitiminin ÇSM'lerin psikometrik özelliklerini geliştirici bir etkiye sahip olduğunu öne sürmüşlerdir.

İlgili arařtırmalar incelendiğinde;

- Madde yazarlıđı eđitiminin, madde yazım kurallarına uygun madde geliřtirebilmeye olumlu katkı sađladıđı,
- Madde yazım kurallarına uygun geliřtirilmiř maddelerin, ođrenci bařarısını ołçmede daha bařarılı; madde gůçlůklerinin daha istendik ve eldiricilerinin daha iřler dűzeyde olduđu,
- Madde yazarlıđı konusunda tecrűbeli madde yazarlarının tecrűbesizlere nazaran daha iyi dűzeyde ołçme yapabilen maddeler geliřtirdiđi,
- Madde yazarlıđı eđitiminin madde gůçlűđüne olumlu yűnde katkı sađladıđı, madde ayırıcılık gűcűne ise olumlu yűnde katkı sađlayıcı etkisinin olduđu fakat farklı alanlarda bu etkinin de farklılařabileceđi gűrűlmektedir.

Bu arařtırma ile alanyazında yapılan ۆnceki arařtırmalarda olduđu gibi madde yazarlıđı eđitiminin geliřtirilen maddelerin psikometrik ۆzelliklerine etkisi incelenmiřtir. Sűz konusu arařtırmalardan farklı olarak ise;

- Yaratıcılıđın madde psikometrik ۆzelliklerine etkisi incelenmiř,
- Tek bir alan yerine iki farklı alanda madde yazarlıđının madde psikometrik ۆzelliklerine etkisi incelenmiř ve karřılařtırılmıř,
- Madde yazarlıđı eđitiminin madde psikometrik ۆzelliklerine etkisinin ne oranda olduđu analiz edilmiř,
- Madde eldiricilerinin iřlerliđi tek bir indeks ile karřılařtırılmıř,
- Madde yazım kuralları tek bir kaynak yerine alan taranarak ortaya konmuř,
- Yalnız madde psikometrik ۆzellikleri deđil bu maddelerle oluřturulan test psikometrik ۆzellikleri de incelenmiřtir.

3. YÖNTEM

Bu bölümde araştırmanın yöntemi, çalışma grubu ve özellikleri, veri toplama araçları ve bu araçların uygulanışı, verilerin işlenmesi ve çözümlenmesi, araştırmanın iç ve dış geçerliği konu başlıklarına yer verilmiştir.

3.1. Araştırmanın Yöntemi

Bu çalışma, bağımsız değişkenlerin (yaratıcılık ve madde yazarlığı eğitimi) bağımlı değişken (madde ve test psikometrik özellikleri) üzerindeki etkilerini test etmeye yönelik bir araştırma olduğundan nicel araştırma türlerinden yarı deneysel araştırma kapsamına girmektedir. Araştırmada gerçek deneysel desenlerin özelliklerinden olan deney ve kontrol gruplarının oluşturulması ile seçkisizlik kurallarına uyulmuştur. Fakat özellikle sosyal bilimlerde bağımlı değişken üzerine etki eden bağımsız değişkenlerin tümünün tam anlamıyla belirlenememesi ve araştırmaya dahil edilememesi nedeniyle yarı deneysel tanımlamasının uygun olacağı düşünülmüştür. Araştırmacı tarafından oluşturulan farkların bağımlı değişken üzerindeki etkisini sınımaya yönelik çalışmalar olan deneysel araştırmalarda temel amaç, değişkenler arasında oluşturulan neden sonuç ilişkisini test etmektir (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2016, s.195). Deneysel desenlerde, bağımsız değişken değişimlemesine tâbi tutulacak gruplarla işleme kontrollü olarak alınmayacak gruplar seçkisiz atanır. Bu süreç işletilirken dışsal değişkenler kontrol altına alınır (Cash, Stankovic & Storga, 2016, s.6). Deneysel araştırmaların hemen hemen tüm araştırmacıların üzerinde hemfikir oldukları dört temel özelliği bulunmaktadır (Büyüköztürk vd., 2016, s.196).

Grupların karşılaştırılması: Deney grubu bağımlı değişkende etkisi sınanacak işleme tâbi tutulurken, kontrol grubu ya bu işleme hiç tâbi tutulmaz ya da farklı bir işleme tâbi tutulur. Sonuçta, işlemlerin etkisi gruplar arasında kıyaslanarak, başlangıçta tanımlanan araştırma sorusuna yanıt aranır. Bu çalışmada madde yazarlığı eğitiminin madde ve test psikometrik özelliklerine etkisi araştırıldığından, deney grubuna madde yazarlığı eğitimi verilirken, kontrol grubuna bu işlem uygulanmayarak, iki grubun geliştirdikleri maddelerin psikometrik özellikleri incelenerek araştırma sorusu yanıtlanmıştır.

Bağımsız Değişkenin Manipüle Edilmesi: Araştırmacının kasıtlı olarak deney ve kontrol grupları üzerine bağımsız değişkeni uygulama biçimini modellemesi ve

sunu olarak bağımlı deęişkende yaratacağı etkiyi gözlemlemesidir. Bu çalışmada deneklere madde yazarlığı eğitimi verilmiş, kazanılan bilgi ve deneyim ile bağımsız deęişkendeki deęişim gözlemlenerek, deęişimin manipüle edilen bağımsız deęişkenden kaynaklandığı ileri sürülmüştür.

Seçkisizlik: Deneklerin gruplara seçkisiz, rastgele, kasıtlı olmayacak şekilde atanmasıdır. Ankara ilinde görev yapmakta olan tüm 8. sınıf matematik ve Türkçe öğretmenleri (MEB, 2016a) numaralandırılmış, bilgisayara rastgele sayı seçtirilerek elde edilen sayı karşılığı öğretmenlerle çalışma grubu oluşturulmuştur. Deney ve kontrol gruplarının da belirlenmesinde aynı yöntem kullanılmıştır. Araştırmada öğretmenler kontrol ve deney gruplarına, seçkisiz atanarak, deneysel araştırmanın temel kuralına uyulmuştur.

Dışsal Deęişkenlerin Kontrolü: Deneysel araştırmalarda amaç, bağımlı deęişkendeki deęişimin bağımsız deęişkenden kaynaklandığını ileri sürmek olduğundan, başka deęişkenlerin bağımlı deęişkende deęişim yaratmadığından emin olmak gerekmektedir. Yoksa gruplar arası farkların yalnız uygulanan işlemde kaynaklandığını iddia etmek uygun olmayacaktır. Deneysel çalışmalarda, anılan dışsal deęişkenlerin etkileri; seçkisizlik, dışsal deęişkeni sabit tutmak, dışsal deęişkeni çalışmaya dâhil etmek, eşleştirme ve kovaryans analizi gibi yöntemlerle minimize edilebilir (Fraenkel, Wallen & Hyun, 2011, s.236). Bu araştırmada, dışsal deęişken olarak; madde yazarlığında eğitilmiş olma, TYDT tecrübesi, geniş ölçekli testlerde madde yazarlığı tecrübesi ve meslekî kıdem belirlenmiştir. Deneklerin önceden madde yazarlığı eğitimi almamış olması, TYDT'ne gerek katılımcı gerekse puanlayıcı olarak katılmamış olmaları ve herhangi bir bölgesel ya da ülke genelinde geniş ölçekli bir testte madde yazarlığı yapmamış olmalarının beyanlarınca tespiti ile araştırmanın, ilk üç dışsal deęişkenin etkisine maruz kalmadığı anlaşılmıştır. Meslekî kıdem dışsal deęişkeni ise üzerine madde geliştirilen alana hâkimiyet ve öğrencilerin düşünce yapılarını bilme konusundaki tecrübe düzeylerini etkilediğinden, araştırmada etkisini minimize edebilmek amacıyla, seçkisizlik ve istatistikî olarak kontrol altına alacak yöntemler kullanılmıştır. Araştırma süreci Şekil 3.1'de verilmiştir.

Safha	Çalışma Grubu	İşlem	Grup	İşlem	Elde Edilen
BİRİNCİ SAFHA	8. Sınıf Matematik Öğretmenleri (n=40)	Yaratıcılığın ölçülmesi	Yaratıcılığı Yüksek grup (20)	(1) Her öğretmene bir kazanım atanması	Form I
			Yaratıcılığı Düşük grup (20)	(2) Öğretmenlerin kazanım için madde geliştirmesi	Form II
İKİNCİ SAFHA	8. Sınıf Türkçe Öğretmenleri (n=40)	Seçkisiz	Deney Grubu (20)	(1) Madde Yazarlığı Eğitimi	Form III
			Form V*	(2) Başlangıçta atanan kazanım için bir maddenin daha geliştirilmesi	
			Kontrol Grubu (20)	-	Form IV*

* Bu formlar için yeniden madde geliştirilmemiştir. Deney ve Kontrol grubuna seçkisiz atanan öğretmenlerin, araştırmanın birinci safhasında kendilerine verilen kazanımlara dönük geliştirdikleri maddelerle oluşturulmuştur.

VERİLERİN ELDE EDİLMESİ

Veri Toplama Aracı	Elde Ediliş	Kime Uygulandı	Kim Uyguladı	Uygulama Yeri	Elde Edilen Veri
TYDT	Mevcut TYDT Türkçe Versiyonu A Formu	Öğretmenlere (40 Matematik ve 40 Türkçe)	Araştırmacı	Öğretmenler Odası	Yaratıcılık İndeksi
MSEF	Öğretmenlerin Geliştirdiği Maddeler	8.Sınıf Öğrencisine (n=1313)	Araştırmacı ve Öğrencilerin Öğretmenleri	Derslikler	(1) Madde ve Test Puanları (2) MOAS Puanları
MUF	Araştırmacı Tarafından	Geliştirilen Maddelerle Oluşturulan Formlara (Form I, II, III)	Alan Uzmanları (14 Matematik 12 Türkçe)	Çalışma Odaları	Maddelerin KGO ve KGI'leri

Şekil 3.1 Araştırma Süreci

3.2. Çalışma Grubu

Evren seçiminde, araştırma sahası olarak araştırmanın yapılabilirliği adına araştırmacının eğitim gördüğü Ankara ili belirlenmiştir. Araştırma, iki farklı evrenden, Ankara İl Millî Eğitim Müdürlüğü'nden alınan izinle (EK 3) örneklem ve çalışma grubu alınmak suretiyle yapılmıştır. Araştırmanın birinci evreni 2016-2017 eğitim ve öğretim yılında Ankara İl Millî Eğitim Müdürlüğü devlet okullarında görev yapan tüm 8.sınıf matematik (n=1082) ve Türkçe (n=2364) öğretmenleridir. İkincisi ise 2016-2017 eğitim ve öğretim yılında Ankara ilindeki 8.sınıf 1.dönem Merkezî Ortak Asıl Sınavına (MOAS) katılmış engelli olmayan (n=70.408) 8.sınıf öğrencileridir.

Çalışma grubunun örneklem büyüklüğünün belirlenmesinde;

- ortak bir kapsama dayalı madde geliştirebilmek,
- uyum geçerliğine kanıt olarak sunulabilecek geniş ölçekli bir test puanını kullanabilmek,
- kazanım esaslı bir ölçme aracı ile araştırmayı sürdürebilmek için TEOG sistemi içerisindeki Merkezî Ortak Sınavı (MOS) alt testlerinin madde sayısı (n=20) esas alınmıştır. Yaratıcılığın etkisinin belirlenmesi için, iki alanda, alt test madde sayılarının iki katı kadar büyüklükte çalışma grubu oluşturulmuştur. Araştırmanın çalışma grubunu Ankara İl Millî Eğitim Müdürlüğü 2016-2017 eğitim ve öğretim yılında, devlet okullarında görev yapan 8.sınıf matematik ve Türkçe öğretmenlerinden (MEB, 2016a) seçkisiz alınan 40 matematik, 40 Türkçe toplam 80 öğretmen oluşturmaktadır. 8.sınıf öğrencilerinin sayısının belirlenmesinde sürekli değişkenlerde tahmin için Büyüköztürk vd. (2016, s.95) tarafından önerilen aşağıdaki formül kullanılmıştır.

$$n = \frac{n_o}{1 + \frac{n_o}{N}} \quad (\text{eşitlik 7})$$

$$n_o = [(t \times S)/d]^2$$

n : Örneklem büyüklüğü

t : Güven düzeyine karşılık gelen t değeri

S : Standart sapma

d : Sapma miktarı (örnekleme hatası)

N : Evren büyüklüğü

70.408 öğrenci için, form bazında alınacak puanlar temelinde ortalamanın tahmini için sapma (d) miktarı 0,05 puan ve S=0,4 puan olarak tahminle, güven düzeyi (1-α)=0,95 alındığında (t=1,96);

$$n_o = [(1,96 \times 0,4)/0,05]^2 = 246$$

$$n = \frac{246}{1 + \frac{246}{70408}} = 245$$

elde edilmiştir. Toplam 6 form için 245 x 6 = 1470 öğrenci hesaplanmıştır. Bu nedenle, araştırmanın ikinci çalışma grubunu ise 2016-2017 eğitim ve öğretim yılında Ankara ilindeki 8.sınıf 1.dönem MOAS'na katılmış 8.sınıf öğrencilerinden okul bazında rastgele seçilmiş toplam 1470 öğrenci oluşturmaktadır.

3.2.1. Çalışma Grubunun Özellikleri

Araştırmanın çalışma grubunda bulunan öğretmenlerin tümünün ilgili alan mezunu ve branşlarında görev yapmakta oldukları tespit edilmiştir. Önceden bu çalışmada belirlenen seviyede madde yazarlığı eğitimi almadıkları, TYDT'ne gerek cevaplayıcı gerekse puanlayıcı olarak katılmadıkları, herhangi bir geniş ölçekli testte madde yazarlığı yapmadıkları beyanlarınca tespit edilmiştir. Araştırmada öğretmenlerce geliştirilen maddelerden oluşan testleri cevaplayan öğrencilerin demografik bilgileri alınmamıştır. Öğrenciler, geliştirilen madde ve test puanlarının elde edildiği, MOS alt test puan bilgilerinin beyanlarınca alındığı grup olarak çalışmada yer almışlardır.

3.2.2. Çalışma Grubunu Oluşturan Öğretmenlere İlişkin Demografik Bilgiler

Araştırmaya katılan öğretmenlerin cinsiyete, yaşa ve meslekî kıdeme göre dağılımları sıralı başlıklarda sunulmuştur.

3.2.2.1. Çalışma Grubunu Oluşturan Öğretmenlerin Cinsiyete Göre Dağılımları

Araştırmaya katılan kadın (n=41) ve erkek (n=39) öğretmenlerin cinsiyete göre dağılımları; branşları, araştırma gruplarındaki mevcutları ile birlikte Tablo 3.1'de gösterilmiştir.

Tablo 3.1: Çalışma Grubunu Oluşturan Öğretmenlerin Cinsiyete Göre Dağılımları

<i>Branşı</i>	<i>Cinsiyet</i>	<i>Deney Grubu</i>		<i>Kontrol Grubu</i>		<i>Toplam</i>	
		<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>
Matematik	Kadın	11	55	7	35	18	45
	Erkek	9	45	13	65	22	55
	Toplam	20	100	20	100	40	100
Türkçe	Kadın	10	50	13	65	23	58
	Erkek	10	50	7	35	17	42
	Toplam	20	100	20	100	40	100
Toplam		40	50	40	50	80	100

3.2.2.2. Çalışma Grubunu Oluşturan Öğretmenlerin Yaşa Göre Dağılımları

Araştırmaya 21-30 yaş aralığında (n=13), 31-40 yaş aralığında (n=25), 41-50 yaş aralığında (n=37) ve 51-55 yaş aralığında (n=5) katılan öğretmenlerin yaş aralıklarına göre dağılımları; branşları, araştırma gruplarındaki mevcutları ile birlikte Tablo 3.2’de gösterilmiştir.

Tablo 3.2: Çalışma Grubunu Oluşturan Öğretmenlerin Yaş Aralıklarına Göre Dağılımları

<i>Branşı</i>	<i>Yaş Aralığı</i>	<i>Deney Grubu</i>		<i>Kontrol Grubu</i>		<i>Toplam</i>	
		<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>
Matematik	21-30	4	20	6	30	10	25
	31-40	4	20	10	50	14	35
	41-50	12	60	4	20	16	40
	Toplam	20	100	20	100	40	100
Türkçe	21-30	2	10	1	5	3	8
	31-40	5	25	7	35	12	30
	41-50	10	50	10	50	20	50
	51-55	3	15	2	10	5	12
	Toplam	20	100	20	100	40	100
Toplam		40	50	40	50	80	100

3.2.2.3. Çalışma Grubunu Oluşturan Öğretmenlerin Meslekî Kıdeme Göre Dağılımları

Araştırmaya 0-5 yıl meslekî kıdem aralığında (n=13), 6-10 yıl meslekî kıdem aralığında (n=14), 11-15 yıl meslekî kıdem aralığında (n=17), 16 yıl ve üstü meslekî kıdem aralığında (n=36) katılan öğretmenlerin (n=80) meslekî kıdeme

göre dağılımları; branşları, araştırma gruplarındaki mevcutları ile birlikte Tablo 3.3'de gösterilmiştir.

Tablo 3.3: Çalışma Grubunu Oluşturan Öğretmenlerin Meslekî Kıdem Aralıklarına Göre Dağılımları

Branşı	Mesleki Kıdem	Deney Grubu		Kontrol Grubu		Toplam	
		f	%	f	%	f	%
Matematik	0-5	4	20	6	30	10	25
	6-10	4	20	4	20	8	20
	11-15	2	10	6	30	8	20
	16 ve üstü	10	50	4	20	14	35
	Toplam	20	100	20	100	40	100
Türkçe	0-5	2	10	1	5	3	8
	6-10	2	10	3	15	5	13
	11-15	4	20	6	30	10	25
	16 ve üstü	12	60	10	50	22	54
	Toplam	20	100	20	100	40	100
Toplam		40	100	40	100	80	100

3.3. Veri Toplama Araçları

Bu araştırmada veri toplamak amacıyla üç ölçme aracı kullanılmıştır. İlki öğretmenlerin yaratıcılıklarını belirlemek amacıyla kullanılan TYDT'dir. İkincisi öğretmenlerce geliştirilen maddelerden oluşan ve öğrencilere uygulanarak madde-test puanları ve öğrencilerden MOAS puanları elde edilen Merkezî Sınav Eş Formları'dır (MSEF). Üçüncüsü ise geliştirilen testlerin kapsam geçerliğini sınamak için, araştırmacı tarafından geliştirilen MUF'dur.

3.3.1. Torrance Yaratıcı Düşünme Testi

Künye

TYDT, Guilford'un (1956) akıcılık, esneklik, orijinallik ve zenginleştirme alt boyutlarıyla tanımladığı iraksak düşünme kavramı esas alınarak (Krumm, Lemos, & Filippetti, 2014), yaratıcılığı bir bütün olarak ölçmek amacıyla ilk kez 1966 yılında Torrance tarafından geliştirilmiştir (Torrance, 1966).

Ölçtüğü Yapı

Yaratıcılığı ölçmeyi amaçlayan test bataryası, üç temel varsayım esas alınarak geliştirilmiştir;

1. Yaratıcılık doğal bir yetenektir,

2. Yaratıcı birey olgulardaki uyumsuzluk ve eksikliklere karşı duyarlıdır,
3. Birey bilinçli veya bilinçsiz olarak yeni çözümler üreterek, problem çözer (Torrance, 1974, s.8).

Test bataryası ile bireyin hem yaratıcılık tarzı hem de yaratıcılık potansiyeli ölçülmektedir (Cheng, Kim & Hull, 2010). Torrance, geliştirdiği testin;

1. İnsan zekâsı, işlevi ve gelişimi üzerine yapılan temel araştırmalarda,
2. Bireysel eğitim için etkili temelleri keşfetmeyi amaçlayan çalışmalarda,
3. Psikoterapi ve düzeltici programlar için ipucu toplamada,
4. Çeşitli deneysel programların, yeni geliştirilen eğitim programı ve materyallerin, eğitim politikalarının, öğretme tekniklerinin etkilerini belirlemede,
5. Yaratıcı potansiyelin fark edilmesinde kullanılabileceğini önermiştir (Torrance, 1974, s.5).

Testin Yapısı

TYDT, 7 maddeli sözel alt test ve 3 maddeli şekilsel alt test olmak üzere toplam 10 faaliyetten (maddeden) oluşmaktadır.

Sözel alt test, formlardaki sıralamayla;

- soru sorma,
- nedenlerini tahmin etme,
- sonuçları tahmin etme,
- ürün geliştirme,
- alışılmadık kullanımlar,
- alışılmadık sorular,
- farz edin ki;

Şekilsel alt test ise sırasıyla;

- resim oluşturma,
- resim tamamlama,
- paralel çizgiler/daireler

maddelerinden meydana gelmektedir. Test bataryasının A ve B olmak üzere iki eş formu bulunmaktadır.

Uygulama

TYDT'nde bulunan maddeler açık uçlu yanıtlanmaktadır. Sözel alt test metinsel tepki gerektirirken, şekilsel alt test başlıklandırma haricinde çizim yapmayı gerektirir. Bireysel veya grup halinde uygulanabilen testin süresi sözel bölüm için 45, şekilsel bölüm için ise 30 dakika olarak önerilmektedir (Torrance, 1974, s.4).

Puanlama

Sözel alt testte bulunan norm dayanaklı ölçütlerden akıcılık, ilgili ve yorumlanabilir cevapların sayısıdır. Orijinallik cevabın alışılmış dışında, sık görülmeyen özellikte olmasına göre puanlanmaktadır. Esneklik, cevabın, farklı bakış açılarından ifade edilebilmesinin değerlendirmesidir.

Sözel alt testten farklı olarak, şekilsel alt testte bulunan norm dayanaklı ölçütlerden başlıkların soyutluğu, üretme yeteneğinde sentez ve organizasyonun; zenginleştirme, detayla ifade ve hayal edebilmenin; erken kapamaya direnç, özgün fikir üretmeyi sağlayacak zihnin açık tutulabilirliğinin ölçütleri olarak puanlandırılırlar. Kriter dayanaklı ölçütler ise norm dayanaklı ölçütlerden farklı olarak, figürlerdeki ipuçlarından yola çıkılarak puanlanır ve bonus puanı adı altında toplam puanla hesaplanır. Test bataryasının puanlanması, en az testin uygulanması kadar süre almakta ve özel bir eğitim gerektirmektedir. Bu sebeple detaylı bir puanlama rehberi oluşturulmuştur (Torrance, 1974; Torrance & Ball, 1984).

Test bataryasının güncel versiyonunun puanlanmasında, norm dayanaklı ölçütler ile kriter dayanaklı ölçütler kullanılmaktadır. Sözel alt test için akıcılık, esneklik ve orijinallik ölçütleri puanlanmaktadır. Şekilsel alt test ise; akıcılık, orijinallik, başlıkların soyutluğu, zenginleştirme, erken kapamaya direnç norm dayanaklı ölçütleri ile 13 kriter dayanaklı ölçüt (duygusal ifadeler, hikaye anlatma, hareket/faaliyet, başlıkların açıklayıcılığı, tamamlanmamış şekillerin birleştirilmesi, çizgilerin/dairelerin sentezi, alışılmadık görselleştirme, içsel görselleştirme, sınırları uzatma/geçme, mizah, hayalgücü zenginliği, hayalgücünün renkliliği, fantezi) esas alınarak puanlanmaktadır.

Testten elde edilen puanlar sözel ve şekilsel alt testler için ayrı ayrı kullanılmaktadır. Orijinal testin 37.814 denek üzerinden elde edilmiş norm çalışması ile belirlenmiş ham puanların yaş grupları için standart puanlara dönüşüm tabloları mevcuttur. Torrance (1974, s.56), tavsiye etmemekle birlikte testten tek bir yaratıcılık puanı elde edilmek istendiğinde, ham puanların toplanması yerine, T-puanlarının kullanılmasının uygun olacağını belirtmiştir (Torrance, 1974; Torrance & Ball, 1984).

Norm Çalışmaları

Çeşitli yaş grupları, meslek ve eğitim seviyelerindeki bireyler için kullanılabilen (Torrance, 1967) test bataryası, 1966 yılında geliştirilmiş ve 1974, 1984, 1990, 1998 ve 2008 yıllarında olmak üzere toplam beş norm çalışması ile güncellenmiştir (Torrance, 1966; 1974; 1990; 1998; 2008; Torrance & Ball, 1984).

Güvenirlilik Kanıtları

Test bataryasının güvenirlik çalışmaları kapsamında, bir öğretmen grubuna yalnızca puanlama rehberi verilerek 25 adet test bataryasını puanlamaları istenmiş, elde edilen puanların, eğitimli puanlayıcıların puanları ile korelasyonuna bakılmıştır. Sözel alt test puanları ortalama korelasyon katsayılarının en düşüğü orijinalliğe (0,94) en yükseği akıcılığa (0,99) ait hesaplanmıştır. Şekilsel alt test puanları ortalama korelasyon katsayılarının ise en düşüğü orijinallik (0,88) en yükseği ise akıcılık (0,96) için hesaplanarak, puanlayıcılar arası güvenirliğe kanıt olarak sunulmuştur (Torrance, 1974, s.17).

118 ilköğretim öğrencisine 1-2 haftalık aralıkla uygulanan AB/BA alternatif formlarının test tekrar test güvenirliği için elde edilen korelasyon katsayılarının en düşüğü şekilsel akıcılığa ($r=0,50$) en yükseği ise sözel akıcılığa ($r=0,93$) aittir (Torrance, 1974, s.19). Aynı örneklem grubuna uygulanan A ve B Formları puanları arasındaki korelasyon katsayılarının en düşüğü şekilsel akıcılık ($r=0,50$); en yükseği ise sözel akıcılık ($r=0,93$) için hesaplanmıştır. Ayrıca sözel ve şekilsel alt test arasında çok düşük bir korelasyon değeri ($r=0,06$) tespit edilmiştir (Torrance, 1990). Şekilsel alt testin iç tutarlık katsayıları ise 0,89 ile 0,94 arasında kestirilmiştir (Torrance, 1998).

Gök ve Erdoğan (2011) tarafından Hacettepe Üniversitesi Sınıf Öğretmenliği Anabilim Dalında eğitim gören 103 birinci sınıf öğrencisine uygulanan TYDT A Formu sözel alt testine ait iç tutarlık güvenirlik katsayısı 0,85 olarak kestirilmiştir.

Geçerlik Kanıtları

Yordama Geçerliği

TYDT puanları için öngörülen geçerlilik çalışmaları, 7 yıllık, 12 yıllık, 22 yıl ve 40 yıllık boylamsal çalışmalarda yaratıcı başarı ile anlamlı korelasyon göstermiştir (Torrance & Wu, 1981). 40 yıllık boylamsal çalışmaya göre, 1958-1998 yılları arasında yapılan orijinal çalışmadaki 391 öğrenciden 101'inin ilkokuldaki yaratıcılık puanı ile 40 yıl sonra elde edilen yaratıcılık kalitesi puanı arasında yüksek bir korelasyon (erkeklerde 0,45; kızlar için 0,41) elde edilerek TYDT'nin yordama geçerliğine kanıt olarak gösterilmiştir (Kim, 2006). TYDT yaratıcı başarıyı, iraksak düşünmeyi ve diğer potansiyel yaratıcılığı ölçen testlere göre daha iyi düzeyde ($r=0,33$, $p<0,0001$) yordamaktadır (Kim, 2008). Test bataryası, benzer yaratıcılık testlerine nazaran; büyük sayıda örneklem gruplarla norm çalışmaları (sözel alt test için $n = 10.127$, şekilsel alt test için $n = 37.814$) ile boylamsal geçerleme ve geniş bir yaş aralığındaki yüksek seviyeli yordama geçerliği kanıtları sebepleri ile daha kuvvetli bir yordayıcı olarak değerlendirilmektedir (Kim, 2006).

Yapı Geçerliği

Guilford (1959), yaratıcı düşünebilmeyi çok boyutlu olarak görmüş, birçok araştırmacı da yaratıcılığın çeşitli psikolojik faktörlerden oluştuğunu öne sürmüştür. Benzer şekilde, Torrance (1966; 1974) tek bir puanın kullanılmasının yanıltıcı olabileceği çünkü her alt puanın bağımsız bir anlamı olduğu konusunda uyarıda bulunarak tek bir yaratıcılık indeksi elde edilmesi konusunda tereddütlü hareket etmiştir. Fakat birçok araştırma yaratıcılığın tek boyutlu olduğunu ve bu nedenle ölçüldüğünden dolayı, tek bir puanın kullanılabileceğine dönük bulgulara ulaşmıştır (Clapham, 1998; Heausler & Thompson, 1988; Hocevar, 1979; Hocevar & Michael, 1979; Runco & Marz, 1992). Clapham (1998) ile Heausler & Thompson (1988) TYDT puanları üzerine yaptıkları faktör analizleri sonucu tek bir genel faktör elde etmişlerdir.

Fakat TYDT'nin tek boyutlu olmadığına dair bulgular elde eden araştırmalar da mevcuttur (Aranguren, 2014; Kim, 2006). Kim (2006) tarafından 6. sınıf 500

öğrenciye uygulanan A Formu şekilsel alt testinden elde edilen puanlara uygulanan Doğrulayıcı faktör analizi (DFA) sonucu, TYDT'nin *uyumluluk* ve *yenilikçilik* olarak adlandırdığı iki faktörlü yapıda olduğu bulgusuna ulaşmıştır. Bart, Hokanson & Can (2017) 1744 ortaöğretim öğrencisinin A formu şekilsel alt testi yanıtlarından elde edilen puanlara uyguladıkları DFA sonucu Kim (2006)'in iki faktörlü yapısını doğrulamışlardır.

Kapsam Geçerliği

Test bataryasının kapsam geçerliği için Guilford (1956) başta olmak üzere yaratıcılığın ne anlama geldiği ile ilgili araştırmalar (Simpson,1922; Spearman, 1931; Newell, Shaw & Simon, 1959) esas alınarak teorik kanıtlar sunulmuştur (Torrance, 1974; Torrance & Ball, 1984).

Uyarlama

İlk defa Toğrol (1999) tarafından Türkçeleştirilen testin, uyarlama çalışmaları Aslan (2001) tarafından yapılarak dilsel eşdeğerlik, güvenirlik ve geçerliğine kanıt sunulmuştur. Dilsel eş değerlik çalışmaları önce İngilizce sonra Türkçe olacak şekilde 15 gün ara ile her iki dili bilen 30 kişilik çalışma grubuyla; güvenirlik ve geçerlik çalışmaları ise anaokulu, ilköğretim, lise, üniversite ve çeşitli meslek gruplarından seçkisiz toplam 922 kişilik çalışma grubu ile gerçekleştirilmiştir. Çalışma analiz sonuçları Tablo 3.4'de sunulmuştur (Aslan, 2001). Türkçe versiyona ait norm çalışması ise yapılmamıştır.

Ölçme Aracı Olarak Seçilme Gerekçeleri

Yaratıcılık potansiyelini ölçen, 35'den fazla dile tercüme edilmiş (Kim, 2006), yarım asırdır yaratıcılık alanında en fazla araştırma yapılan (Cramond, Matthews-Morgan, Bandalos & Zuo, 2005), gerek eğitim gerekse iş dünyasında kullanılabilen bir test bataryası olan TYDT mevcut yaratıcılık testleri içerisinde en göze çarpanıdır. Çalışma grubunun yaş, eğitim ve mesleki özelliklerine uygunluğu (Torrance, 2003); aynı yapıyı ölçen benzer testlere göre psikometrik özelliklerinin daha iyi düzeyde olması; yaratıcılık alanında en yaygın (Plucker & Renzulli, 1999; Runco & Acar, 2012), tüm yaratıcılık testleri içerisinde en fazla alıntı yapılan (Fishkin, Cramond, Olszewski-Kubilius & Rejskind, 2000), kültürel farklılıklardan bağımsız (Rudowicz, Lok & Kitto, 1995), Türkiye'de birçok araştırmada kullanılan (Ergen ve Akyol, 2012; Gök ve Erdoğan, 2011; Karataş ve Özcan, 2010; Tok ve

Sevinç, 2012) yaratıcılık testi olması; uyarlama, dilsel eşdeğerlik, güvenilirlik ve geçerlik çalışmalarının ülkemiz için yapılmış olması sebepleriyle, bu araştırmada madde yazarlarının yaratıcılık düzeylerinin tespitinde veri toplama aracı olarak TYDT'nin kullanılması uygun görülmüştür. Çalışmada yaratıcılık düzeylerinin ölçülmesinde test bataryasının A formunun sözel ve şekilsel alt testleri birlikte işe koşulmuştur.

Tablo 3.4: TYDT Türkçe Uyarlama Çalışma Sonuçları

<i>Uyarlama Basamağı</i>	<i>Veri Kaynağı</i>	<i>İstatistik</i>	<i>Bulgular</i>
Dilsel Eşdeğerlik	Deneklerin iki dilde aldıkları test puanları	PMÇK	$0,50 < r < 0,96$
	Test puanları ortalamaları	İlişkili Gruplar t Testi	$p < 0,01$ düzeyinde anlamlı**
Güvenirlik /İç Tutarlılık	Test puanları	Guttman Korelasyonu	$0,48 < r < 0,89$
	Test puanları (Şekilsel+Sözel Alt Test)	Cronbach Alfa Korelasyonu	$0,56 < r < 0,86$
	Test puanları (Şekilsel+Sözel Alt Test)	Spearman Brown Korelasyonu	$0,38 < r < 0,89$
Geçerlik/Uyum	Test puanları- Sıfat Listesi Puanları	PMÇK	$-0,41 < r < -0,35^*$
	Test puanları- Wechsler Yetişkinler formu (WAIS) Puanları	SBSFK	$0,59 < p < 0,61$ ($p < 0,05$)
	Test puanları- Wonderlic Personel Testi (Genel Yetenek Testi) Puanları	SBSFK	$p = 0,79$ ($p < 0,05$)
Geçerlik/Yapı	Test puanları (Şekilsel Alt Test)	Madde Toplam	$p < 0,01$ düzeyinde anlamlı**
	Test puanları (Şekilsel Alt Test)	Madde Hariç	$p < 0,01$ düzeyinde anlamlı**
	Sıralanmış Test Puanları	Madde Ayırıcılığı İlişkisiz Gruplar t Testi	$p < 0,01$ düzeyinde anlamlı**

* Ters yapılar ölçüldüğünden negatif korelasyon elde edilmiştir.

** Değerler belirtilmemiştir.

3.3.2. Merkezî Sınav Eş Formları (MSEF)

Ortak bir kapsamda madde geliştirebilmek ve geliştirilen maddelerle oluşturulan teste uyum geçerliği kanıtı sunabilmek için 2016-2017 eğitim yılı 8.sınıf MOAS matematik ve Türkçe alt testleri ile aynı kapsamda ölçme yapabilecek eş formlar, çalışma grubundaki öğretmenler tarafından geliştirilen maddelerle oluşturulmuş, ilgili aşamalar sıralı başlıklarda açıklanmıştır.

3.3.2.1. Belirtke Tablosunun Oluşturulması

Uzman görüşlerine göre MOAS matematik ve Türkçe alt testlerindeki maddelerin ölçtüğü öğrenme alanları, kazanımları ve bilişsel süreç basamaklarını tespit etmek amacıyla Madde-Kazanım Eşleştirme Formu oluşturulmuştur. Araştırmada bu aşamada matematik (n=14) ve Türkçe (n=12) alan uzmanları görev almıştır. Görev alan uzmanlara ait branş ve mezuniyet bilgileri Tablo 3.5’de sunulmuştur.

Tablo 3.5: Alan Uzmanlarının Mezuniyet Durumlarına Göre Dağılımları

<i>Branşı</i>	<i>Alan Lisans Mezunu</i>		<i>Alan Tezsiz Yüksek Lisans Mezunu</i>		<i>Alan Yüksek Lisans Mezunu</i>		<i>Toplam</i>
	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	
Matematik	8	57	4	29	2	14	14
Türkçe	7	58	2	17	3	25	12
Toplam	15	100	6	100	5	100	26

Alt testlerin A kılavuzlarının maddeleri bir çizelgede sıralanarak, öğretim programındaki kazanım listesi ile birlikte alan uzmanlarına verilmiş ve maddelerin karşılıklarına, maddelerin hangi kazanımı ölçtüğünü tespit ederek, işaretlemeleri istenmiştir. Eşleştirme frekans ve yüzdeleri matematik alt testi için EK 5’de, Türkçe alt testi için EK 6’da sunulmuştur.

Uzman görüşleri alınarak, 2016-2017 yılı 8. Sınıf 1. Dönem MOAS matematik alt testi maddelerinin (TTKB, 2016), öğretim programında (MEB, 2013) belirlenen 15 kazanımla olan eşleştirme sonuçlarına göre en yüksek eşleştirme %100, en düşük %79 olmuştur.

2016-2017 yılı 8. Sınıf 1. Dönem MOAS Türkçe alt testi maddelerinin, öğretim programında (MEB, 2015a) belirlenen 77 kazanımla olan eşleştirme sonuçlarına göre en yüksek eşleştirme %100, en düşük %58 olmuştur.

Alan uzmanlarının ne derecede uyumlu eşleştirme yaptıklarının tespiti yoklanmıştır. Elde edilen veriler kategorik (kesikli) olduğundan, eşleştirmelerin ne düzeyde uyum gösterdiğinin tespiti için Fleiss Kappa Katsayıları hesaplanmıştır. Bu analiz için Fleiss Kappa uzantısı yüklenmiş SPSS 24.0 programı kullanılmıştır. Landis ve Koch’a (1977) göre Kappa değeri Tablo 3.6’a göre yorumlanmaktadır.

Tablo 3.6: Kappa İstatistiği Yorum Ölçütleri

<i>K</i>	<i>Yorum</i>
< 0	Hiç uyuşma olmaması
0,0 - 0,20	Önemsiz uyuşma olması
0,21 – 0,40	Orta derecede uyuşma olması
0,41 – 0,60	Ekseriyetle uyuşma olması
0,61 – 0,80	Önemli derecede uyuşma olması
0,81 – 1,00	Neredeyse mükemmel uyuşma olması

Alt test madde-kazanım eşleştirmesi Fleiss Kappa Katsayısı matematik testi için 0,929; Türkçe testi için 0,836 olarak bulunmuştur. Bu değerler alan uzmanları görüşleri arasında madde-kazanım eşleştirmesi için neredeyse mükemmel uyum olduğuna işaretir.

Puanlayıcılar arası güvenilirlik kanıtı sunularak, uzman görüşlerine dayalı MOAS matematik ve Türkçe alt testlerine ait madde-kazanım eşleştirmelerinin bir araya getirilmesi ile bu çalışmada geliştirilecek testlerin belirtke tabloları elde edilmiştir.

3.3.2.2. Maddelerin Yazılması ve Eş Formların Elde Edilmesi

Araştırmanın birinci safhasında her bir öğretmene seçkisiz olarak bir kazanım atanmış ve bu kazanımı yoklayacak madde geliştirmeleri istenmiştir. İki alanda yaratıcılığı yüksek grup öğretmenlerin geliştirdiği maddelerle Form I; yaratıcılığı düşük öğretmenlerin geliştirdiği maddelerle de Form II elde edilmiştir. Araştırmanın ikinci safhasında deney grubundaki öğretmenlere okullarında öğretmenler odası/toplantı salonlarında, bireysel ve/veya grup halinde, ortalama 4 saat süre ile Tablo 1.7’de sunulan konu başlıkları, konu kapsamaları, süreleri ve başarı ölçütleri sunulan madde yazarlığı eğitimi verilmiştir. Eğitim sonunda, öğretmenlerden araştırmanın başlangıcında kendilerine verilen kazanımı yoklayacak bir madde daha geliştirmeleri istenmiştir. Bu maddelerle de iki alanda Form III elde edilmiştir. Matematik alanında geliştirilen maddelerden bir kısmı EK 7’de, Türkçe alanında geliştirilen maddelerden bir kısmı ise EK 8’de sunulmuştur.

Maddeler formlara yerleştirilirken, madde sıralamasının (Tan, 2009b) ve seçeneklerin yerinin (Atılğan vd., 2009, s.250; Cizek, 1994; Hohensinn & Baghaei, 2017) madde istatistiklerine muhtemel etkisini önleyebilmek adına, MOAS A Formlarındaki aynı sıra izlenmiş ve maddelerin seçeneklerinin yerleri tüm cevaplayıcılar için aynı olacak şekilde yerleştirilmiş, değişikliğe gidilmemiştir.

Öğretmenler tarafından geliştirilen maddelerin dış kontrol kapsamında, öğrencilerin yanlış öğrenmelerine sebep olmamak için, bilimsel yönden bir yanlışının bulunup bulunmadığı Tablo 3.5’de mezuniyet durumları verilen alan uzmanlarınca; psikolojik durumlarına etki edebilecek durumlarının bulunup bulunmadığı ise 3 rehber öğretmence kontrol edilmiştir. Öğrencilere eğitimsel ve psikolojik zararı olabilecek herhangi bir maddenin bulunmadığı tespit edilmiştir. Dış kontrol kapsamında yapılması gereken diğer kontroller ise araştırmada analiz edildiğinden dolayı yapılmamıştır.

3.3.3. Madde Uygunluk Formu

MSEF’nun her bir maddesi için, alan uzmanları görüşlerine dayalı kapsam geçerliği çalışmasında kullanılmak üzere, düşey bölümünde maddelerin, yatay bölümünde ise belirtke tablosuna uygunluk ölçütlerinin bulunduğu form araştırmacı tarafından hazırlanmış ve Tablo 3.7’de bir maddeli örneği sunulmuştur.

Tablo 3.7: Madde Uygunluk Formu Örneği

Form	Madde No	Madde	Ölçtüğü Kazanım	Bu madde ile ölçülen kazanım GEREKLİ	Bu madde ile ölçülen kazanım YARARLI/ YETERSİZ	Bu madde ile ölçülen kazanım GEREKSİZ
I	3	Aşağıdakilerden hangisi 720 sayısının çarpanlarından biri değildir? a) 12 b) 16 c) 18 d) 22	8.1.1.1. Verilen pozitif tam sayıların çarpanlarını bulur; pozitif tam sayıları üslü ifade ya da üslü ifadelerin çarpımı şeklinde yazar. (U)			

3.4. Veri Toplama Araçlarının Uygulanışı

TYDT Türkiye yasal hak sahibinden, testin uygulama ve puanlamasına yönelik eğitim görülmüş, uygulama yapmak amacıyla izin alınmıştır (EK 5). Araştırma için Hacettepe Üniversitesi Etik Komisyonu’ndan Etik Onayı (EK 1), Ankara İl Millî Eğitim Müdürlüğü’nden araştırma izni alınmıştır (EK 2). Araştırmaya katılan öğretmenlere bilgilendirilmiş gönüllü onam formu ıslak imzalı teslim edilmiştir (EK 4).

3.4.1. Torrance Yararıcı Düşünme Testinin Uygulanışı

TYDT, araştırma izni alındıktan sonra, 80 öğretmene, bireysel ve grup halinde olmak üzere uygulanmıştır. Uygulama öncesinde öğretmenlerin kimlik bilgileri alınmadan, demografik bilgileri alınmıştır. Testin ölçtüğü yapı, madde adedi, yanıtlama yöntemi ve süre sınırlamasını içeren yönerge verilmiş, kitapçık numaralı formlar öğretmenlere verilerek, kitapçıkta bulunan yönergeleri takip etmeleri istenmiştir. Sözel alt testin her bir maddesi için 5'er, şekilsel alt testin her bir maddesi için ise 10'ar toplam 65 dakika sürede test uygulanarak, yaratıcılık indeksleri için veri toplanmıştır.

3.4.2. Merkezî Sınav Eş Formlarının Uygulanışı

Matematik ve Türkçe öğretmenlerince geliştirilen maddelerle 6 MSEF (I, II, III); MOAS'ndaki yönerge, madde-seçenek adedi, biçim, yazı tipi ve puntosuna benzer şekilde oluşturulmuştur. Örneklem grubundaki 1470 öğrenciye ulaşmak üzere veri toplama işlemine başlanmış, çeşitli sebeplerle (izin, yer değişikliği, rapor, mazeret sınavına katılma) 1313 öğrenciye ulaşılmıştır. MOAS sonuçlarının açıklanmasının ardından, toplamda 5 hafta sürede, 8.sınıf öğrencilerine (n=1313) 6 MSEF uygulanmıştır. Uygulama şubedeki öğrencilere aynı alanda 3 farklı formun birlikte şube mevcutlarına ve MOAS puanları durumlarına göre eşit miktarda olacak şekilde yapılmıştır. Örneğin bir şubede matematik eş formu uygulanırken; MOAS puanı 1-5 aralığında 3 öğrenci varsa birine Form I, birine Form II, birine de Form III verilmiştir. Form IV ile Form V'ler için ayrıca madde geliştirilmemiş, Form I ve Form II'nin maddeleriyle oluşturulmuş, öğrencilere uygulanmamış; ilgili maddelerin puanlarından analizler yapılmıştır. Çalışmaya dahil olan öğrenci sayıları, formlar ve branş bazında Tablo 3.8'de gösterilmiştir.

Tablo 3.8: Geliştirilen Formların Uygulandığı Öğrenci Sayıları

<i>Form</i>	<i>Form Maddelerini Geliştiren Öğretmen Grubu</i>	<i>Matematik</i>	<i>Türkçe</i>
Form I	Yüksek Yaratıcılık Düzeyi Grubu	219	219
Form II	Düşük Yaratıcılık Düzeyi Grubu	219	219
Form III	Deney Grubu (Madde Yazarlığı Eğitimi Alan)	219	218
Form IV	Kontrol Grubu (Madde Yazarlığı Eğitimi Almayan)	0	0
Form V	Deney Grubu (Eğitim Öncesi)	0	0
Toplam		657	656

Uygulama öncesi, öğrencilere testlerin hangi amaçla hazırlandığı, uyacakları kurallar, yanıtlama yöntemi ve süre konusunda bilgilendirme yapılmış, MOAS'dan aldıkları puanı cevap kâğıtlarına yazmaları istenmiştir. Öğrencileri motive edebilmek adına testin 26-27 Nisan 2017 tarihinde uygulanacak 2.dönem MOS konu kapsamının bir kısmını oluşturduğu hatırlatılmış, 1.dönem bilgilerinin bir tekrarı olarak kendilerine yararlı olacağı izah edilmiştir. Uygulama, MOS'nda olduğu gibi sınıflarda 40 dakikalık sürede yapılarak, madde ve test puanları için veri toplanmıştır.

3.4.3. Madde Uygunluk Formlarının Uygulanışı

Her bir MSEF alan uzmanlarına verilerek, her maddenin karşısına, ilgili madde ile hedeflenen kazanım ölçülüyor ise; "gerekli", kazanım ile ilişkili ancak gereksiz ise; "yararlı/yetersiz", kazanımı ölçmüyor ise; "gereksiz" seçeneklerinden birini işaretlemeleri istenmiştir. Maddenin birden fazla kazanımı yoklayabileceği düşünülerek, uzmanlardan bu durumda olan maddeyi, kazanım listesindeki kazanımlardan hangisine daha yakın olarak değerlendiriyorlar ise o kazanımı seçmeleri belirtilmiştir. 14 matematik ve 12 Türkçe alan uzmanı tarafından tüm eş formlar için bu analiz yapılarak kapsam geçerlik çalışması için veri toplanmıştır.

3.5. Verilerin İşlenmesi ve Çözümlemesi

Bu safhada öncelikle alt problemlerin yanıtlanmasında kullanılacak veriler elde edilmiş, müteakiben alt problemlerin analizlerine yer verilmiştir.

3.5.1. TYDT Uygulamasından Elde Edilen Verilerin İşlenmesi ve Analizi

Çalışma grubundaki öğretmenlerin, TYDT A Formu'na verdikleri yanıtlar, yanlılığı azaltabilmek için, test bataryasının puanlanmasında eğitimli üç puanlayıcı tarafından birbirlerinden bağımsız olarak puanlamışlardır. Puanlayıcılardan ikisi doktora öğrencisi, üçüncüsü ise yüksek lisans öğrencisi olan araştırmacıdır. Sözel alt test, Normlar-Teknik El Kitabı (Torrance, 1974) esaslarınca Sözel Puanlama Kâğıdı; şekilsel alt test ise Streamlined El Kitabı (Torrance & Ball, 1984) uyarınca Akıcı Puanlama Kâğıdı kullanılarak puanlanmıştır. Böylece her ölçüt için üç puanlayıcıdan olmak üzere toplam 9X3'lük veri seti elde edilmiştir. Elde edilen verilerin ne derecede güvenilir olduğunun tespiti için ham puanlar üzerinden puanlayıcı güvenilirliğinin kestirimi yapılmıştır.

3.5.1.1. Matematik Öğretmenleri TYDT Puanlarına Ait Güvenirlik Çalışması

İç Tutarlık Güvenirliği

Puanlayıcıların 9 ölçüt için, matematik öğretmenlerinin TYDT'ne verdikleri yanıtlara yaptıkları puanlamaların betimsel istatistikleri Tablo 3.9'da sunulmuştur.

Tablo 3.9: Matematik Öğretmenleri TYDT Ham Puanlarına Ait Betimsel İstatistikler

	Ölçütler	Puanlayıcı	Xmin	Xmax	\bar{x}	S _x	Çarpıklık	Basıklık
Sözel Alt Test	Akıcılık	1	13	107	45,73	20,59	0,689	0,493
		2	13	107	45,73	20,67	0,663	0,455
		3	11	106	45,73	20,55	0,642	0,417
	Esneklik	1	9	56	28,78	13,51	0,336	-1,182
		2	9	52	28,60	13,29	0,325	-1,267
		3	9	53	28,73	13,44	0,321	-1,247
	Orijinallik	1	9	95	35,55	21,63	1,055	0,913
		2	9	94	34,45	20,76	0,916	0,492
		3	9	94	35,15	20,89	0,942	0,618
Şekilsel Alt Test	Akıcılık	1	5	40	23,23	8,60	0,095	0,019
		2	5	40	23,40	8,53	0,076	0,155
		3	4	40	22,90	8,78	0,084	0,115
	Orijinallik	1	1	51	17,90	11,49	0,557	0,009
		2	1	46	17,68	11,38	0,431	-0,677
		3	1	46	17,20	11,34	0,441	-0,730
	Başlıkların Soyutluğu	1	0	12	4,38	3,16	0,483	-0,589
		2	0	12	4,15	3,36	0,737	-0,589
		3	0	12	4,10	2,93	0,743	0,163
	Zenginleştirme	1	2	10	5,75	1,92	0,284	-0,227
		2	2	9	5,43	1,77	0,101	-0,605
		3	2	10	5,45	1,85	0,287	-0,353
	Erken Kapamaya Direnç	1	0	8	3,38	2,33	0,109	-0,792
		2	0	7	3,18	2,21	0,038	-1,212
		3	0	8	3,18	2,29	0,207	-0,766
Bonus	1	0	28	7,90	7,28	1,472	2,387	
	2	0	21	6,35	5,71	1,056	1,147	
	3	0	18	6,20	5,07	0,673	0,070	

Tablo 3.9 incelendiğinde matematik öğretmenlerine ait sözel alt testte akıcılığın 11-107; esnekliğin 9-56; orijinalliğin 9-95; şekilsel alt testte akıcılığın 4-40; orijinalliğin 1-51; başlıkların soyutluğunun 0-12; zenginleştirmenin 2-10; erken kapamaya direncin 0-8; bonusun 0-28 arası değerler aldığı gözlemlenmiştir. Puan

ortalamlarının en yükseği sözel akıcılığa ($\bar{x}=45,73$) en düşüğü ise erken kapamaya dirence ($\bar{x}=3,18$) aittir. Puan dağılımlarının standart sapmaları incelendiğinde, en fazla farklılaşmanın sözel orijinalliğe ($S_x=21,63$); en az farklılaşmanın zenginleştirmeye ($S_x=1,77$) ait olduğu görülmektedir.

Her bir puanlayıcının, dokuz ölçüte verdiği puanlar için Cronbach alfa güvenilirlik katsayısı SPSS 24.0 programı ile hesaplanmıştır. Cronbach alfa güvenilirlik katsayıları birinci puanlayıcı için 0,848; ikinci puanlayıcı için 0,843 ve üçüncü puanlayıcı için ise 0.840 olarak bulunmuştur. Elde edilen bu katsayılara göre, yaratıcılığın ölçülmesinde yer alan ölçütlerin, yaratıcılığı birbirleriyle tutarlı bir şekilde ölçtükleri yorumu yapılabilir.

Puanlayıcı Güvenirliđi

Puanlayıcıların birbirleri ile ne derecede uyumlu puanlama yaptıklarının belirlenmesi için iki alanda da sırasıyla; ölçüt puanları arasındaki korelasyonların analizleri yapılmış, puan ortalamaları arasındaki fark sınanmış ve G çalışması ile puan varyanslarındaki puanlayıcı etkisi analiz edilmiştir.

Analizlerde kullanılacak yöntem karar vermek için puan dağılımlarının normalliği yoklanmıştır. Puanlayıcıların öğretmenlere ölçüt bazında verdikleri puanların grup büyüklüğü 50'den küçük ve küçük örneklerde güçlü bir istatistik olması (Öztuna, Elhan & Tüccar, 2006) sebebiyle Shapiro-Wilk Normallik Testi (S-W) ile normallik sınanmıştır. "Puan dağılımı normal dağılımdan anlamlı farklılık göstermez" şeklinde kurulan yokluk hipotezinin analizi sonucu hesaplanan p değeri $\alpha=0,05$ 'den küçük olması başka bir ifade ile anlamlı sonuçlanması durumunda dağılımlar için çarpıklık katsayılarının standart hatasına bölünmesi ile elde edilecek z-istatistiğinin $\alpha=0,05$ için $\pm 1,96$ aralığında olup olmadığı yoklanmıştır (Büyüköztürk, 2015, s.42). S-W SPSS 24.0 ile z-istatistikleri ise Microsoft Office 2007 Excel programıyla hesaplanmış, sonuçlar Tablo 3.10'da sunulmuştur.

Tablo 3.10: Matematik Öğretmenlerine Ait TYDT Puan Dağılımları Normallik Testi Sonuçları

<i>Ölçüt</i>	<i>Puanlayıcı</i>	<i>İstatistik</i>	<i>p</i>	<i>Z_{çarpıklık}</i>
Sözel Akıcılık	1	0,956	0,119	
	2	0,957	0,128	
	3	0,962	0,192	
Sözel Esneklik	1	0,933	0,020	0,90 [#]
	2	0,923	0,009	0,87 [#]
	3	0,923	0,009	0,86 [#]
Sözel Orijinallik	1	0,896	0,001	2,82
	2	0,916	0,006	2,45
	3	0,918	0,007	2,52
Şekilsel Akıcılık	1	0,971	0,376	
	2	0,961	0,181	
	3	0,968	0,319	
Şekilsel Orijinallik	1	0,943	0,045	1,49 [#]
	2	0,950	0,075	
	3	0,935	0,024	1,18 [#]
Başlıkların Soyutluğu	1	0,948	0,064	
	2	0,898	0,002	1,97
	3	0,940	0,034	1,99
Zenginleştirme	1	0,966	0,260	
	2	0,955	0,114	
	3	0,963	0,204	
Erken Kapamaya Direnç	1	0,941	0,036	0,29 [#]
	2	0,929	0,015	0,10 [#]
	3	0,936	0,026	0,56 [#]
Bonus	1	0,837	0,000	3,94
	2	0,869	0,000	2,83
	3	0,903	0,002	1,80 [#]

[#] $z < 1,96$

Sözel akıcılık, şekilsel akıcılık ve zenginleştirmenin S-W sonucu anlamlı olmaması; sözel esneklik, erken kapamaya direnç z-istatistiğinin kabul edilen aralıkta olması sebebiyle normal dağılımlı olduğu kararı verilmiştir. Şekilsel orijinallik puanlayıcı-2 için normallik testi sonucu ve puanlayıcı-1,3 için ise z-istatistikleri sonucu normal dağıldığı görülmüştür. Söz konusu ölçütler için parametrik; sözel orijinallik, başlıkların soyutluğu ve bonus ölçütleri normallik testinin anlamlı çıkması sebebiyle parametrik olmayan yöntemlerle analiz yapılması gerektiği anlaşılmıştır.

Puanlayıcılar arası güvenilirlik için her bir puanlayıcının verdiği puanlar ile diğer bir puanlayıcının verdiği puanlar arasındaki korelasyon katsayıları; sözel orijinallik, başlıkların soyutluğu ve bonus ölçütleri için Spearman Brown Sıra Farkları Korelasyonu (SBSFK); diğer 6 ölçüt için Pearson Momentler Çarpımı Korelasyonu (PMÇK) ile, SPSS 24.0 programı kullanılarak hesaplanmıştır. Puanlayıcıların 9

ölçüt üzerinden verdikleri puanlar arasındaki korelasyon değerleri Tablo 3.11’de sunulmuştur.

Tablo 3.11: Üç Puanlayıcının Matematik Öğretmenlerine Ait Ölçüt Puanları Arasındaki Korelasyon Katsayıları

	<i>Puanlayıcı-1</i>	<i>Puanlayıcı-2</i>	<i>Puanlayıcı-3</i>		<i>Puanlayıcı-1</i>	<i>Puanlayıcı-2</i>	<i>Puanlayıcı-3</i>		<i>Puanlayıcı-1</i>	<i>Puanlayıcı-2</i>	<i>Puanlayıcı-3</i>	
		Sözel Akıcılık				Sözel Esneklik				Sözel Orijinallik		
Puanlayıcı-1	-	0,988**	1,000**	-	0,988**	0,991**	-	0,968**	0,976**			
Puanlayıcı-2		-	0,997**		-	0,994**		-	0,961**			
		Şekilsel Akıcılık				Şekilsel Orijinallik				Başlıkların Soyutluğu		
Puanlayıcı-1	-	0,962**	0,983**	-	0,976**	0,970**	-	0,965**	0,971**			
Puanlayıcı-2		-	0,962**		-	0,966**		-	0,942**			
		Zenginleştirme				Erken Kapamaya Direnç				Bonus		
Puanlayıcı-1	-	0,902**	0,927**	-	0,926**	0,893**	-	0,931**	0,908**			
Puanlayıcı-2		-	0,888**		-	0,889**		-	0,949**			

** $p < 0,01$

Tablo 3.11’de görüldüğü gibi, üç puanlayıcının 9 ölçüte verdiği puanlar arasındaki korelasyon katsayıları 0,89 ile 1,00 arasında değişen yüksek değerlere, anlamlı düzeyde ($p < 0,01$) sahiptir. Elde edilen bu değerler puanlayıcılar arası uyumun olduğu şeklinde yorumlanabilir. Ancak puanlayıcılar arası korelasyon katsayıları ortalamalardan bağımsız olarak hesaplandığı için, puanlayıcıların verdikleri puanların ortalamaları arasındaki farklılıkları ortaya koyamamaktadır. Bu sebeple puanlayıcıların puan ortalamaları arasındaki farkın test edilmesi önerilmektedir (Goodwin, 2001). Üç puanlayıcının 40 öğretmene 9 ölçütte verdiği puanların ortalamaları arasındaki farklılığın yoklanması için puan dağılımlarına uygun olarak; sözel orijinallik, başlıkların soyutluğu ve bonus ölçütleri için ilişkisiz ölçümler için Kruskal Wallis H-Testi, diğer 6 ölçüt için ilişkisiz örneklem için tek yönlü varyans analizinin (ANOVA) yapılmasına karar verilmiştir. ANOVA varsayımları; bağımlı değişkene ait puanların en az aralık ölçeğinde olması, ortalama puanları karşılaştırılacak grupların ilişkisiz olması, puanların bağımlı değişkende etkisi araştırılan faktörün her bir düzeyinde normal dağılım göstermesi ve bağımlı değişkene ilişkin varyansların her bir örneklem için eşit olmasıdır (Büyüköztürk, 2015, s.48). Eşit aralık ölçeğinde, ilişkisiz, normal dağılımlı 6 ölçütün varyanslarının eşit olup olmadığına yönelik yapılan Levene testi sonucu

varyansları anlamlılık düzeyi sınırları içerisinde eşit görülmüştür ($p>0,05$). Ölçüt puanların ortalamaları arasında farklılığa dönük SPSS 24.0 programı ile yapılan analiz sonuçları Tablo 3.12’de sunulmuştur.

Tablo 3.12: Üç Puanlayıcının Matematik Öğretmenlerine Ait Ölçüt Puanları ANOVA Sonuçları

	<i>Varyansın Kaynağı</i>	<i>Kareler Toplamı</i>	<i>Sd</i>	<i>Kareler Ortalaması</i>	<i>F/χ²</i>	<i>p</i>
Sözel Akıcılık	Gruplararası	49601,26	39	1271,827		
	Gruplarıçi	0	2	0,000	0,000	1,000
	Hata	56,667	78	0,726		
	Toplam	49657,93	119			
Sözel Esneklik	Gruplararası	20833,87	39	536,766		
	Gruplarıçi	0,65	2	0,325	0,207	0,814
	Hata	122,683	78	1,573		
	Toplam	20957,2	119			
Şekilsel Akıcılık	Gruplararası	8554,658	39	219,350		
	Gruplarıçi	5,15	2	2,575	1,107	0,336
	Hata	181,517	78	2,327		
	Toplam	8741,325	119			
Şekilsel Orijinallik	Gruplararası	14913,66	39	382,401		
	Gruplarıçi	10,217	2	5,108	1,341	0,268
	Hata	297,117	78	3,809		
	Toplam	15220,99	119			
Zenginleştirme	Gruplararası	373,792	39	9,584		
	Gruplarıçi	2,617	2	1,308	4,020	0,022
	Hata	25,383	78	0,325		
	Toplam	401,792	119			
Erken Kapamaya Direnç	Gruplararası	565,325	39	14,496		
	Gruplarıçi	1,067	2	0,533	1,051	0,355
	Hata	39,6	78	0,508		
	Toplam	605,992	119			
Sözel Orijinallik			2		0,055	0,973
Başlıkların Soyutluğu			2		0,255	0,880
Bonus			2		0,887	0,642

Tablo 3.12’e göre; puanlayıcılar arası ölçüt puanlar bazında anlamlı fark bulunmamıştır ($p>0,01$).

Yaratıcılığın ölçülmesiyle ulaşılan puanların, G çalışması ile elde edilen varyanslarını ve varyans yüzdelerini hesaplamak amacıyla, tümüyle çaprazlanmış b (birey-öğretmen) xö (ölçüt) xp (puanlayıcı) modeli uygulanmıştır. Analiz için G String IV 6.1.1 programı kullanılmıştır. 40 öğretmen, 3 puanlayıcı ve 9 ölçüt üzerinde yapılan G çalışması için; kestirilen varyans bileşenleri ve toplam varyansı açıklama yüzdeleri b, ö ve p ana etkileri ile bö, bp, öp ve böp ortak etkileri Tablo 3.13’de sunulmuştur.

Tablo 3.13: G Çalışması ile Matematik Öğretmenleri Yaratıcılıkları Ölçmesinin Kestirilen Varyansları ve Toplam Varyansı Açıklama Oranları

<i>Varyans Kaynağı</i>	<i>Sd</i>	<i>Toplam Kareler</i>	<i>Kareler Ortalaması</i>	<i>Varyans</i>	<i>%</i>
b	39	67794,261	1738,314	54,388	14,36
ö	8	225829,275	28228,659	232,976	61,54
p	2	42,739	21,369	0,044	0,01
bö	312	83941,083	269,042	88,994	23,51
bp	78	222,965	2,859	0,089	0,02
öp	16	74,344	4,647	0,065	0,02
böp	624	1285,285	2,060	2,060	0,54
Toplam	1079				100

Tablo 3.13'de verilen G çalışması sonucunda, kestirilen varyans ve toplam varyansı açıklama oranları incelendiğinde, puanlayıcı ana etkisi için kestirilen varyans bileşeni (0.044) toplam varyansın % 0,01'ini açıklayarak, ana etkiler içinde en düşük varyans bileşeni olmaktadır. Puanlayıcı ana etkisinin G çalışması ile kestirilen varyans oranının düşük çıkması, puanlayıcıların tüm bireyler için yaptıkları puanlamalar arasında büyük bir farklılık bulunmadığının, puanlamalar arasında bir tutarlılığın söz konusu olduğunun göstergesi olarak sunulabilir. Birey x puanlayıcı (bp) ortak etkisi (0.089) toplam varyansın % 0,02'sini açıklamaktadır. Buna göre, bu ölçme için birey x puanlayıcı ortak etkisinden kaynaklanan farklılığın küçük olduğu, belli bireylerin bağıl durumlarının bir puanlayıcıdan diğerine çok değişmediği yorumu yapılabilir. Ölçüt x puanlayıcı (öp) ortak etkisi (0.065) toplam varyansın % 0,02'sini açıklamaktadır. Bu sonuca göre, ölçüt x puanlayıcı ortak etkisinden kaynaklanan farklılığın neredeyse olmadığı, belli ölçütlere verilen puanların puanlayıcıdan puanlayıcıya değişmediği yorumunda bulunulabilir. Birey x ölçüt x puanlayıcı ortak etkisi varyans bileşeninin de (2,060) toplam varyansı açıklama oranı % 0,54 olarak görülmektedir. Birey x ölçüt x puanlayıcı varyansının küçük olması; birey, puanlayıcı ve ölçüt ortak etkisi ve / veya tesadüfî hataların küçük olabileceğinin bir göstergesi olabilir. Çaprazlanmış desen analizi sonucu G (Genellenebilirlik) ve Phi "Φ" (Güvenirlilik) katsayıları sırasıyla 0,845 ve 0,602 olarak kestirilmiştir. Bu veriler ışığında puanlayıcıların elde edilen puanlara neredeyse etkisi olmadığı, puanlayıcılar arası güvenirliliğin uygun düzeyde olduğu ileri sürülebilir.

3.5.1.2. Türkçe Öğretmenleri TYDT Puanlarına Ait Güvenirlik Çalışması

İç Tutarlık Güvenirliği

Puanlayıcıların, 9 ölçüt için, Türkçe öğretmenlerinin TYDT'ne verdikleri yanıtlara yaptıkları puanlamaların betimsel istatistikleri Tablo 3.14'de sunulmuştur.

Tablo 3.14: Türkçe Öğretmenleri TYDT Ham Puanlarına Ait Betimsel İstatistikler

	Ölçütler	Puanlayıcı	Xmin	Xmax	\bar{x}	S _x	Çarpıklık	Basıklık
Sözel Alt Test	Akıcılık	1	18	78	40,78	16,07	0,689	-0,420
		2	18	78	40,78	16,07	0,689	-0,420
		3	18	78	40,78	16,07	0,689	-0,420
	Esneklik	1	11	44	26,45	8,49	0,389	-0,689
		2	13	44	26,55	8,36	0,480	-0,672
		3	12	44	26,60	8,28	0,389	-0,652
	Orijinallik	1	4	65	31,03	15,21	0,461	-0,432
		2	3	65	30,45	14,95	0,597	-0,201
		3	7	65	31,08	15,16	0,524	-0,511
Şekilsel Alt Test	Akıcılık	1	16	40	29,25	8,44	0,041	-1,115
		2	16	40	29,25	8,40	0,056	-1,105
		3	16	40	28,95	8,61	0,087	-1,189
	Orijinallik	1	3	36	19,15	7,85	0,009	-0,316
		2	4	37	19,45	7,77	0,007	-0,129
		3	2	37	19,15	7,86	0,009	-0,040
	Başlıkların Soyutluğu	1	1	15	6,23	4,01	0,733	-0,215
		2	0	16	6,25	4,25	0,780	-0,035
		3	0	17	6,45	4,37	0,803	-0,275
	Zenginleştirme	1	3	8	4,60	1,50	0,687	-0,459
		2	3	9	4,58	1,57	0,885	0,248
		3	3	8	4,58	1,55	0,852	-0,234
	Erken Kapamaya Direnç	1	0	9	3,38	2,14	0,583	-0,105
		2	0	7	3,35	2,20	0,288	-1,241
		3	0	8	3,35	2,33	0,439	-0,867
Bonus	1	1	39	7,28	7,29	2,500	8,409	
	2	1	35	6,48	6,41	2,664	9,431	
	3	1	39	6,60	6,87	2,985	12,110	

Tablo 3.14 incelendiğinde Türkçe öğretmenlerine ait sözel alt testte akıcılığın 18-78; esnekliğin 11-44; orijinalliğin 3-65; şekilsel alt testte akıcılığın 16-40; orijinalliğin 2-37; başlıkların soyutluğunun 0-17; zenginleştirmenin 3-9; erken kapamaya direncin 0-9; bonusun 1-39 arası değerler aldığı gözlemlenmiştir. Puan ortalamalarının en yükseği sözel akıcılığa ($\bar{x}=40,78$) en düşüğü ise erken kapamaya dirence ($\bar{x}=3,35$) aittir. Puan dağılımlarının standart sapmaları incelendiğinde, en fazla farklılaşmanın sözel akıcılığa ($S_x=16,07$); en az farklılaşmanın zenginleştirmeye ($S_x=1,50$) ait olduğu görülmektedir.

Her bir puanlayıcının, dokuz ölçüte verdiği puanlar için Cronbach alfa güvenilirlik katsayıları birinci puanlayıcı için 0,814; ikinci puanlayıcı için 0,818 ve üçüncü puanlayıcı için ise 0,814 olarak bulunmuştur. Elde edilen bu katsayılara göre, yaratıcılığın ölçülmesinde yer alan ölçütlerin, yaratıcılığı birbirleriyle tutarlı bir şekilde ölçtükleri yorumu yapılabilir.

Puanlayıcı Güvenirliği

Türkçe öğretmenlerinin 9 ölçüte 3 puanlayıcı tarafından verilen puanların S-W sonuçları ile z-istatistikleri Tablo 3.15’de sunulmuştur.

Tablo 3.15: Türkçe Öğretmenlerine Ait TYDT Puan Dağılımları Normallik Testi Sonuçları

<i>Ölçüt</i>	<i>Puanlayıcı</i>	<i>İstatistik</i>	<i>p</i>	<i>Z_{çarpıklık}</i>
Sözel Akıcılık	1	0,935	0,023	1,84 [#]
	2	0,935	0,023	1,84 [#]
	3	0,935	0,023	1,84 [#]
Sözel Esneklik	1	0,961	0,181	
	2	0,951	0,081	
	3	0,966	0,268	
Sözel Orijinallik	1	0,967	0,296	
	2	0,958	0,138	
	3	0,956	0,125	
Şekilsel Akıcılık	1	0,912	0,004	-0,10 [#]
	2	0,913	0,005	-0,06 [#]
	3	0,907	0,003	0,04 [#]
Şekilsel Orijinallik	1	0,985	0,863	
	2	0,981	0,722	
	3	0,987	0,927	
Başlıkların Soyutluğu	1	0,921	0,008	1,96
	2	0,926	0,012	2,09
	3	0,906	0,003	2,15
Zenginleştirme	1	0,880	0,001	1,84 [#]
	2	0,867	0,000	2,37
	3	0,857	0,000	2,28
Erken Kapamaya Direnç	1	0,942	0,042	1,56 [#]
	2	0,912	0,004	0,77 [#]
	3	0,933	0,021	1,17 [#]
Bonus	1	0,741	0,000	6,69
	2	0,725	0,000	7,13
	3	0,699	0,000	7,99

[#]-1,96<z<1,96

Sözel esneklik, sözel orijinallik ve şekilsel orijinalliğin S-W sonucunun anlamlı olmaması; sözel akıcılık, şekilsel akıcılık ve erken kapamaya direnç z-istatistiklerinin kabul edilen aralıkta olması sebebiyle normal dağılımlı olduğu kararı verilmiştir. Başlıkların soyutluğu, zenginleştirme ve bonus puanları üç puanlayıcının en az biri tarafından normallik kriterlerinin dışında kaldığından, bu ölçütlere ait analizlerin parametrik olmayan yöntemlerle yapılmasına, diğer 6 ölçüt için parametrik yöntemlerin kullanılmasına karar verilmiştir.

Puanlayıcılar arası güvenilirlik için her bir puanlayıcının verdiği puanlar ile diğer bir puanlayıcının verdiği puanlar arasındaki korelasyon katsayıları; başlıkların soyutluğu, zenginleştirme ve bonus ölçütleri için SBSFK; diğer 6 ölçüt için PMÇK ile, SPSS 24.0 programı kullanılarak hesaplanmıştır. Puanlayıcıların 9 ölçüt üzerinden verdikleri puanlar arasındaki korelasyon değerleri Tablo 3.16'da sunulmuştur.

Tablo 3.16: Üç Puanlayıcının Türkçe Öğretmenlerine Ait Ölçüt Puanları Arasındaki Korelasyon Katsayıları

	<i>Puanlayıcı-1</i>	<i>Puanlayıcı-2</i>	<i>Puanlayıcı-3</i>	<i>Puanlayıcı-1</i>	<i>Puanlayıcı-2</i>	<i>Puanlayıcı-3</i>	<i>Puanlayıcı-1</i>	<i>Puanlayıcı-2</i>	<i>Puanlayıcı-3</i>
		Sözel Akıcılık		Sözel Esneklik			Sözel Orijinallik		
Puanlayıcı-1	-	1,000**	1,000**	-	0,995**	0,994**	-	0,980**	0,998**
Puanlayıcı-2		-	1,000**		-	0,991**		-	0,979**
		Şekilsel Akıcılık		Şekilsel Orijinallik			Başlıkların Soyutluğu		
Puanlayıcı-1	-	1,000**	0,982**	-	0,986**	0,990**	-	0,901**	0,911**
Puanlayıcı-2		-	0,982**		-	0,979**		-	0,884**
		Zenginleştirme		Erken Kapamaya Direnç			Bonus		
Puanlayıcı-1	-	0,901**	0,896**	-	0,943**	0,965**	-	0,962**	0,960**
Puanlayıcı-2		-	0,850**		-	0,956**		-	0,955**

** $p < 0,01$

Tablo 3.16'da görüldüğü gibi, üç puanlayıcının 9 ölçüte verdiği puanlar arasındaki korelasyon katsayıları 0,850 ile 1,000 arasında değişen yüksek değerlere, anlamlı düzeyde ($p < 0,01$) sahiptir. Elde edilen bu değerler puanlayıcılar arası uyumun olduğu şeklinde yorumlanabilir. Üç puanlayıcınının 40 öğretmene 9 ölçütte verdiği puanların ortalamaları arasındaki farklılık; başlıkların soyutluğu, zenginleştirme ve bonus ölçütleri için ilişkisiz ölçümler için Kruskal Wallis H-Testi ile, varyansları

anlamlılık düzeyi sınırları içerisinde eşit görülen ($p>0,05$) 6 ölçüt için ilişkisiz örneklem için tek yönlü varyans analizi (ANOVA) ile SPSS 24.0 programında test edilmiş, sonuçları Tablo 3.17’de her bir ölçüt puanı için sunulmuştur.

Tablo 3.17: Üç Puanlayıcının Türkçe Öğretmenlerine Ait Ölçüt Puanları ANOVA Sonuçları

	<i>Varyansın Kaynağı</i>	<i>Kareler Toplamı</i>	<i>Sd</i>	<i>Kareler Ortalaması</i>	<i>F</i>	<i>p</i>
Sözel Akıcılık	Gruplararası	30200,93	39	774,383		
	Gruplarıçi	0	2	0,000	0,000	1,000
	Hata	0	78	0,000		
	Toplam	30200,93	119			
Sözel Esneklik	Gruplararası	8171,867	39	209,535		
	Gruplarıçi	0,467	2	0,233	0,485	0,618
	Hata	37,533	78	0,481		
	Toplam	8209,867	119			
Sözel Orijinallik	Gruplararası	26448,63	39	678,170		
	Gruplarıçi	9,65	2	4,825	1,476	0,235
	Hata	255,017	78	3,269		
	Toplam	26713,3	119			
Şekilsel Akıcılık	Gruplararası	7945,592	39	203,733		
	Gruplarıçi	2,4	2	1,200	1,398	0,253
	Hata	66,933	78	0,858		
	Toplam	8014,925	119			
Şekilsel Orijinallik	Gruplararası	7090,5	39	181,808		
	Gruplarıçi	2,4	2	1,200	1,307	0,276
	Hata	71,6	78	0,918		
	Toplam	7164,5	119			
Erken Kapamaya Direnç	Gruplararası	561,592	39	14,400		
	Gruplarıçi	0,017	2	0,008	0,036	0,965
	Hata	17,983	78	0,231		
	Toplam	579,592	119			
Başlıkların Soyutluğu			2		0,038	0,981
Zenginleştirme			2		0,032	0,984
Bonus			2		0,373	0,830

Tablo 3.17’e göre; puanlayıcılar arası ölçüt puanlar bazında anlamlı fark bulunmamıştır ($p>0,01$).

40 öğretmen, 3 puanlayıcı ve 9 ölçüt üzerinde yapılan G çalışması için; kestirilen varyans bileşenleri ve toplam varyansı açıklama yüzdeleri b, ö ve p ana etkileri ile bö, bp, öp ve böp ortak etkileri Tablo 3.18’de sunulmuştur.

Tablo 3.18: G Çalışması ile Türkçe Öğretmenleri Yaratıcılıkları Ölçmesinin Kestirilen Varyansları ve Toplam Varyansı Açıklama Oranları

<i>Varyans Kaynağı</i>	<i>Sd</i>	<i>Toplam Kareler</i>	<i>Kareler Ortalaması</i>	<i>Varyans</i>	<i>%</i>
b	39	35745,393	916,549	27,744	10,10
ö	8	183999,308	22999,914	190,263	69,30
p	2	2,252	1,126	0*	0,00
bö	312	52270,174	167,533	55,501	20,21
bp	78	74,341	0,953	0*	0,00
öp	16	28,731	1,796	0,019	0,01
böp	624	643,343	1,031	1,031	0,38
Toplam	1079				100

* *Negatif varyans bileşenleri sıfır olarak alınmıştır (Brennan, 2001, s.84).*

Tablo 3.18'de verilen G çalışması sonucunda, kestirilen varyans ve toplam varyansı açıklama oranları incelendiğinde, puanlayıcı ana etkisi için kestirilen varyans bileşeni (0) yoktur. Puanlayıcı ana etkisinin G çalışması ile kestirilen varyans oranının sıfır çıkması, puanlayıcıların tüm bireyler için yaptıkları puanlamalar arasında farklılık bulunmadığının, puanlamalar arasında bir tutarlılığın söz konusu olduğunun göstergesi olarak sunulabilir. Birey x puanlayıcı (bp) ortak etkisi (0) yoktur, buna göre, bu ölçme için birey x puanlayıcı ortak etkisinden kaynaklanan farklılığın olmadığı, belli bireylerin bağıl durumlarının bir puanlayıcıdan diğerine değişmediği yorumu yapılabilir. Ölçüt x puanlayıcı (öp) ortak etkisi (0.019) toplam varyansın % 0,01'ini açıklamaktadır. Bu sonuca göre, ölçüt x puanlayıcı ortak etkisinden kaynaklanan farklılığın neredeyse olmadığı, belli ölçütlere verilen puanların puanlayıcıdan puanlayıcıya değişmediği yorumunda bulunulabilir. Birey x ölçüt x puanlayıcı ortak etkisi varyans bileşeninin de (1,031) toplam varyansı açıklama oranı % 0,38 olarak görülmektedir. Birey x ölçüt x puanlayıcı varyansının küçük olması; birey, puanlayıcı ve ölçüt ortak etkisi ve / veya tesadüfî hataların küçük olabileceğinin bir göstergesi olabilir. Çaprazlanmış desen analizi sonucu ise G (Genellenebilirlik) ve Phi "Φ" (Güvenirlik) katsayıları sırasıyla 0,817 ve 0,504 olarak kestirilmiştir. Bu veriler ışığında puanlayıcıların elde edilen puanlara neredeyse etkisi olmadığı, puanlar arası güvenirliliğin uygun düzeyde olduğu ileri sürülebilir.

3.5.1.3. TYDT Puanlarına Yönelik Yapılan Yapı Geçerliği Çalışması

Matematik ve Türkçe öğretmenlerinin yaratıcılıklarını belirten tek bir puana, öğretmenleri gerek birbirleriyle kıyaslayabilmek gerekse yüksek/düşük düzey ayrımı yapabilmek için ihtiyaç duyulmaktadır. Sözel ve şekilsel alt testlerin ölçüt puanlarının birleşimi ile elde edilebilecek herhangi bir kompozit puan hesaplaması bulunmamaktadır (Torrance, 1974; Torrance & Ball, 1984). Ayrıca Türkiye için norm çalışmasının yapılmamış olması nedeniyle elde edilen puanların karşılaştırılabileceği norm kriterleri de mevcut değildir (Aslan, 2001). TYDT’de bulunan toplam 9 ölçüt, yaratıcılığı farklı pencerelerden fakat bütün olarak ölçmektedir. Herhangi bir ölçütün bir diğerine göre bir önceliği ya da ağırlığı bulunmamaktadır (Torrance, 1974; Torrance & Ball, 1984). Üç puanlayıcıdan elde edilen puanlar güvenilir olduğundan hareketle her bir ölçüt için üç puanlayıcının verdiği puanların aritmetik ortalaması alınmıştır. Elde edilen aritmetik ortalamalardan tek bir puan elde edilip edilmeyeceğinin kararını verebilmek için TYDT puanlarının tek boyutlu olup olmadığının sınılanması gerektiği düşünülmüştür. Bunun için matematik öğretmenlerinin ortalama ölçüt puanları üzerinden Açıklayıcı Faktör Analizi (AFA), Türkçe öğretmenlerinin ortalama ölçüt puanları üzerinden ise DFA yapılmasına karar verilmiştir.

Matematik öğretmenlerinin ortalama ölçüt puanlarının faktör analizine uygunluğunu sınamak için örneklem uygunluk ölçüsü olan Kaiser-Meyer-Olkin (KMO) ile küreselliği (eş varyanslılık) sınavan bir ki-kare testi olan Bartlett testi kullanılmıştır. Bartlett küresellik testinin anlamlı çıkması (Bartlett, 1950) ve KMO değerinin 0,50’den düşük çıkmaması beklenir (Kaiser, 1974). KMO değerlerinin yorumlanması Tablo 3.19’da sunulmuştur.

Tablo 3.19: KMO Değerlerinin Yorumlanma Ölçütleri

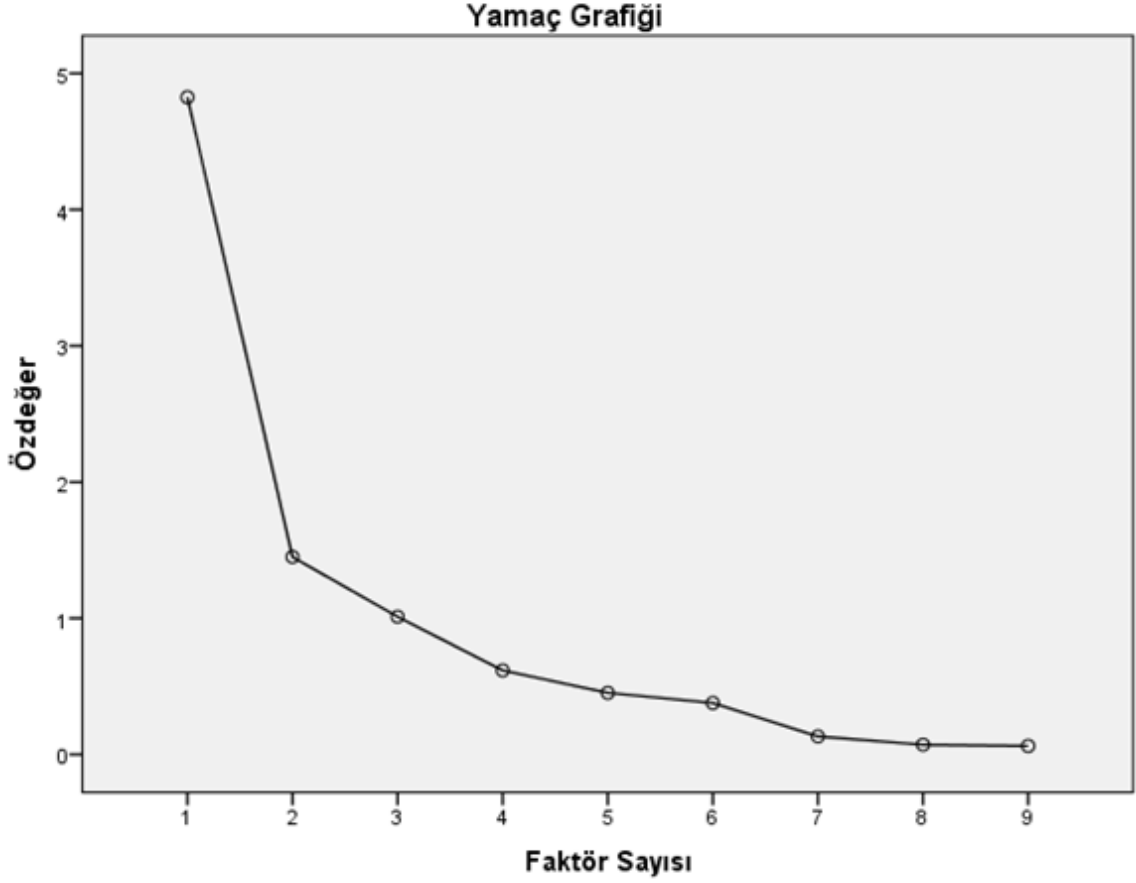
<i>KMO</i>	<i>Yorum</i>
0,90’lar içerisinde	Mükemmel (Harikulade)
0,80’ler içerisinde	Çok İyi
0,70’ler içerisinde	İyi
0,60’lar içerisinde	Orta Düzey
0,50’ler içerisinde	Zayıf
0,50’den düşük ise	Kabul Edilemez

Veri analizi sonucu KMO değeri 0,78 ile iyi ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=271,7$; $sd=36$; $p=0,00$). Bu değerler, verilerin faktör analizine uygun olduğunu göstermektedir. Gizil değişkenler arasındaki ilişkinin ortaya çıkarılmak istenmesi (Widaman, 1993), alanyazındaki uygulamalardan farklı olarak tek bir yaratıcılık indeksi elde edileceği için bir bakıma teori üretildiğinden ve eldeki veriler normal dağıldığından dolayı, AFA maksimum olabilirlik yöntemi ile yapılmıştır (Fabrigar, Wegener, MacCallum & Strahan, 1999). Ortalama ölçüt puanların korelasyon matrisine dayalı, maksimum olabilirlik yöntemiyle SPSS 24.0 programı kullanılarak uygulanan AFA sonucu elde edilen faktör sayısı, faktör öz değerleri, açıklanan varyans oranları ve maddelerin tek faktöre verdiği yükleri gösteren değerler Tablo 3.20’de verilmiştir.

Tablo 3.20: Matematik Öğretmenleri TYDT Ortalama Ölçüt Puanlarına Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplam Varyans (%)</i>	<i>Ölçüt Adı</i>	<i>Faktör Yük Değerleri</i>
1	4,83	53,61	53,61	Sözel Akıcılık	0,971
2	1,45	16,11	69,73	Sözel Esneklik	0,942
3	1,01	11,24	80,96	Sözel Orijinallik	0,947
				Şekilsel Akıcılık	0,422
				Şekilsel Orijinallik	0,719
				Erken Kapamaya Direnç	0,524
				Başlıkların Soyutluğu	0,341
				Zenginleştirme	0,437
				Bonus	0,455

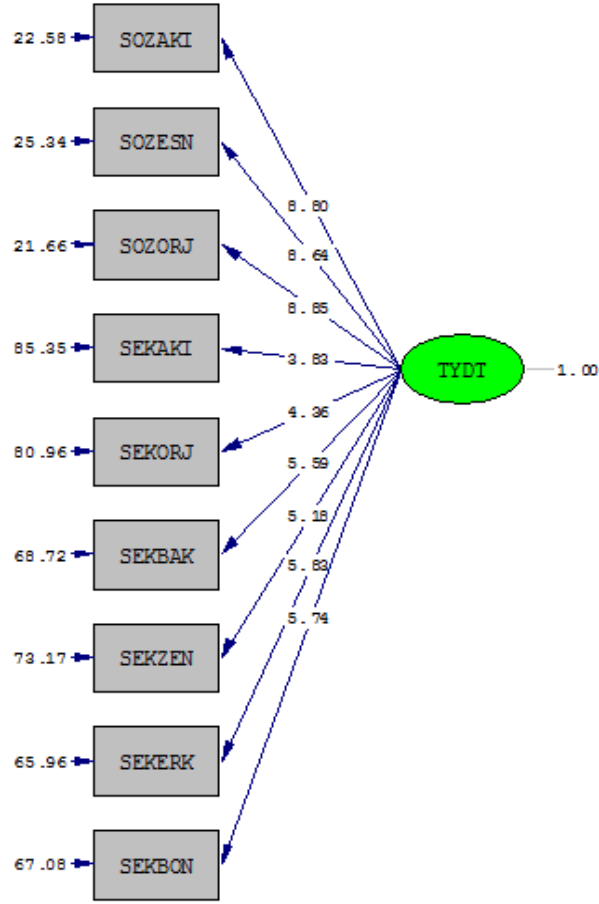
Toplam açıklanan varyansa bakıldığında ölçekte özdeğeri 1’den büyük olan 3 faktör görülmektedir. İlk faktör (özdeğeri 4,83) varyansın %53.61’ini açıklarken ikinci faktör (özdeğeri 1,45) ise varyansın %16,11’ini açıklamaktadır. Faktör sayısına karar vermek için özdeğeri 1’den büyük olan faktörler dikkate alındığında test 3 faktörlü görülmektedir. Şekil 3.2’de verilen yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir.



Şekil 3.2: Matematik Öğretmenleri TYDT Ortalama Ölçüt Puanları AFA Sonucu Elde Edilen Yamaç Grafiği

Döndürülmemiş sonuçlar için özdeğerler ve faktör yükleri bir arada dikkate alınırsa birinci ve ikinci özdeğer arasındaki farka, birinci faktörün açıklanan varyansın %54'ünü açıklamasına dayanarak ve yamaç grafiğinde birinci faktörden sonra ani düşüşün olmasından dolayı testin tek boyutlu olduğu söylenebilir. Maddelerin birinci faktör üzerindeki yükleri 0,341 ile 0,971 arasındadır. Bu değerlere dayanarak ölçüt puanlarının birlikte yaratıcılığı uygun düzeyde ölçtüğü yorumu yapılabilir. Faktör analizi sonucu uyum indeksi $\chi^2/sd=75,653/27=2,80$ elde edildiğinden, bu değer 2-3 arasında olması sebebiyle verilerin tek faktörlü modele anlamlı derecede kabul edilebilir uyum gösterdiği söylenebilir (Schermele-Engel, Moosbrugger & Müller, 2003).

Kurulan bu model Türkçe öğretmenleri TYDT ortalama ölçüt puanları üzerinden LISREL 8.80 programı kullanılarak DFA ile test edilmiş ve modelin gösterimi Şekil 3.3'de verilmiştir.



Chi-Square=80.18, df=27, P-value=0.00000, RMSEA=0.225

Şekil 3.3: TYDT'nin Faktör Yapısına İlişkin Tanımlanan Tek Faktörlü Ölçme Modeli

DFA sonucu alanyazında kullanılan birçok uyum indeksi bulunmaktadır. Kline (2015b, s.269); ki-kare testi, RMSEA, CFI ve SRMR indekslerinin raporlanmasını önermektedir. DFA'dan elde edilen χ^2/sd oranı (=80,18/27) 2,97 olduğundan verilerin modele kabul edilebilir uyum gösterdiği söylenebilir (Schermelleh-Engel vd., 2003). RMSEA ise 0,225 olarak elde edilmiştir. Hu ve Bentler'e (1999) göre RMSEA değerinin 0,06 ve daha altında olması kabul edilebilir bir model için gereklidir. Bu ölçütle kıyaslandığında modelin uyumsuz olduğuna dair iddiada bulunulabilir fakat RMSEA gerçek populasyon modellerini küçük örneklem büyüklüğünde aşma eğilimindedir (Hu & Bentler, 1999). Eldeki örneklem grubu n=40 gibi küçük bir grup olduğundan RMSEA dikkate alınmamıştır. CFI değeri

0,92 ile modelin kabul edilebilir uyum; SRMR ise 0,04 ile modelin iyi uyum gösterdiğine işarettir (Schermelleh-Engel vd., 2003). AFA ve DFA sonuçları birlikte değerlendirildiğinde; TYDT'nin tek boyutlu olduğu iddia edilebilir. Bu bulgu Heausler ve Thompson (1988) ile Clapham (1998)'in bulguları ile de uyuşmaktadır.

Yaratıcılık İndeksinin Hesaplanması

Tek boyutlu TYDT puanlarının tek bir yaratıcılık indeksi ile gösterilebilmesi için her branşta ortalama puanların ölçüt bazında standartlaştırılmasının uygun olacağı düşünülmüş ve ortalaması 50; standart sapması 10 olan T-puanına aşağıdaki formülle çevrilmiştir (Furr & Bacharach, 2013, s.47).

$$T = \frac{X - \bar{X}}{S_x} \cdot 10 + 50 \quad (\text{eşitlik 8})$$

T : T puanı

X : T puanı hesaplanacak ham puan

\bar{X} : Ham puanların aritmetik ortalaması

S_x : Ham puanların standart sapması

Çevirme işlemi sonucu elde edilen 9 ölçüt için T-puanları toplamı ile her bir öğretmen için yaratıcılık indeksleri elde edilmiştir. Araştırmada yaratıcılık değişkeninin nicel değeri olarak alınacak yaratıcılık indekslerinin betimsel istatistikleri ve S-W sonuçları Tablo 3.21'de sunulmuştur.

Tablo 3.21: Öğretmenlerin Yaratıcılık İndeksleri Betimsel İstatistikleri ve Normallik Testi Sonuçları

Branş	n	Xmin	Xmax	\bar{x}	S_x	Çarpıklık	Basıklık	S-W	p
Matematik	40	346,43	594,15	450,00	64,73	-0,034	-0,766	0,950	0,073
Türkçe	40	352,08	631,15	450,00	64,00	0,661	0,265	0,962	0,191

Tablo 3.21 incelendiğinde matematik öğretmenleri yaratıcılık indekslerinin 346,43 ile 594,15; Türkçe öğretmenlerinin ise 352,08 ile 631,15 aralığında değerler aldığı, her iki branş ortalamalarının 450 olduğu görülmektedir. Matematik öğretmenleri yaratıcılık indeksleri puan dağılımlarının çok hafif basık; Türkçe öğretmenlerinin puan dağılımları ise sağa çarpık ve sivri dağılımlı olduğu anlaşılmaktadır. Ancak bu değerlerin ± 1 aralığında bulunması nedeniyle normal dağılımdan sapmanın

önemli olmadığı söylenebilir. Matematik öğretmenlerinin yaratıcılık indekslerinin yarısından çok az fazlasının, ortalama puandan yüksek puan aldığı ve puanların öğretmenler bazında farklılaştığı; Türkçe öğretmenlerinin yaratıcılık indekslerinin yarısından fazlasının, ortalama puandan düşük puan aldığı ve puanların öğretmenler bazında farklılaşmadığı, homojen bir dağılıma sahip olduğu anlaşılmıştır. S-W sonuçlarına göre matematik (0,950; $p>0,05$) ve Türkçe (0,962; $p>0,05$) öğretmenlerine ait yaratıcılık indekslerinin normal dağılımlı oldukları görülmüş, müteakip analizlerde parametrik yöntemlerin kullanılmasına karar verilmiştir.

3.5.1.4. Yaratıcılık Düzeylerine Göre Gruplandırma

Matematik ve Türkçe öğretmenleri, hesaplanan yaratıcılık indekslerinin değerine göre büyükten küçüğe sıralanmış ve ilk 20'lik grup yaratıcılığı yüksek, sonraki 20'lik grup yaratıcılığı düşük olarak adlandırılmıştır. Oluşturulan bu iki grubun yaratıcılık indeksleri arasında anlamlı bir farklılığın olup olmadığı, bağımsız gruplar t testi ile SPSS 24.0 programı kullanılarak analiz edilmiş, sonuçları Tablo 3.22'de verilmiştir.

Tablo 3.22: Öğretmenlerin Yaratıcılık İndekslerinin Yaratıcılık Düzey Gruplarına Göre Bağımsız Gruplar t Testi Sonuçları

<i>Branş</i>	<i>Grup</i>	<i>n</i>	\bar{x}	S_x	<i>Sd</i>	<i>t</i>	<i>p</i>	<i>Cohen's d</i>
Matematik	Yaratıcılığı Yüksek	20	503,79	31,63	38	9,603	0,000	3,03
	Yaratıcılığı Düşük	20	396,21	38,85				
Türkçe	Yaratıcılığı Yüksek	20	500,57	47,99	38	8,227	0,000	2,60
	Yaratıcılığı Düşük	20	399,43	26,81				

Matematik branşında, yaratıcılığı yüksek ($\bar{x}=503,79$) ve yaratıcılığı düşük ($\bar{x}=396,21$) grupların yaratıcılık indeksleri aritmetik ortalamaları arasında fark yaratıcılığı yüksek grup lehine manidar bulunmuştur ($t_{38}=9,603$; $p<0,05$). Bu kıyaslamaya ait etki büyüklüğü 3,03 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu ve bu farkın örneklem büyüklüğünden kaynaklanmadığı söylenebilir.

Tablo 3.22 incelendiğinde, Türkçe branşında yaratıcılığı yüksek ($\bar{x}=500,57$) ve yaratıcılığı düşük ($\bar{x}=399,43$) grupların yaratıcılık indeksleri aritmetik ortalamaları arasında fark yaratıcılığı yüksek grup lehine manidar bulunmuştur ($t_{38}=8,227$;

$p < 0,05$). Bu kıyaslamaya ait etki büyüklüğü 2,60 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu söylenebilir.

Matematik ve Türkçe öğretmenlerinin yaratıcılık indekslerine göre oluşturulan yüksek ve düşük yaratıcılık gruplarının arasında anlamlı bir farklılaşma olduğu ve bu farklılığın örneklem büyüklüğünden ziyade gruptaki öğretmenlerin yaratıcılık indekslerinin nicel büyüklüklerinden kaynaklandığı görülmüştür.

Grupların tecrübe değişkeninden etkilenebileceği düşünülerek (Baş ve Beyhan, 2016); oluşturulan grupların meslekteki yılları arasında anlamlı bir farklılığın olup olmadığı yoklanmıştır. Öğretmenlerin meslekteki yılları dağılımlarının normallik varsayımını karşılamasından dolayı SPSS 24.0 programı kullanılarak yapılan bağımsız gruplar t testi sonuçları Tablo 3.23’de sunulmuştur.

Tablo 3.23: Öğretmenlerin Meslekteki Yıllarının Yaratıcılık Düzey Gruplarına Göre Normallik Testi ve Bağımsız Gruplar t Testi Sonuçları

<i>Branş</i>	<i>Grup</i>	<i>n</i>	<i>S-W</i>	<i>p</i>	\bar{x}	S_x	<i>Sd</i>	<i>t</i>	<i>p</i>
Matematik	Yaratıcılığı Yüksek	20	0,926	0,130	14,65	8,02	38	2,273	0,029
	Yaratıcılığı Düşük	20	0,930	0,152	9,45	6,35			
Türkçe	Yaratıcılığı Yüksek	20	0,955	0,447	16,05	7,25	38	0,094	0,926
	Yaratıcılığı Düşük	20	0,942	0,266	15,85	6,16			

Tablo 3.23’de görüldüğü üzere 4 grubun da mesleki yıl değişkenlerinin normal dağılımlı olduğu görülmektedir. Matematik branşında, yaratıcılığı yüksek ($\bar{x}=14,65$) ve yaratıcılığı düşük ($\bar{x}=9,45$) grupların meslekteki yılları aritmetik ortalamaları arasında anlamlı fark bulunmamıştır ($t_{38}=2,273$; $p > 0,01$). Aynı şekilde Türkçe branşında, yaratıcılığı yüksek ($\bar{x}=16,05$) ve yaratıcılığı düşük ($\bar{x}=15,85$) grupların meslekteki yılları aritmetik ortalamaları arasında anlamlı fark bulunmamıştır ($t_{38}=0,094$; $p > 0,01$). Alt problemlerin yanıtlanmasına yönelik analizlerde, grupların söz konusu özellikleri karşılaştırılırken; yaratıcılık indekslerinin, yaratıcılık değişkenini uygun şekilde temsil edeceği, yorumların meslekî yıllardan bağımsız olacağı değerlendirilmiştir.

3.5.1.5. Deney ve Kontrol Gruplarının Oluşturulması

Matematik ve Türkçe öğretmenleri, madde yazarlığı eğitimine tâbi tutulmak üzere deney ve kontrol gruplarına seçkisiz atanmışlardır. Deney grupları, matematik alanında 11 yaratıcılığı yüksek, 9 yaratıcılığı düşük; Türkçe alanında 10 yaratıcılığı

yüksek, 10 yaratıcılığı düşük öğretmenden oluşmuştur. Oluşturulan bu iki grubun yaratıcılık indeksleri arasında anlamlı bir farklılığın olup olmadığı, bağımsız gruplar t testi ile SPSS 24.0 programı kullanılarak analiz edilmiş, sonuçları Tablo 3.24'de gösterilmiştir.

Tablo 3.24: Deney ve Kontrol Gruplarına Uygulanan Bağımsız Gruplar t Testi Sonuçları

<i>Brans</i>	<i>Grup</i>	<i>n</i>	\bar{x}	S_x	<i>Sd</i>	<i>t</i>	<i>p</i>
Matematik	Deney Grubu	20	450,55	65,27	38	0,053	0,958
	Kontrol Grubu	20	449,45	65,89			
Türkçe	Deney Grubu	20	438,46	58,37	38	-1,145	0,260
	Kontrol	20	461,54	68,69			

Matematik branşında, deney ($\bar{x}=450,55$) ve kontrol ($\bar{x}=449,45$) gruplarının yaratıcılık indeksleri aritmetik ortalamaları arasında manidar fark bulunmamıştır ($t=0,053$; $p>0,05$). Aynı şekilde Türkçe branşında, deney ($\bar{x}=438,46$) ve kontrol ($\bar{x}=461,54$) gruplarının yaratıcılık indeksleri aritmetik ortalamaları arasında manidar fark bulunmamıştır ($t=-1,145$; $p>0,05$). Matematik ve Türkçe öğretmenlerinden seçkisiz oluşturulan deney ve kontrol gruplarının arasında anlamlı bir farklılaşma olmadığı görülmüştür. Bu sebeple alt problemlerin yanıtlanmasına yönelik analizlerde elde edilecek sonuçların, yaratıcılık indekslerinden bağımsız olacağı söylenebilir.

Grupların mesleki tecrübeden etkilenebileceği düşünülerek; deney ve kontrol grubundaki öğretmenlerin meslekteki yılları arasında anlamlı bir farklılığın olup olmadığı yoklanmıştır. Deney ve kontrol grubundaki öğretmenlerin meslekteki yılları dağılımlarının normallik varsayımını karşılamaından dolayı SPSS 24.0 programı kullanılarak yapılan bağımsız gruplar t testi sonuçları Tablo 3.25'de gösterilmiştir.

Tablo 3.25: Öğretmenlerin Meslekteki Yıllarının Yaratıcılık Düzey Gruplarına Göre Bağımsız Gruplar t Testi Sonuçları

<i>Brans</i>	<i>Grup</i>	<i>n</i>	<i>S-W</i>	<i>p</i>	\bar{x}	S_x	<i>Sd</i>	<i>t</i>	<i>p</i>
Matematik	Deney	20	0,913	0,071	13,90	8,48	38	1,565	0,126
	Kontrol	20	0,953	0,420	10,20	6,31			
Türkçe	Deney	20	0,945	0,296	15,65	6,91	38	-0,282	0,779
	Kontrol	20	0,965	0,565	16,25	6,53			

Tablo 3.25’de görüldüğü üzere 4 grubun da mesleki yıl değişkenlerinin normal dağılımlı olduğu görülmektedir. Matematik branşında, deney ($\bar{x}=13,90$) ve kontrol ($\bar{x}=10,20$) gruplarının meslekteki yılları aritmetik ortalamaları arasında anlamlı fark bulunmamıştır ($p>0,01$). Benzer şekilde Türkçe branşında, deney ($\bar{x}=15,65$) ve kontrol ($\bar{x}=16,25$) gruplarının meslekteki yılları aritmetik ortalamaları arasında anlamlı fark bulunmamıştır ($p>0,01$). Bu sonuçlar; yapılacak analizlerde elde edilecek bulguların yorumlanmasında; öğretmenlerin meslekteki yıllarının veya başka bir deyişle tecrübe değişkeninin etkili olmayacağına kanıt olarak sunulabilir.

3.5.2. Merkezî Sınav Eş Formları Uygulamasından Elde Edilen Verilerin İşlenmesi ve Analizi

Öğrencilerin MSEF’larına verdikleri yanıtlar, yalnız tek seçeneğin işaretlendiği doğru yanıtlar için “1”, diğer tüm yanıtlar için “0” olacak şekilde puanlanmıştır. Hiçbir yanıtı olmayan, belirli bir örüntü ile yanıtlanan cevap kağıtları ($n=76$) analizlere dahil edilmemiştir. Analizlerin örneklem büyüklüklerinden etkilenmemeleri ve Form IV ile Form V’lerin oluşturulabilmesi için matematik formları 203’er, Türkçe formları 206’sar öğrenci esas alınarak madde ve test psikometrik özellikleri hesaplanmıştır. Cevap kâğıtları form bazında en aza eşitleme yapılırken; test puanı uç değerlerde bulunan öğrencilerin cevap kâğıtlarının araştırma dışı bırakılması esas alınmıştır.

3.5.2.1. Madde Psikometrik Özelliklerinin Hesaplanması

Madde güçlük ve ayıricılık gücü indekslerinin hesaplanmasında düzeltilmiş nokta çift serili korelasyon tekniğini esas alan TAP 16.11.13 programı; madde çeldiricilik indekslerinin hesaplanmasında ise eşitlik (5)’deki formülün girildiği Microsoft Office 2007 Excel programı kullanılmıştır. 5 forma ait madde güçlük, ayıricılık gücü, çeldiricilik indekslerinden matematik alanına ait olanlar Tablo 3.26’da; Türkçe alanına ait olanlar Tablo 3.27’de sunulmuştur.

Tablo 3.26: Matematik Eş Formları Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri

Madde	p_i	r_{jx}	ζ	p_i	r_{jx}	ζ	p_i	r_{jx}	ζ	p_i	r_{jx}	ζ	p_i	r_{jx}	ζ
	Form I			Form II			Form III			Form IV			Form V		
1	0,73	0,25	0,61	0,71	0,31	0,71	0,53	0,28	0,93	0,71	0,23	0,71	0,73	0,22	0,61
2	0,43	0,30	0,70	0,50	0,36	0,66	0,43	0,40	0,98	0,50	0,29	0,66	0,43	0,20	0,70
3	0,67	0,29	0,81	0,61	0,28	0,63	0,50	0,23	0,92	0,61	0,21	0,63	0,67	0,22	0,81
4	0,75	0,21	0,71	0,50	0,41	0,70	0,47	0,09	0,85	0,50	0,33	0,70	0,75	0,10	0,71
5	0,67	0,30	0,72	0,35	0,20	0,74	0,47	0,35	0,78	0,67	0,18	0,72	0,35	0,15	0,74
6	0,75	0,29	0,65	0,78	0,12	0,71	0,48	0,21	0,80	0,75	0,05	0,65	0,78	0,09	0,71
7	0,39	0,29	0,85	0,79	0,28	0,62	0,49	0,33	0,94	0,39	0,26	0,85	0,79	0,06	0,62
8	0,65	0,22	0,73	0,80	0,22	0,73	0,51	0,23	0,97	0,65	0,09	0,73	0,80	0,03	0,73
9	0,33	0,29	0,77	0,79	0,17	0,44	0,49	0,41	0,75	0,33	0,17	0,77	0,79	0,12	0,44
10	0,42	0,26	0,66	0,62	0,29	0,61	0,47	0,30	0,79	0,62	0,20	0,61	0,42	0,17	0,66
11	0,29	0,24	0,72	0,54	0,36	0,63	0,51	0,37	0,89	0,54	0,29	0,63	0,29	0,19	0,72
12	0,62	0,27	0,67	0,85	0,22	0,71	0,51	0,35	0,91	0,85	0,22	0,71	0,62	0,15	0,67
13	0,70	0,23	0,78	0,60	0,28	0,61	0,53	0,32	0,86	0,60	0,28	0,61	0,70	0,14	0,78
14	0,51	0,26	0,66	0,51	0,29	0,64	0,49	0,29	0,87	0,51	0,11	0,66	0,51	0,07	0,64
15	0,44	0,25	0,75	0,78	0,40	0,67	0,51	0,23	0,93	0,44	0,05	0,75	0,78	0,21	0,67
16	0,48	0,36	0,75	0,64	0,35	0,68	0,46	0,38	0,89	0,64	0,24	0,68	0,48	0,26	0,75
17	0,75	0,28	0,68	0,61	0,43	0,64	0,48	0,38	0,92	0,61	0,32	0,64	0,75	0,17	0,68
18	0,52	0,30	0,67	0,70	0,37	0,74	0,51	0,30	0,95	0,70	0,24	0,74	0,52	0,19	0,67
19	0,80	0,24	0,79	0,38	0,28	0,60	0,46	0,37	0,90	0,80	0,20	0,79	0,38	0,16	0,60
20	0,76	0,21	0,65	0,46	0,12	0,69	0,48	0,48	0,84	0,76	0,04	0,65	0,46	0,04	0,69

Tablo 3.26'da görüldüğü üzere Form I'de 4, Form II'de 7 maddenin ayırıcılık gücü indeksleri 0,30 değeri ve üzerindedir. Bu sonuçlar öğretmenlerin ölçme ve değerlendirme konusunda kendilerini yetersiz görmelerine yönelik görüşlerini ortaya koyan araştırma sonuçlarını destekler niteliktedir (Baş ve Beyhan, 2016; Çakan, 2004; Şenel, 2008).

Tablo 3.27: Türkçe Eş Formları Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri

Madde	p_i	r_{jx}	Ç	p_i	r_{jx}	Ç	p_i	r_{jx}	Ç	p_i	r_{jx}	Ç	p_i	r_{jx}	Ç
	Form I			Form II			Form III			Form IV			Form V		
1	0,69	0,37	0,82	0,72	0,29	0,62	0,34	0,29	0,94	0,69	0,22	0,82	0,72	0,37	0,62
2	0,74	0,23	0,67	0,75	0,23	0,87	0,35	0,24	0,92	0,75	0,12	0,87	0,74	0,20	0,67
3	0,64	0,29	0,94	0,78	0,23	0,60	0,41	0,39	0,90	0,64	0,16	0,94	0,78	0,20	0,60
4	0,63	0,28	0,65	0,69	0,34	0,66	0,36	0,36	0,95	0,69	0,19	0,66	0,63	0,31	0,65
5	0,50	0,27	0,68	0,61	0,25	0,68	0,48	0,40	0,90	0,61	0,10	0,68	0,50	0,33	0,68
6	0,86	0,20	0,70	0,47	0,43	0,65	0,37	0,28	0,91	0,47	0,30	0,65	0,86	0,21	0,70
7	0,91	0,15	0,67	0,67	0,29	0,73	0,18	0,56	0,91	0,91	0,05	0,67	0,67	0,33	0,73
8	0,61	0,17	0,68	0,80	0,31	0,67	0,63	0,37	0,70	0,61	0,00	0,68	0,80	0,34	0,67
9	0,84	0,23	0,71	0,42	0,35	0,57	0,66	0,31	0,86	0,42	0,23	0,57	0,84	0,27	0,71
10	0,69	0,22	0,59	0,89	0,40	0,60	0,46	0,31	0,91	0,69	0,09	0,59	0,89	0,43	0,60
11	0,81	0,09	0,70	0,58	0,21	0,69	0,44	0,43	0,90	0,58	0,03	0,69	0,81	0,11	0,70
12	0,88	0,28	0,49	0,67	0,35	0,72	0,61	0,29	0,81	0,67	0,14	0,72	0,88	0,30	0,49
13	0,52	0,43	0,78	0,33	0,22	0,69	0,83	0,18	0,89	0,52	0,20	0,78	0,33	0,26	0,69
14	0,78	0,21	0,64	0,75	0,28	0,76	0,22	0,26	0,95	0,75	0,25	0,76	0,78	0,29	0,64
15	0,63	0,31	0,73	0,42	0,29	0,69	0,42	0,31	0,89	0,42	0,11	0,69	0,63	0,37	0,73
16	0,82	0,20	0,75	0,63	0,22	0,64	0,63	0,35	0,78	0,82	0,04	0,75	0,63	0,28	0,64
17	0,76	0,28	0,76	0,41	0,26	0,73	0,22	0,28	0,74	0,41	0,19	0,73	0,76	0,36	0,76
18	0,58	0,23	0,71	0,68	0,17	0,61	0,34	0,28	0,67	0,58	0,07	0,71	0,68	0,18	0,61
19	0,43	0,31	0,69	0,74	0,26	0,73	0,58	0,38	0,82	0,43	0,18	0,69	0,74	0,27	0,73
20	0,67	0,33	0,68	0,81	0,24	0,65	0,59	0,28	0,80	0,67	0,18	0,68	0,81	0,30	0,65

Tablo 3.27’de görüldüğü üzere Form I’de 5, Form II’de 6 maddenin ayırıcılık güçleri 0,30 değeri ve üzerindedir. Bu sonuçlar öğretmenlerin ölçme aracı olarak ÇSM geliştirebilme yeteneklerinin yetersizliğine işaretler ve matematik alanında elde edilen sonuçlarla paraleldir.

3.5.2.2. Test Psikometrik Özelliklerinin Hesaplanması

Öğrencilerin eş formlara verdikleri yanıtlar TAP 16.11.13 programına girilerek; formlara ait KR-20 güvenilirlikleri, varyansları, ortalamaları; Microsoft Office 2007 Excel programına girilerek kapsam geçerlik indeksleri; MOAS puanları ile birlikte SPSS 24.0 programına girilerek uyum geçerliğine dönük kanıt olarak PMÇK katsayıları; Factor 10.4 programına girilerek AFA sonucu elde edilen özdeğerleri ve döndürülmemiş madde faktör yükleri elde edilmiştir. Aynı kazanımı ölçen maddelerle oluşturulan testlerin, MOAS puanları benzer düzeyde olan, bir başka deyişle başarı bakımından benzer gruplara aynı koşullarda uygulanması sonucunda benzer sonuçların elde edilmesi beklenir. Benzer düzeydeki özelliklerin

ölçülmesi sonucu farklı ölçme sonuçlarının elde edilmesi, ölçme aracının hatası veya hatalarından kaynaklanacaktır. Bu düşünceden hareketle, madde yazarı öğretmenlerin geliştirdikleri maddelerle geliştirilen testlerin, başarı durumları arasında anlamlı farklılıkları olmayan öğrencilere uygulanması ile elde edilecek puanların ve bu puanlarla kestirilen özelliklerin benzeşmesi başka bir ifade ile bu testlerin paralel olması beklenir. İki testin paralel kabul edilebilmesi için ise bu testlerin;

- Madde sayıları, güvenilirlikleri, varyansları, ortalamaları eşit olmalıdır,
- Aynı tipte maddelerden oluşmalıdır (Gulliksen, 1950, s.11),
- Aynı kapsamı ya da yapıyı ölçmelidir,
- Her iki test de içerdiği maddelerle ölçülecek yapıyı eşit şekilde ölçmelidir (Geisinger vd., 2013).

Yukarıda sıralı varsayımlar esas alınarak, aynı tipte (çoktan seçmeli) eşit maddeli (n=20) eş formların karşılaştırılmasında; güvenilirlikleri, varyansları, ortalamaları, kapsam geçerlikleri, uyum geçerlikleri, yapı geçerlikleri (boyutları ve faktör analizi sonucu elde edilen döndürülmemiş madde faktör yükleri) esas alınmıştır. Söz konusu istatistiklerin anlamlı derecede farklılaşıp farklılaşmadığı yoklanarak alt problemlere yanıt aranmıştır.

3.5.2.3. Öğrenci MOAS Puanlarının İşlenmesi ve Analizi

2016-2017 eğitim ve öğretim yılı 1.dönem MOAS alt testlerinin ne derecede güvenilir ve geçerli ölçme araçları olduğuna dönük yapılan literatür taraması sonucu, bu araştırmanın süresi dahilinde alt testler bazında puan ortalamaları (matematik için 48,61; Türkçe için 54,56) haricinde herhangi bir analiz bulgusuna ulaşılamamıştır. Geçmiş yıllara ait yapılan taramalarda ise iç tutarlık güvenilirliği 2015-2016 eğitim ve öğretim yılı 1.dönem matematik alt testi için 0,870; Türkçe alt testi için 0,837 (MEB, 2016b); 2014-2015 eğitim ve öğretim yılı 1.dönem matematik alt testi için 0,868; Türkçe alt testi için 0,853 olarak açıklanmıştır (MEB, 2015b). Bu değerlere dayanarak MOS alt testlerinin güvenilir ölçme araçları olduğu kabul edilerek araştırma sürdürülmüştür.

Öğrencilerin MOAS matematik ve Türkçe alt testlerinden aldıkları puanların, cevapladıkları eş formlara göre farklılaşıp farklılaşmadığını belirlemek amacıyla

ilişkisiz örneklemeler için tek faktörlü varyans analizi (one-way anova) yapılmıştır. Bu analizden önce analizin varsayımları kontrol edilmiştir. Eşit aralık ölçeğinde ve ilişkisiz öğrenci gruplarından elde edilen MOAS puanlarının; grup büyüklüklerinin 50'den büyük olması sebebiyle normallik Kolmogorov-Smirnov Normallik Testi (K-S) ile ve varyanslarının eşit olup olmadığına yönelik yapılan Levene F testi ile yapılan analiz sonuçları Tablo 3.28'de verilmiştir.

Tablo 3.28: Öğrenci MOAS Puanları Normallik ve Homojenlik Testi Sonuçları

Öğrenci Grubu	n	\bar{X}	S_x	K-S		Levene	
				İstatistik	p	İstatistik	P
Matematik FI	203	10,53	4,563	0,055	0,200	0,000	1,000
Matematik FII	203	10,51	4,558	0,056	0,200		
Matematik FIII	203	10,49	4,550	0,058	0,094		
Türkçe FI	206	11,11	4,164	0,060	0,066	0,001	0,999
Türkçe FII	206	11,09	4,139	0,061	0,060		
Türkçe FIII	206	11,08	4,140	0,059	0,080		

Tablo 3.28'de görüleceği üzere; öğrencilerin MOAS matematik ve Türkçe alt test puanlarının uygulanan her bir "eş formdaki dağılımı normal dağılımdan anlamlı farklılık göstermez" şeklinde kurulan H_0 hipotezleri K-S ile kabul edilmiştir. Her bir eş form alan öğrenci grubunun MOAS alt test puanları normal dağılmaktadır. Matematik ve Türkçe alanlarındaki 3'er form gruplarının homojenliği için uygulanan Levene testi sonuçlarına göre varyansları anlamlılık düzeyi sınırları içerisinde eşit görülmüştür ($p > 0,05$). ANOVA'nın dört varsayımı karşılandığından, aynı alanda eş form yanıtlayan 8.sınıf öğrencilerinin MOAS puanlarının anlamlı şekilde farklılaşp farklılaşmadığını yoklayan analiz sonuçları Tablo 3.29'da sunulmuştur.

Tablo 3.29: Öğrenci MOAS Puanlarının Uygulanan Eş Formlara Göre ANOVA Sonuçları

Alan	Varyansın Kaynağı	Kareler Toplamı	Sd	Kareler Ortalaması	F	p
Matematik	Gruplararası	0,122	2	0,061	0,003	0,997
	Gruplarıçi	12600,059	606	20,792		
	Toplam	12600,181	608			
Türkçe	Gruplararası	0,204	2	0,102	0,006	0,994
	Gruplarıçi	10575,078	615	17,195		
	Toplam	10575,282	617			

Analiz sonuçları, öğrencilerin MOAS matematik alt test puanları arasında, uygulanan eş form bakımından anlamlı bir farkın olmadığını göstermektedir ($F_{2,606}=0,003$; $p > 0,05$). Aynı şekilde Türkçe alt testi puanları arasında da, uygulanan eş form bakımından anlamlı bir farkın olmadığı görülmektedir

($F_{2,615}=0,006$; $p>0,05$). Bu sonuçlara göre, alt problemlere yanıt aranan analizlerde söz konusu gruplardan elde edilecek verilerin grupların alan başarı farkından bağımsız olacağı söylenebilir.

3.5.3. MUF Uygulamasından Elde Edilen Verilerin İşlenmesi ve Analizi

Matematik ve Türkçe öğretmenlerince gerek araştırmanın başlangıcında gerekse madde yazarlığı eğitimi sonrası geliştirdikleri maddelerle oluşturulan testlerin kapsam geçerliklerini incelemek amacıyla matematik alanında 14, Türkçe alanında 12 uzmanının görüşlerinden elde edilen veriler kullanılmıştır. Uzmanlardan gelen dönütlerle her bir madde için Lawshe (1975) tarafından geliştirilen Kapsam Geçerlik Oranı (KGO) elde edilmiştir. KGO elde edilirken her bir madde için toplam kaç uzmanın “gerekli” seçeneğini işaretlediği esas alınmıştır. KGO’larının kritik değerlerine göre maddelerin KGO ortalaması ile Kapsam Geçerlik İndeksi (KGI) hesaplanmıştır. Ayre ve Scally (2014) tarafından yapılan çalışmada, 10 üzeri uzman görüşünde Lawshe (1975) tarafından belirlenen kritik değerlerin dağılımsal hatası giderilerek yeni bir kritik değer tablosu oluşturulmuştur. Oluşturulan tabloya göre kritik değerler; 0,05 anlamlılık düzeyinde, 14 uzman görüşü alınan matematik alanında KGO için 0,571 (asgari 11 uzman), 12 uzman görüşü alınan Türkçe alanında ise 0,667 (asgari 10 uzman) olarak belirlendiğinden, çalışmada bu değerler ölçüt alınmıştır.

3.5.4. Alt Problemlerin Yanıtlanmasına Yönelik Analizler

Alt problemlerin yanıtlanması için yapılan analizlere ait; değişkenler, verilerin nitelikleri, analiz tekniği, tekniğe ait varsayımlar, yorum ölçütleri, kullanılan analiz programı ve anlamlı farklılaşma elde analizler için etki büyüklükleri Tablo 3.30’da sunulmuştur. Tüm farklılaşmaların yorumlanmasında $\alpha = 0,01$ manidarlık düzeyi esas alınmıştır. Analiz programları vasıtasıyla elde edilen sonuçların ne derece doğru olduğundan emin olabilmek adına, her analiz için %5’i kadar işlem araştırmacı tarafından elle hesaplanarak kontrol edilmiştir.

Tablo 3.30: Alt Problemlerin Yanıtlanmasında Kullanılan Analizler

BAGIMSIZ/BIRINCIL DEĞİŞKEN	VERİ NİTELİK(LER)i	BAGIMLI/İKİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)i	ANALİZ / TEKNİK	TEKNİĞİN VARSAYIM(LAR)I	YORUM ÖLÇÜTLERİ	ANALİZ PROGRAMI	ETKİ BÜYÜKLÜĞÜ
1.Öğretmenlerin yaratıcılık düzeyleri ile geliştirdikleri ÇSM ve testlerin psikometrik özellikleri arasında anlamlı bir ilişki var mıdır?								
a. Matematik ve Türkçe Öğretmenlerine Ait Yaratıcılık İndeksleri ile Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasında Anlamlı Bir İlişki Var mıdır?								
Yaratıcılık İndeksi (Matematik ve Türkçe Öğretmenleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Madde Güçlük İndeksi (Form I, II maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Korelasyon Analizi/ PMÇK (Pearson, 1895)	* İki değişken de sürekli * İki değişken de normal dağılımlı (Büyüköztürk, 2015, s.31)	Korelasyon katsayısının mutlak değeri; 0,00-0,29 arasında ise düşük; 0,30-0,69 arasında ise orta; 0,70 ile 1,00 arasında ise yüksek düzeyde ilişki vardır (Büyüköztürk, 2015, s.32)	SPSS 24.0	-
		Madde Ayırıcılık Gücü İndeksi (Form I, II maddeleri)						-
		Madde Çeldiricilik İndeksi (Form I, II maddeleri)						-
Madde Güçlük İndeksi (Form I maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Oran İstatistiği	Madde Güçlük İndeksi (Form II maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Oran İstatistiği	Farklılık/ İlişkisiz Zoran testi	* İki değişken de oran istatistiği * İki değişken ilişkisiz (Baykul, 1999, s.329)	α= 0,01 anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Baykul, 1999, s.274).	çevrimiçi program (http://www.socscistatistics.com/tests/ztest/Default2.aspx)	-
Madde Çeldiricilik Gücü İndeksi (Form I maddeleri)		Madde Çeldiricilik İndeksi (Form II maddeleri)						-
Madde Ayırıcılık Gücü İndeksi (Form I maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Korelasyon İstatistiği	Madde Ayırıcılık Gücü İndeksi (Form II maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Korelasyon İstatistiği	Farklılık/ İlişkisiz Z _{korelasyon} testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)		çevrimiçi program (http://vassarstats.net/rdiff.html)	-

BAĞIMSIZ/BİRİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)İ	BAĞIMLI/İKİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)İ	ANALİZ / TEKNİK	TEKİNİN VARSAYIM(LAR)I	YORUM ÖLÇÜTLERİ	ANALİZ PROGRAMI	ETKİ BÜYÜKLÜĞÜ
b. Matematik ve Türkçe Öğretmenlerinin Geliştirdikleri Maddelerle Oluşturulan Testlerin Güvenirlikleri, Varyansları, Aritmetik Ortalamaları, Kapsam Geçerlikleri, Uyum Geçerlikleri Yaratıcılık Düzeylerine Göre Anlamlı Bir Şekilde Farklılaşmakta Mıdır?								
İç Tutarlık Güvenirliği (KR-20) (Form I)	Korelasyon İstatistiği	İç Tutarlık Güvenirliği (KR-20) (Form II)	Korelasyon İstatistiği	Farklılık/ İlişkisiz $Z_{korelasyon}$ testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Baykul, 1999, s.274).	çevrimiçi program (http://vassarstat.s.net/rdiff.html)	-
Test Varyansı (Form I)	Normal Dağılımlı	Test Varyansı (Form II)	Normal Dağılımlı	Farklılık/ İki Varyansın Oranının Test Edilmesinde Kullanılan F Testi	* İlişkisiz varyanslar * Varyansların elde edildiği puanlar normal dağılımlı (Baykul, 1999, s.402)	Varyansların oranının F tablosundaki alt ve üst değer aralığında olması yorumlanır (Baykul, 1999, s.403).	-	-
Aritmetik Ortalama (Form I)	* Normal Olmayan Dağılımlı * Sürekli * Eşit Aralıklı	Aritmetik Ortalama (Form II)	* Normal Olmayan Dağılımlı * Sürekli * Eşit Aralıklı	Farklılık/ İlişkisiz Ölçümler için Mann Whitney U testi	* İlişkisiz örneklemelerden elde edilmiş ortalamalar * En az sıralama ölçeğinde (Büyüköztürk, 2015, s.165)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Büyüköztürk, 2015, s.171)	SPSS 24.0	Cohen's d
Kapsam Geçerliği (KGO/KGİ) (Form I)	Oran İstatistiği	Kapsam Geçerliği (KGO/KGİ) (Form II)	Oran İstatistiği	Farklılık/ İlişkisiz Z_{oran} testi	* İki değişken de oran istatistiği * İki değişken ilişkisiz (Baykul, 1999, s.329)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Büyüköztürk, 2015, s.171)	çevrimiçi program (http://www.socscistatistics.com/ests/ztest/Default2.aspx)	-
Uyum Geçerliği (MOAS Puanları ile Korelasyon Katsayısı) (Form I)	Korelasyon İstatistiği	Uyum Geçerliği (MOAS Puanları ile Korelasyon Katsayısı) (Form II)	Korelasyon İstatistiği	Farklılık/ İlişkisiz $Z_{korelasyon}$ testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Büyüköztürk, 2015, s.171)	çevrimiçi program (http://vassarstat.s.net/rdiff.html)	-
c. Yaratıcılığı Yüksek ve Düşük Madde Matematik ve Türkçe Öğretmenlerinin Geliştirdikleri Maddelerle Oluşturulan Testlerin Faktör Analizi Sonuçları Nasıldır?								
Yapı Geçerlikleri (Madde Faktör Yükleri) (Form I)	Korelasyon İstatistiği	Yapı Geçerlikleri (Madde Faktör Yükleri) (Form II)	Korelasyon İstatistiği	Farklılık/ İlişkisiz $Z_{korelasyon}$ testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Büyüköztürk, 2015, s.171)	çevrimiçi program (http://vassarstat.s.net/rdiff.html)	-

BAĞIMSIZ/BİRİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)i	BAĞIMLI/İKİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)i	ANALİZ / TEKNİK	TEKNİĞİN VARSAYIM(LAR)ı	YORUM ÖLÇÜTLERİ	ANALİZ PROGRAMI	ETKİ BÜYÜKLÜĞÜ
2.Madde Yazarlığı Eğitimi Alan Öğretmenlerin, Eğitim Öncesi Geliştirdikleri ÇSM ve Testlerin Psikometrik Özellikleri, Eğitim Sonrası Anamlı Bir Şekilde Farklılaşmakta Mıdır?								
a. Matematik ve Türkçe Öğretmenlerinin Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü, Çeldiricilikleri Eğitim Sonrası ve Öncesine Göre Anamlı Bir Şekilde Farklılaşmakta Mıdır?								
Madde Güçlük İndeksi (Form III maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Oran İstatistiği	Madde Güçlük İndeksi (Form V maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Oran İstatistiği	Farklılık/ İlişkisiz Z _{oran} testi	* İki değişken de oran istatistiği * İki değişken ilişkisiz (Baykul, 1999, s.329)	α= 0,01 anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Baykul, 1999, s.274).	çevrimiçi program (http://www.socscistatistics.com/tests/ztest/Default2.aspx)	-
Madde Çeldiricilik İndeksi (Form III maddeleri)		Madde Çeldiricilik İndeksi (Form V maddeleri)					-	
Madde Ayırıcılık Gücü İndeksi (Form III maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Korelasyon İstatistiği	Madde Ayırıcılık Gücü İndeksi (Form V maddeleri)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı * Korelasyon İstatistiği	Farklılık/ İlişkisiz Z _{korelasyon} testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)		çevrimiçi program (http://vassarstats.net/rdiff.html)	-
Madde Güçlük İndeksleri Ortalaması (Form III)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Madde Güçlük İndeksleri Ortalaması (Form V)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Farklılık/ Bağımsız Gruplar t Testi	* Veriler en az aralık ölçeğinde * Normal Dağılımlı * Gruplar İlişkisiz (Büyüköztürk, 2015, s.39)		SPSS 24.0	Cohen's d
Madde Ayırıcılık Gücü İndeksleri Ortalaması (Form III)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Madde Ayırıcılık Gücü İndeksleri Ortalaması (Form V)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı					
Madde Çeldiricilik İndeksleri Ortalaması (Form III)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Madde Çeldiricilik İndeksleri Ortalaması (Form V)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı					

BAĞIMSIZ/BİRİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)İ	BAĞIMLI/İKİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)İ	ANALİZ / TEKNİK	TEKİNİN VARSAYIM(LAR)I	YORUM ÖLÇÜTLERİ	ANALİZ PROGRAMI	ETKİ BÜYÜKLÜĞÜ
b. Matematik ve Türkçe Öğretmenlerinin Geliştirdikleri Maddelerle Oluşturulan Testlerin Güvenirlikleri, Varyansları, Aritmetik Ortalamaları, Kapsam Geçerlikleri, Uyum Geçerlikleri Eğitim Sonrası ve Öncesine Göre Anlamlı Bir Şekilde Farklaşmakta Mıdır?								
İç Tutarlık Güvenirliği (KR-20) (Form III)	Korelasyon İstatistiği	İç Tutarlık Güvenirliği (KR-20) (Form V)	Korelasyon İstatistiği	Farklılık/ İlişkisiz $Z_{korelasyon}$ testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Baykul, 1999, s.274).	çevrimiçi program (http://vassarstat.s.net/rdiff.html)	Cohen's q
Test Varyansı (Form III)	Normal Dağılımlı	Test Varyansı (Form V)	Normal Dağılımlı	Farklılık/ İki Varyansın Oranının Test Edilmesinde Kullanılan F Testi	* İlişkisiz varyanslar * Varyansların elde edildiği puanlar normal dağılımlı (Baykul, 1999, s.402)	Varyansların oranının F tablosundaki alt ve üst değer aralığında olması yorumlanır (Baykul, 1999, s.403).	-	Cohen's d
Aritmetik Ortalama (Form III)	* Normal Olmayan Dağılımlı * Sürekli * Eşit Aralıklı	Aritmetik Ortalama (Form V)	* Normal Olmayan Dağılımlı * Sürekli * Eşit Aralıklı	Farklılık/ İlişkisiz Ölçümler için Mann Whitney U testi	* İlişkisiz örneklemelerden elde edilmiş ortalamalar * En az sıralama ölçeğinde (Büyüköztürk, 2015, s.165)		SPSS 24.0	Cohen's d
Kapsam Geçerliği (KGO/KGİ) (Form III)	Oran İstatistiği	Kapsam Geçerliği (KGO/KGİ) (Form V)	Oran İstatistiği	Farklılık/ İlişkisiz Z_{oran} testi	* İki değişken de oran istatistiği * İki değişken ilişkisiz (Baykul, 1999, s.329)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Büyüköztürk, 2015, s.171)	çevrimiçi program (http://www.socscistatistics.com/ests/ztest/Default2.aspx)	-
Uyum Geçerliği (MOAS Puanları ile Korelasyon Katsayısı) (Form III)	Korelasyon İstatistiği	Uyum Geçerliği (MOAS Puanları ile Korelasyon Katsayısı) (Form V)	Korelasyon İstatistiği	Farklılık/ İlişkisiz $Z_{korelasyon}$ testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)		çevrimiçi program (http://vassarstat.s.net/rdiff.html)	Cohen's q
c. Madde Yazarlığı Eğitimi Alan Matematik ve Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerle Oluşturulan Testlerin Faktör Analizi Sonuçları Nasıldır?								
Yapı Geçerlikleri (Madde Faktör Yükleri) (Form I)	Korelasyon İstatistiği	Yapı Geçerlikleri (Madde Faktör Yükleri) (Form II)	Korelasyon İstatistiği	Farklılık/ İlişkisiz $Z_{korelasyon}$ testi	* İki değişken de korelasyon istatistiği * İki değişken ilişkisiz (Eid, Gollwitzer & Schmidt, 2011, s.112)	$\alpha = 0,01$ anlamlılık düzeyinde farklılık olup olmadığına göre yorumlanır (Büyüköztürk, 2015, s.171)	çevrimiçi program (http://vassarstat.s.net/rdiff.html)	-

BAĞIMSIZ/BİRİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)İ	BAĞIMLI/İKİNCİL DEĞİŞKEN	VERİ NİTELİK(LER)İ	ANALİZ / TEKNİK	TEKNIĞIN VARSAYIM(LAR)I	YORUM ÖLÇÜTLERİ	ANALİZ PROGRAMI	ETKİ BÜYÜKLÜĞÜ
3.Yaratıcılık ve Madde Yazarığı Eğitimi ÇSM Psikometrik Özelliklerinin Anlamli Birer Yordayıcısı Mıdır?								
a. Matematik ve Türkçe Öğretmenlerinin Yaratıcılık İndeksleri ve Madde Yazarığı Eğitim Durumları Geliştirdikleri Maddelerin Güçlük İndekslerinin Anlamli Birer Yordayıcısı Mıdır?								
* Yaratıcılık İndeksi * Madde Yazarığı Durumu	Sürekli Kesikli (1-0 ile yapay sürekli) (Büyüköztürk, 2015, s.92)	Madde Güçlük İndeksi (Form III, IV)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Çoklu Regresyon (Standart Regresyon)	* Veriler en az aralık ölçeğinde, * Veriler sürekli, * Değişkenler arası doğrusal ilişkili, * Çok değişkenli normal dağılımlı olmalıdır. * Uç değerler, * Çoklu bağlantı (multi-collinearity), * Tolerans değeri 0,1'den küçük, * Varyans Şişkinlik Değeri 5'den büyük olmamalıdır. Büyüköztürk, 2015, ss.98-100)	Anlamli yordayıcıların varyans açıklama oranları esas alınır. Büyüköztürk, 2015, ss.98-100)	SPSS 24.0	-
b. Matematik ve Türkçe Öğretmenlerinin Yaratıcılık İndeksleri ve Madde Yazarığı Eğitim Durumları Geliştirdikleri Maddelerin Ayırıcılık Gücü İndekslerinin Anlamli Birer Yordayıcısı Mıdır?								
* Yaratıcılık İndeksi * Madde Yazarığı Durumu	Sürekli Kesikli (1-0 ile yapay sürekli) (Büyüköztürk, 2015, s.92)	Madde Ayırıcılık Gücü İndeksi (Form III, IV)	* Sürekli * Eşit Aralıklı * Normal Dağılımlı	Çoklu Regresyon (Standart Regresyon)	* Veriler en az aralık ölçeğinde, * Veriler sürekli, * Değişkenler arası doğrusal ilişkili, * Çok değişkenli normal dağılımlı olmalıdır. * Uç değerler, * Çoklu bağlantı (multi-collinearity), * Tolerans değeri 0,1'den küçük, * Varyans Şişkinlik Değeri 5'den büyük olmamalıdır. Büyüköztürk, 2015, ss.98-100)	Anlamli yordayıcıların varyans açıklama oranları esas alınır. Büyüköztürk, 2015, ss.98-100)	SPSS 24.0	-

3.6. Araştırmanın İç ve Dış Geçerliliği

Araştırmanın başlangıcında güvenilir ve geçerli bulgular elde etmek amacıyla bağımlı değişkenlere (madde ve test psikometrik özellikleri) ait varyansların kontrol edilebileceği bir desen planlanmıştır. Bağımlı değişkenler üzerinde etkisi araştırılan bağımsız değişkeninin (madde yazarlığı eğitimi) sebep olduğu varyansı artıran, dışsal değişkenlerden kaynaklanabilecek varyansı (meslekî tecrübe) kontrol edebilen bir araştırma deseni kurgulanmıştır.

3.6.1. Araştırmanın İç Geçerliliği

İç geçerlik, bağımlı değişkende gözlenen değişmelerin, bağımsız değişkenle açıklanabilirlik derecesidir (Büyüköztürk vd., 2016, s.174). Araştırmada madde yazarlığı eğitimine tâbi tutulan öğretmenlerin, eğitim öncesi ve sonrası geliştirdikleri maddelerin psikometrik özelliklerindeki değişimin, madde yazarlığı eğitimi dışındaki başka değişkenlerden etkilenip etkilenmediği irdelenmiştir. İç geçerliği tehdit eden faktörler ve bu faktörlere alınmış olan tedbirler aşağıda sıralanmış ve açıklanmıştır (Büyüköztürk vd., 2016, s.175).

- (1)*deneklerin seçimi*, (2)*deneklerin geçmişi*, (3)*istatistiksel regresyon*, (4)*etkileşim etkisi*; öğretmenler gruplara yansız atanarak 4 tehlide tedbir üretilmiştir.

- (5)*deneklerin olgunlaşması*; öğretmenlerin madde yazarlığı eğitimi öncesi ve sonrası geliştirdikleri maddelerin psikometrik özelliklerindeki değişimlemeyi açıklayacak, başka bir öğrenme veya olgunlaşma gözlemlenmemiş veya tespit edilmemiştir.

- (6)*veri toplama aracı*; gerek TYDT gerekse eş formlar, tek elden ve araştırmacı tarafından aynı esaslarda uygulanmış, farklı hareket tarzlarından sakınılmış, böyle hareket edilerek de bu tehlide tedbir üretilmiştir.

- (7)*denek kaybı etkisi*; araştırma sürecinde öğretmen ve öğrencilerden çalışmadan ayrılan olmadığından bu tehdit belirmemiştir.

- (8)*öntest etkisi*; araştırmada madde yazarlığı eğitimi öncesi öğretmenlerden bir kazanımı yoklayan madde geliştirmesi, eğitim sonrası ise aynı kazanımdan bir madde daha geliştirmesi istenmiştir. Madde yazarlığı eğitimi sonrasında öncesi ile aynı kazanımda madde geliştirmenin bir alışıklık

yaratabileceđi ve bu durumun eđitim sonrası lehine avantaj sađlayacađı iddia edilebilir. Bu tehdidin madde ve test psikometrik özelliklerinde etkisi olmuş olabileceđi düşünölmektedir.

- (9)*beklentilerin etkisi*; öđretmenlere aldıkları TYDT puanı ile geliřtirdikleri maddelerin psikometrik özellikleri söylenmemiřtir. Arařtırmanın bařlangıcında ve eđitim sonrası her seçmenin yansız ve rastgele yapıldıđı belirtilmiřtir. Bu sebeple “siz yaratıcı çıktınız daha iyi madde geliřtirmenizi bekliyorum” gibi söylemlerle öđretmenlerde farklı motivasyon yaratabilecek veya beklenti içerisine sokulacak uygulamalardan sakınılarak bu tehvide önlem alınmıřtır.

Yukarıda sıralanan iç geçerliđe tehdit oluşturabilecek 9 faktörden 8'inin etkisi arındırıldıđı söylenebilir. Bu sebeple madde ve test psikometrik özelliklerindeki olası deđiřim, madde yazarlıđı eđitiminden kaynaklıdır ve madde yazarlıđı eđitimince açıklanmaktadır denebilir. Bu da arařtırmanın iç geçerliđine kanıt olarak sunulabilir.

3.6.2. Arařtırmanın Dıř Geçerliđi

Dıř geçerlik, sonuçların deneklerin seçildiđi büyük gruplara, evrene genellenebilirlik derecesidir (Büyüköztürk vd., 2016, s.174). Arařtırmanın bulgularının örneklem grubunun yansız seçilmesi sebebiyle arařtırma evreni olan Ankara ilinde devlet okullarındaki tüm 8.sınıf matematik ve Türkçe öđretmenlerine genellenebileceđi söylenebilir.

4. BULGULAR VE TARTIŞMA

Bu bölümde, alt problem sırasına göre verilmiş araştırma bulguları ve bu bulgularla ilgili değerlendirmeler yer almaktadır.

4.1. Yaratıcılık ile Madde ve Test Psikometrik Özellikleri İlişisine Yönelik Bulgular

4.1.1. Yaratıcılık İndeksleri ile Madde Psikometrik Özellikleri

Matematik

Matematik öğretmenlerinin yaratıcılık indeksleri ile geliştirdikleri maddelerin güçlük, ayırıcılık gücü ve çeldiricilik indeksleri arasında yapılan korelasyon analizi sonuçları Tablo 4.1’de sunulmuştur.

Tablo 4.1: Matematik Öğretmenlerine Ait Yaratıcılık İndeksleri ile Madde Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki Korelasyon Analizi Sonuçları

		p_i	r_{jx}	ζ
Yaratıcılık İndeksi	r	-0,222	-0,034	0,665
	p	0,169	0,834	0,000
	r^2	0,049	0,001	0,442

Tablo 4.1 incelendiğinde yaratıcılık indeksi ile madde güçlüğü arasında düşük düzeyde, negatif, anlamlı olmayan ($r=-0,222$; $p>0,01$) bir ilişki olduğu görülmektedir. Yaratıcılık indeksi ile madde ayırıcılık gücü indeksi arasında çok düşük, negatif, anlamlı olmayan ($r=-0,034$; $p>0,01$) bir ilişki bulunmaktadır. Yaratıcılık indeksi ile madde çeldiricilik indeksleri arasında orta düzeyde, pozitif ve anlamlı bir ilişki olduğu görülmektedir ($r=0,665$; $p<0,01$). Buna göre yaratıcılık arttıkça madde çeldiriciliğinin de arttığı söylenebilir. Determinasyon katsayısı ($r^2=0,44$) dikkate alındığında, madde çeldiriciliğindeki toplam varyansın %44’ünün yaratıcılıktan kaynaklandığı söylenebilir. Fakat bu neden-sonuç bağlamında bir kanıt olarak ileri sürülemeyeceğinden, yapılan yorum sadece değişimleşmenin ne kadar açıklandığının bir tespittir.

Yaratıcılığın madde psikometrik özellikleri ile olan ilişkisini daha iyi yorumlayabilmek için, aynı kazanımı ölçmeye yönelik madde geliştiren, farklı yaratıcılık düzeylerindeki matematik öğretmenlerinin geliştirdiği maddelerin

psikometrik özellikleri arasında anlamlı farklılıkların olup olmadığı test edilmiş ve sonuçları Tablo 4.2'de sunulmuştur.

Tablo 4.2: Farklı Yaratıcılık Düzeylerinde Aynı Kazanıma Yönelik Madde Geliştiren Matematik Öğretmenlerinin Geliştirdikleri Maddelerin Güçlük, Ayırcılık Gücü, Çeldiricilik İndeksleri Arasındaki z Testi Sonuçları

Madde No	Form	Güçlük			Ayırcılık			Çeldiricilik		
		p_i	z	p	r_{jx}	z	p	Ç	z	p
1	Form-I	0,73	0,45	0,65	0,25	-0,65	0,52	0,61*	-2,13	0,03
	Form-II	0,71			0,31			0,71		
2	Form-I	0,43	-1,41	0,16	0,30	-0,67	0,50	0,70	0,86	0,39
	Form-II	0,50			0,36			0,66		
3	Form-I	0,67	1,26	0,21	0,29	0,11	0,91	0,81	4,04	0,00
	Form-II	0,61			0,28			0,63		
4	Form-I	0,75	5,2	0,00	0,21	-2,22	0,03	0,71	0,22	0,83
	Form-II	0,50			0,41			0,70		
5	Form-I	0,67	6,45	0,00	0,30	1,07	0,28	0,72	-0,45	0,65
	Form-II	0,35			0,20			0,74		
6	Form-I	0,75	-0,71	0,48	0,29	1,78	0,08	0,65*	-1,30	0,19
	Form-II	0,78			0,12			0,71*		
7	Form-I	0,39	-8,19	0,00	0,29	0,11	0,91	0,85	5,25	0,00
	Form-II	0,79			0,28			0,62*		
8	Form-I	0,65	-3,38	0,00	0,22	0,00	1,00	0,73	0,00	1,00
	Form-II	0,80			0,22			0,73*		
9	Form-I	0,33	-9,34	0,00	0,29	1,27	0,20	0,77	6,8	0,00
	Form-II	0,79			0,17			0,44*		
10	Form-I	0,42	-4,03	0,00	0,26	-0,32	0,75	0,66	1,05	0,29
	Form-II	0,62			0,29			0,61		
11	Form-I	0,29	-5,11	0,00	0,24	-1,32	0,19	0,72	1,94	0,05
	Form-II	0,54			0,36			0,63		
12	Form-I	0,62	-5,25	0,00	0,27	0,53	0,60	0,67	-0,87	0,38
	Form-II	0,85			0,22			0,71*		
13	Form-I	0,70	2,11	0,03	0,23	-0,53	0,60	0,78	3,72	0,00
	Form-II	0,60			0,28			0,61		
14	Form-I	0,51	0,00	1,00	0,26	-0,32	0,75	0,66	0,42	0,67
	Form-II	0,51			0,29			0,64		
15	Form-I	0,44	-7,02	0,00	0,25	-1,68	0,09	0,75	1,78	0,08
	Form-II	0,78			0,40			0,67*		
16	Form-I	0,48	-3,25	0,00	0,36	0,11	0,91	0,75	1,56	0,12
	Form-II	0,64			0,35			0,68		
17	Form-I	0,75	3,02	0,00	0,28	-1,72	0,09	0,68*	0,85	0,40
	Form-II	0,61			0,43			0,64		
18	Form-I	0,52	-3,72	0,00	0,30	-0,79	0,43	0,67	-1,55	0,12
	Form-II	0,70			0,37			0,74		
19	Form-I	0,80	8,6	0,00	0,24	-0,43	0,67	0,79*	4,16	0,00
	Form-II	0,38			0,28			0,60		
20	Form-I	0,76	6,2	0,00	0,21	0,93	0,35	0,65*	-0,86	0,39
	Form-II	0,46			0,12			0,69		

*Ç<Çmin

Tablo 4.2 incelendiğinde, madde güçlük indekslerinin yaratıcılığı yüksek grup tarafından geliştirilen maddelerle oluşturulan form için 0,29 ile 0,80; yaratıcılığı düşük grup için 0,35 ile 0,85 arasında değiştiği görülmektedir. Yaratıcılığı yüksek gruptaki öğretmenler tarafından geliştirilen çok kolay ve çok zor madde bulunmazken; yaratıcılığı düşük grup için çok kolay 1 madde (8.madde) bulunmakta, çok zor madde ise bulunmamaktadır. İlk grupta 11 kolay, 6 orta güçlükte, 3 zor; ikinci grupta 11 kolay, 6 orta güçlükte, 2 zor madde vardır. Aynı kazanımdaki madde güçlük indeksleri arasındaki z testi sonuçlarının 8 tanesi pozitif yani yaratıcılığı yüksek gruptaki öğretmenlerin geliştirdiği maddeler lehine, 11 tanesinin ise negatif olarak yaratıcılığı düşük grup lehine olduğu görülmektedir. Anlamlı farklılıkların 5 tanesi yaratıcılığı yüksek grup lehine, 9 tanesi ise yaratıcılığı düşük grup lehinedir. Bu veriler yaratıcılığı düşük gruptaki öğretmenlerin geliştirdikleri kolay madde sayısının, yaratıcılığı yüksek gruptan fazla olduğuna işaret etmektedir. Korelasyon sonuçları ile birlikte değerlendirildiğinde hissedilir bir farklılaşmanın olduğu fakat anlamlı ve güçlü bir ilişkinin olmadığı anlaşılmaktadır. Matematik alanında daha yaratıcı madde yazarlarının geliştirdiği maddelerin daha kolay veya daha zor olacağını iddia edebilecek kanıt elde edilmemiştir.

Tablo 4.2 incelendiğinde madde ayıricılık gücü indekslerinin yaratıcılığı yüksek grup tarafından geliştirilen maddelerle oluşturulan form için 0,21 ile 0,36; yaratıcılığı düşük grup için 0,12 ile 0,43 arasında değiştiği görülmektedir. Yaratıcılığı yüksek grupta yer alan öğretmenler tarafından geliştirilen 4 maddenin madde ayıricılık gücü indeksi 0,30 ve üzerinde, yaratıcılığı düşük grupta ise bu düzeyde 8 madde bulunmaktadır. Madde ayıricılık gücü indeksleri arasındaki z testi sonuçlarının 8 tanesi pozitif yani yaratıcılığı yüksek gruptaki öğretmenler lehine, 11 tanesinin ise negatif olarak yaratıcılığı düşük grup lehine olduğu görülmektedir. Yaratıcılığı düşük ve yüksek grubun geliştirdikleri maddelerin ayıricılık gücü indeksleri arasında yapılan z testi sonuçları arasında anlamlı farklılık bulunmamıştır. Korelasyon sonuçları ile birlikte değerlendirildiğinde matematik alanında, yaratıcılığın madde geliştirmede madde ayıricılık gücü ile kuvvetli ve anlamlı bir ilişkisi olmadığı, daha yaratıcı madde yazarlarının, kazanıma sahip bireyler ile sahip olmayanları ayırt edebilecek nitelikte ölçme aracı geliştirebileceği iddiasına kanıt olamayacağı değerlendirilmiştir.

Yaratıcılığın tüm bilişsel süreç düzeylerinde değil de üst bilişsel süreçlerde etkili olabileceği düşünülerek matematik maddelerinden (2, 3, 4, 5, 6, 7, 8, 9, 12, 15, 16, 18, 19. maddeler) bu araştırmadaki belirlenen en üst bilişsel süreç olan uygulama düzeyinde ölçme yapmak üzere geliştirilen maddelerin güçlük ve ayırıcılık gücü indekslerinin, geliştiricilerinin yaratıcılık indeksleri ile ilişkisi sınanmıştır. Verilerin dağılımının normallikten sapma gösterip göstermediği yoklanmıştır. Yaratıcılık indeksleri (S-W=0,956; p=0,325>0,05) ve madde ayırıcılık gücü indekslerinin (S-W=0,971; p=0,651>0,05) dağılımlarının normallikten sapmadığı fakat madde güçlük indeksleri dağılımının normallik sayılıtısını karşılamadığı görülmüştür (S-W=0,905; p=0,02<0,05). Maddeleri geliştiren madde yazarlarının yaratıcılıkları ile bu maddelerin güçlük ve ayırıcılık gücü indeksleri arasında yapılan korelasyon analizi sonucu da yaratıcılık ile madde güçlüğü ($\rho=-0,342$; p=0,087>0,01) ve madde ayırıcılığı ($r=-0,81$; p=0,694>0,01) arasında anlamlı ilişki bulunmamıştır. Bu sonuçlar yaratıcılığın yalnız üst bilişsel süreçlerde ölçme yapmak üzere geliştirilen maddeler dikkate alındığında da, ilişkili bir değişken olmadığı; daha yaratıcı madde yazarlarının üst düzey bilişsel süreçlerde daha iyi madde geliştirebileceği iddiası (Haladyna, 2004, s.85) ile çeliştiği söylenebilir.

Tablo 4.2 incelendiğinde yaratıcılığı yüksek öğretmenlerce geliştirilen 5 maddenin kabul edilebilir çeldiricilik indeksinden daha düşük çeldiricilik indeksine sahip olduğu; yaratıcılığı düşük öğretmenlerde ise bu şekilde madde sayısının 6 olduğu görülmektedir. Madde çeldiricilik indeksleri arasındaki z testi sonuçlarının 13 tanesi pozitif yani yaratıcılığı yüksek gruptaki öğretmenler lehine, 6 tanesinin ise negatif olarak yaratıcılığı düşük grup lehine olduğu görülmektedir. Toplam 5 anlamlı farklılıkların tümünün yaratıcılığı yüksek grup lehine olduğu görülmektedir. Bu veriler yaratıcılığı yüksek grubun geliştirdikleri çeldiriciliği yüksek madde sayısının, yaratıcılığı düşük gruptan fazla olduğuna işaret etmektedir. Korelasyon sonuçları ile birlikte değerlendirildiğinde, matematik alanında, daha yaratıcı madde yazarlarının, daha fazla sayıda dengeli işleyen çeldiricilere sahip maddeler geliştirebildiği yorumu yapılabilir.

Türkçe

Türkçe öğretmenlerinin yaratıcılık indeksleri ile geliştirdikleri maddelerin güçlük, ayıricılık gücü ve çeldiricilik indeksleri arasında yapılan korelasyon analizi sonuçları Tablo 4.3'de sunulmuştur.

Tablo 4.3: Türkçe Öğretmenlerine Ait Yaratıcılık İndeksleri ile Madde Güçlük, Ayıricılık Gücü, Çeldiricilik İndeksleri Arasındaki Korelasyon Analizi Sonuçları

		p_i	r_{jx}	ζ
Yaratıcılık İndeksi	r	0,080	-0,040	0,668
	p	0,625	0,808	0,000
	r^2	0,006	0,002	0,446

Tablo 4.3 incelendiğinde yaratıcılık indeksi ile madde güçlüğü arasında çok düşük düzeyde, pozitif, anlamlı olmayan bir ilişki olduğu görülmektedir ($r=0,080$; $p>0,01$). Ayıricılık gücü indeksleri ile yaratıcılık indeksleri arasında hemen hemen hiç ilişki yoktur ($r=-0,040$; $p>0,01$). Yaratıcılık indeksi ile madde çeldiricilik indeksleri arasında orta düzeyde, pozitif ve anlamlı bir ilişki olduğu görülmektedir ($r=0,668$; $p<0,01$). Buna göre yaratıcılık arttıkça madde çeldiriciliğinin de arttığı söylenebilir. Determinasyon katsayısı ($r^2=0,45$) dikkate alındığında, madde çeldiriciliğindeki toplam varyansın %45'inin yaratıcılıkla açıklanabildiği söylenebilir.

Aynı kazanımı ölçmeye yönelik madde geliştiren, farklı yaratıcılık düzeylerindeki Türkçe öğretmenlerinin geliştirdiği maddelerin psikometrik özelliklerinin anlamlı farklılıklarının olup olmadığı test edilmiş ve sonuçları Tablo 4.4'de sunulmuştur.

Tablo 4.4: Farklı Yaratıcılık Düzeylerinde Aynı Kazanıma Yönelik Madde Geliştiren Türkçe Öğretmenlerinin Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü, Çeldiricilik İndeksleri Arasındaki z Testi Sonuçları

Madde No	Form	Güçlük			Ayırıcılık			Çeldiricilik		
		p_i	z	p	r_{jx}	z	p	ζ	z	p
1	Form-I	0,69	-0,67	0,50	0,37	0,91	0,36	0,82	4,52	0,00
	Form-II	0,72			0,29			0,62*		
2	Form-I	0,74	-0,23	0,82	0,23	0,00	1,00	0,67*	-4,82	0,00
	Form-II	0,75			0,23			0,87		
3	Form-I	0,64	-3,13	0,00	0,29	0,65	0,52	0,94	8,20	0,00
	Form-II	0,78			0,23			0,60*		
4	Form-I	0,63	-1,29	0,20	0,28	-0,67	0,50	0,65	-0,21	0,83
	Form-II	0,69			0,34			0,66		
5	Form-I	0,50	-2,25	0,02	0,27	0,22	0,83	0,68	0,00	1,00
	Form-II	0,61			0,25			0,68		
6	Form-I	0,86	8,39	0,00	0,20	-2,59	0,01	0,70*	1,08	0,28
	Form-II	0,47			0,43			0,65		
7	Form-I	0,91	5,89	0,00	0,15	-1,49	0,14	0,67	-1,33	0,18
	Form-II	0,67			0,29			0,73		
8	Form-I	0,61	-4,23	0,00	0,17	-1,50	0,13	0,68	0,22	0,83
	Form-II	0,80			0,31			0,67*		
9	Form-I	0,84	8,83	0,00	0,23	-1,32	0,19	0,71*	2,96	0,00
	Form-II	0,42			0,35			0,57		
10	Form-I	0,69	-4,98	0,00	0,22	-2,01	0,04	0,59*	-0,21	0,83
	Form-II	0,89			0,40			0,60*		
11	Form-I	0,81	5,07	0,00	0,09	-1,24	0,22	0,70*	0,22	0,83
	Form-II	0,58			0,21			0,69		
12	Form-I	0,88	5,10	0,00	0,28	-0,78	0,44	0,49*	-4,78	0,00
	Form-II	0,67			0,35			0,72		
13	Form-I	0,52	3,90	0,00	0,43	2,38	0,02	0,78	2,07	0,04
	Form-II	0,33			0,22			0,69		
14	Form-I	0,78	0,72	0,47	0,21	-0,75	0,45	0,64*	-2,66	0,01
	Form-II	0,75			0,28			0,76		
15	Form-I	0,63	4,27	0,00	0,31	0,22	0,83	0,73	0,89	0,37
	Form-II	0,42			0,29			0,69		
16	Form-I	0,82	4,32	0,00	0,20	-0,21	0,83	0,75*	2,42	0,02
	Form-II	0,63			0,22			0,64		
17	Form-I	0,76	7,21	0,00	0,28	0,22	0,83	0,76	0,07	0,48
	Form-II	0,41			0,26			0,73		
18	Form-I	0,58	-2,10	0,04	0,23	0,63	0,53	0,71	2,14	0,03
	Form-II	0,68			0,17			0,61*		
19	Form-I	0,43	-6,39	0,00	0,31	0,55	0,58	0,69	-0,89	0,37
	Form-II	0,74			0,26			0,73		
20	Form-I	0,67	-3,24	0,00	0,33	0,99	0,32	0,68	0,65	0,52
	Form-II	0,81			0,24			0,65*		

* $\zeta < \zeta_{min}$

Tablo 4.4 incelendiğinde, madde güçlük indekslerinin yaratıcılığı yüksek grup tarafından geliştirilen maddelerle oluşturulan form için 0,43 ile 0,91; yaratıcılığı düşük grup için 0,33 ile 0,89 arasında değiştiği görülmektedir. Yaratıcılığı yüksek

grup için çok zor ve zor madde bulunmazken; yaratıcılığı düşük grup için 1 adet zor madde bulunmaktadır. İlk grupta 4 orta güçlükte, 10 kolay ve 6 çok kolay madde bulunurken; ikinci grupta 5 orta güçlükte, 9 kolay, 3 çok kolay madde vardır. Aynı kazanımdaki madde güçlük indeksleri arasındaki z testi sonuçlarının 10 tanesi pozitif yani yaratıcılığı yüksek gruptaki öğretmenlerin geliştirdiği maddeler lehine, 10 tanesinin ise negatif yani yaratıcılığı düşük grup lehine olduğu görülmektedir. Anlamlı farklılıkların 9 tanesi yaratıcılığı yüksek grup lehine, 5 tanesi ise yaratıcılığı düşük grup lehinedir. Bu veriler yaratıcılığı yüksek grubun geliştirdikleri kolay güçlükteki madde sayısının, yaratıcılığı düşük gruptan fazla olduğuna işaret etmektedir. Korelasyon sonuçları ile birlikte değerlendirildiğinde hissedilir bir ayrışmanın olduğu fakat anlamlı ve güçlü bir ilişkinin olmadığı anlaşılmaktadır. Türkçe alanında, yaratıcılığın madde geliştirmede madde güçlüğüyle kuvvetli ve anlamlı bir ilişkisi olmadığı söylenebilir. Daha yaratıcı madde yazarlarının daha kolay veya zor madde geliştirdiğini iddia edebilecek kanıtlara ulaşılmamıştır.

Tablo 4.4 incelendiğinde madde ayırıcılık indekslerinin yaratıcılığı yüksek grup tarafından geliştirilen maddelerle oluşturulan form için 0,09 ile 0,43; yaratıcılığı düşük grup için 0,17 ile 0,43 arasında değiştiği görülmektedir. Yaratıcılığı yüksek grup öğretmenlerce geliştirilen 5 maddenin ayırıcılık gücü indeksi 0,30 ve üzerinde, yaratıcılığı düşük grupta ise 6 madde bu düzeyde bulunmaktadır. Madde ayırıcılık gücü indeksi 0,20 ile 0,29 arasında olan yaratıcılığı yüksek grup için 12 madde, yaratıcılığı düşük grup için 13 madde bulunmaktadır. Madde ayırıcılık gücü indeksleri arasındaki z testi sonuçlarının 9 tanesi pozitif yani yaratıcılığı yüksek gruptaki öğretmenler lehine, 10 tanesinin ise negatif olarak yaratıcılığı düşük grup lehine olduğu görülmektedir. Yaratıcılığı düşük ve yüksek grubun geliştirdikleri maddelerin ayırıcılık gücü indeksleri arasında yapılan z testi sonuçları arasında anlamlı farklılık bulunan 1 madde (6.madde) yaratıcılığı düşük grup lehinedir. Korelasyon sonuçları ile birlikte değerlendirildiğinde Türkçe alanında, yaratıcılığın madde geliştirmede madde ayırıcılık gücü ile kuvvetli ve anlamlı bir ilişkisi olmadığı, daha yaratıcı madde yazarlarının, kazanıma sahip bireyler ile sahip olmayanları ayırt edebilecek nitelikte ölçme aracı geliştirebileceği iddiasına kanıt olamayacağı söylenebilir.

Türkçe alanında yalnızca uygulama düzeyinde ölçme yapmak üzere geliştirilen maddelerden (1, 3, 11, 13, 20. maddeler) madde güçlük ve ayırıcılık gücü indekslerinin, geliştiricilerinin yaratıcılık indeksleri ile ilişkisi yoklanmıştır. Analize alınacak verilerin dağılımının normallikten sapma gösterip göstermediği incelendiğinde; yaratıcılık indeksleri (S-W=0,930; $p=0,447>0,05$), madde güçlük indeksleri (S-W=0,904; $p=0,244<0,05$) ve madde ayırıcılık gücü indekslerinin (S-W=0,972; $p=0,912>0,05$) dağılımlarının normallikten sapmadığı görülmüştür. Maddeleri geliştiren madde yazarlarının yaratıcılıkları ile bu maddelerin güçlük ve ayırıcılık gücü indeksleri arasında yapılan korelasyon analizi sonucu da yaratıcılık ile madde güçlüğü ($r=-0,120$; $p=0,742>0,01$) ve madde ayırıcılığı ($r=0,416$; $p=0,232>0,01$) arasında anlamlı ilişki bulunmamıştır. Bu noktada dikkat çeken durum madde ayırıcılık gücü indekslerinin tüm maddeler bazında yaratıcılıkla olan korelasyon değerinin ($r=-0,040$) neredeyse sıfır olduğu düzeyden, yalnızca uygulama basamağındaki maddeler dikkate alındığında orta düzeye ($r=0,416$) çıkmasıdır. Fakat bu ilişki anlamlı değildir.

Tablo 4.4 incelendiğinde yaratıcılığı yüksek öğretmenler tarafından geliştirilen 8 maddenin kabul edilebilir çeldiricilik indeksinden daha düşük çeldiricilik indeksine sahip olduğu; yaratıcılığı düşük öğretmenler tarafından geliştirilen maddeler için bu durumda olan madde sayısının 6 olduğu görülmektedir. Madde çeldiricilik indeksleri arasındaki z testi sonuçlarının 12 tanesi pozitif yani yaratıcılığı yüksek gruptaki öğretmenler lehine, 7 tanesinin ise negatif olarak yaratıcılığı düşük grup lehine olduğu görülmektedir. Toplam 6 anlamlı farklılıkların 3'ü yaratıcılığı yüksek diğer 3'ü ise yaratıcılığı düşük grup lehine olduğu görülmektedir. Korelasyon sonuçları ile birlikte değerlendirildiğinde, Türkçe alanında, daha yaratıcı madde yazarlarının, daha fazla sayıda işleyen çeldiricili maddeler geliştirebildiği yorumu yapılabilir.

Bu bulgular ışığında, madde yazarı matematik ve Türkçe öğretmenlerinin yaratıcılık indeksleri ile geliştirdikleri maddelerin güçlüğü ve ayırıcılığı arasında anlamlı ilişki bulunmazken, çeldiriciliği ile orta düzeyde ve anlamlı bir ilişki bulunmuştur.

Matematik ve Türkçe alanlarında, öğretmenlerin yaratıcılıklarının bir göstergesi olan yaratıcılık indeksleri ile geliştirdikleri maddelerin güçlük ve ayırıcılık gücü indeksleri arasında yapılan korelasyon analizleri sonucu elde edilen bulgular

anlamli bir iliskinin olmadigini gostermektedir. Bulgular degerlendirildiginde, madde yazarinin yaratıcı olup olmamasinin, geliştirdigi maddenin kolay ya da zor olmasıyla iliskili olmadigi soylenebilir. Ayrıca yoklanan özelliğe sahip olan ve olmayan öğrencileri ayırabilecek güçte madde geliştirme ile yaratıcılığın iliskili olmadigi da öne sürülebilir. ÇSM çeldiricilerinin dengeli işlerliğinin bir ifadesi olan çeldiricilik indeksi ile yaratıcılık indeksleri arasında ise iki alanda orta düzeyde anlamlı iliski olduğuna yönelik bulgular elde edilmiştir. Bu bulgulara dayanarak, ÇSM'den beklenen bir özellik olan çeldiricilerinin eşit düzeyde güçlü olmasının yaratıcılıkla iliskili olduğu bir başka deyişle yaratıcılığı yüksek madde yazarı öğretmenlerin işler çeldirici bulma konusunda yaratıcılığı düşük öğretmenlere nazaran daha başarılı oldukları ve geliştirdigi maddelerin çeldiricilerinin daha dengeli işlediği soylenebilir.

Bu sonuçlar, Ebel'in (1951) öne sürdüğü ve diğer uzmanların desteklediği (Haladyna, 2004; Haladyna vd., 2002; Hambleton, 1979; Wesman, 1971) madde yazarinin iyi maddeler geliştirebilmesi için yaratıcı olması gerektiği görüşü ile çelişmektedir.

4.1.2. Yaratıcılık Düzeyleri ile Test Psikometrik Özellikleri

Matematik

Matematik öğretmenlerinin geliştirdigi maddelerle oluşturulan Form I ve Form II testlerinin öğrencilere uygulanması sonucu elde edilen puanlara ait test psikometrik özelliklerinin farklılık gösterip göstermediği incelenmiş ve Tablo 4.5'de gösterilmiştir.

Tablo 4.5: Farklı Yaratıcılık Düzeyindeki Matematik Öğretmenlerinin Geliştirdigi Maddelerle Oluşturulan Testlerin Psikometrik Özelliklerinin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları

<i>Formlar</i>	<i>KR-20</i>	<i>Güvenirlik z</i>	<i>p</i>	<i>Varyans Sx²</i>	<i>F</i>	<i>X̄</i>	<i>Aritmetik Ortalama U</i>	<i>p</i>	<i>Cohen's d</i>
Form I	0,698	-0,53	0,60	12,94	0,94	11,67	17483,5	0,008	0,26
Form II	0,724			13,66		12,52			

Tablo 4.5 incelendiğinde gerek yaratıcılığı yüksek gerekse yaratıcılığı düşük öğretmenlerce geliştirilen maddelerle oluşturulan testlerin iç tutarlık güvenilirliklerinin, sınıf içi testler için belirlenen ölçüt güvenilirlik değeri 0,75'in altında

olduğu görülmektedir (Davidshofer & Murphy, 2005, s.150). Testlerin KR-20 iç tutarlık güvenilirlikleri arasında uygulanan z testi sonucu ($z=-0,53$; $p>0,01$) anlamlı fark bulunmamıştır. Matematik alanında yaratıcılığı yüksek ve düşük grupların geliştirdiği maddelerle oluşturulan testlerin iç tutarlık güvenilirlikleri bakımından farklılaşmadığı, ölçme sonuçlarına karışan hata miktarları ve maddelerin birlikte matematik başarısını ölçme konusundaki tutarlıkları bakımından benzer düzeyde olduğu söylenebilir.

Testlerin varyansları arasında yapılan F testi sonucu ($S_1^2/S_2^2=12,94/13,66=0,95$; $F_{0,005}=0.947<F_U=1.439$) anlamlı fark bulunmamıştır. Matematik alanında yaratıcılığı yüksek ve düşük grupların geliştirdiği maddelerle oluşturulan testlerin, matematik başarısı yoklanan benzer gruplarda meydana getirdiği ayrışma, başka bir deyişle homojenlik/heterojenlik bakımından farklılaşmadığı, benzer düzeyde olduğu söylenebilir.

Testlerin aritmetik ortalamalar arasındaki yapılan Mann Whitney U testi sonucu yaratıcılığı düşük grubun geliştirdiği maddelerle oluşturulan test lehine anlamlı derecede farklılık görülmüştür ($U=17484$; $p<0,01$; $SO_{FormI}=188,13$; $SO_{FormIV}=218,87$). Yaratıcılığı yüksek ve düşük grupların geliştirdiği maddelerle oluşturulan testlerin, matematik başarısı yoklanan benzer grupların başarısını ölçmede yaratıcılığı düşük grup lehine anlamlı farklılaşması bu gruba ait testin öğrencilere daha kolay geldiğine bir işarettir. Bu karşılaştırmaya ait etki büyüklüğü 0,26 değeri ile küçük düzeyde bulunduğundan elde edilen farklılığın örneklem büyüklüğünden kaynaklandığı söylenebilir.

Yaratıcılığı yüksek ve düşük matematik öğretmenlerinin geliştirdikleri maddelerle oluşturulan testlerin kapsam geçerliklerini incelemek amacıyla 14 alan uzmanının görüşlerinden elde edilen veriler kullanılmıştır. Seçme frekansları, hesaplanan KGO değerleri ve aynı kazanımı yoklayan maddeler için yapılan iki oranı test etmede kullanılan z testi sonuçları Tablo 4.6'da sunulmuştur.

Tablo 4.6: Farklı Yaratıcılık Düzeyindeki Matematik Öğretmenlerinin Geliştirdiği Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları

<i>Madde No</i>	<i>Form</i>	<i>Gerekli</i>	<i>Yararlı/ Yetersiz</i>	<i>Gereksiz</i>	<i>KGO</i>	<i>z</i>	<i>p</i>
1	Form-I	13	1	0	0,857*	-1,47	0,14
	Form-II	14	0	0	1,000*		
2	Form-I	4	10	0	-0,429	-8,37	0,00
	Form-II	14	0	0	1,000*		
3	Form-I	14	0	0	1,000*	3,94	0,00
	Form-II	9	4	1	0,286		
4	Form-I	13	1	0	0,857*	-1.46	0,14
	Form-II	14	0	0	1,000*		
5	Form-I	14	0	0	1,000*	5.29	0,00
	Form-II	7	4	3	0,000		
6	Form-I	14	0	0	1,000*	2,16	0,03
	Form-II	12	2	0	0,714*		
7	Form-I	12	2	0	0,714*	3.94	0.00
	Form-II	7	6	1	0,000		
8	Form-I	6	7	1	-0,143	-6,11	0,00
	Form-II	14	0	0	1,000*		
9	Form-I	11	3	0	0,571*	-2,77	0,00
	Form-II	14	0	0	1,000*		
10	Form-I	14	0	0	1,000*	0,00	1,00
	Form-II	14	0	0	1,000*		
11	Form-I	14	0	0	1,000*	2,77	0,01
	Form-II	11	3	0	0,571*		
12	Form-I	13	1	0	0,857*	4,58	0,00
	Form-II	14	0	0	1,000*		
13	Form-I	14	0	0	1,000*	0,00	1,00
	Form-II	14	0	0	1,000*		
14	Form-I	12	2	0	0,714*	0,00	1,00
	Form-II	12	2	0	0,714*		
15	Form-I	10	4	0	0,429	-3,35	0,00
	Form-II	14	0	0	1,000*		
16	Form-I	9	5	0	0,286	-1,53	0,13
	Form-II	11	3	0	0,571*		
17	Form-I	14	0	0	1,000*	0,00	1,00
	Form-II	14	0	0	1,000*		
18	Form-I	14	0	0	1,000*	0,00	1,00
	Form-II	14	0	0	1,000*		
19	Form-I	14	0	0	1,000*	7,10	0,00
	Form-II	5	6	3	-0,286		
20	Form-I	13	1	0	0,857*	-1,47	0,14
	Form-II	14	0	0	1,000*		

* $p < 0,05$

Tablo 4.6 incelendiğinde 14 uzmanın görüşleri doğrultusunda KGO'na $\alpha=0,05$ anlamlılık düzeyinde katkı yapan maddeler Form-I için 2, 8, 15 ve 16. maddeler haricindeki; Form-II için 3, 5, 7 ve 19. maddeler haricindeki diğer 16'şar maddedir. Aynı kazanımı yoklayan maddelerin KGO değerlerinin z testi sonuçlarına göre 7 tanesi pozitif yani yaratıcılığı yüksek grup lehine; 8 tanesi ise negatif olarak yaratıcılığı düşük grup lehinedir. Anlamlı farklılıkların 7 tanesi yaratıcılığı yüksek grup lehine; 4 tanesi ise yaratıcılığı düşük grup lehinedir.

Test bazında KGİ, kritik değer üzerindeki 16 KGO ortalamalarından hesaplanmış ve Form-I için 0,90; Form-II için ise 0,91 değeri bulunmuştur. Elde edilen KGİ değerlerine ilişkisiz iki oranı test etmede kullanılan z testi uygulanmış; anlamlı fark bulunmamıştır ($z=-0,01$; $p=0,92>0,01$). Gerek aynı kazanımı yoklayan maddelerin KGO değerleri, gerekse tüm teste ait KGİ değerlerinin anlamlı derecede farklılaştığına dair kanıt elde edilmemiştir. Yaratıcılığın kapsamı daha iyi ölçülebilen madde geliştirebilme ile ilgili bir özellik olmadığı, daha yaratıcı madde yazarının kapsamdan daha iyi davranış örneklemini seçebildiği veya tersi gibi bir iddiaya kanıt olamayacağı söylenebilir.

Geliştirilen testleri yanıtlayarak kendilerinden madde ve test puanı alınan 8. sınıf öğrencilerinin 23 Kasım 2016 tarihinde girdikleri ve 03 Ocak 2017 tarihinde sonuçları açıklanan MOAS matematik alt testi puanları ile eş formlardan aldıkları puanların korelasyon analizi yapılarak, eş formların uyum geçerliğine kanıt aranmıştır. Yapılan SBSFK ve iki korelasyonun farklılaşmasına dönük z testi sonuçları Tablo 4.7'de verilmiştir.

Tablo 4.7: Öğrencilerin Merkezî Sınav Matematik Alt Testi Puanları ile Eş Formlardan Aldıkları Matematik Puanlarının Korelasyon Analizi Sonuçları

	<i>MOAS Puanları</i>	<i>z</i>	<i>p</i>
Form I	0,779**	-0,128	0,90
Form II	0,784**		

** $p<0,01$

Tablo 4.7 incelendiğinde uyum geçerliğine kanıt olarak sunulan korelasyon değerlerinin 0,80 değerinin altında kaldığı ve yetersiz olduğu görülmektedir (Ebel & Frisbie, 1991, s.106). Gerek Form-I gerekse Form-II puanlarının MOAS puanları ile yüksek düzeyde, pozitif ve anlamlı bir ilişkisinin olduğu görülmektedir ($\rho_{\text{Form I-}}$

MOAS I=0,779; p=0,00<0,01; $\rho_{\text{Form II-MOAS II}}=0,779$; p=0,00<0,01). İki korelasyon katsayısı arasındaki z testi sonucu (z=-0,128; p>0,01) anlamlı fark bulunmamıştır. Yaratıcılığı yüksek ve düşük madde yazarı matematik öğretmenlerinin geliştirdiği maddelerle oluşturulan testlerin uyum geçerlikleri arasında anlamlı bir fark bulunmamıştır. Yaratıcılığın ölçme aracının, aynı kapsamda ve aynı öğrenci grubuna yönelik önceden geliştirilmiş güvenilir ve geçerli bir ölçme aracıyla uyum gösterme gücüne etkisinin olmadığı söylenebilir.

Türkçe

Türkçe öğretmenlerinin geliştirdiği maddelerle oluşturulan Form I ve Form II testlerinin öğrencilere uygulanması sonucu elde edilen puanlara ait test psikometrik özelliklerinin farklılık gösterip göstermediği incelenmiş ve Tablo 4.8'de gösterilmiştir.

Tablo 4.8: Farklı Yaratıcılık Düzeyindeki Türkçe Öğretmenlerinin Geliştirdiği Testlerin Psikometrik Özelliklerinin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları

Formlar	KR-20	Güvenirlilik z	p	Sx ²	Varyans F	\bar{X}	Aritmetik Ortalama t	Cohen's d
Form I	0,685	-0,51	0,61	11,048	0,868	13,990	3,392**	0,33
Form II	0,711			12,730		12,835		

**p<0,01

Tablo 4.8 incelendiğinde iki testin de güvenirliliklerinin 0,75 değeri altında kalarak, uygun düzeyde olmadığı söylenebilir (Davidshofer & Murphy, 2005, s.150). Testlerin KR-20 iç tutarlık güvenirlilikleri arasında uygulanan ilişkisiz örneklem korelasyonlarının z testi sonucu (z=-0,511; p>0,01) anlamlı fark bulunmamıştır. Türkçe alanında yaratıcılığı yüksek ve düşük grupların geliştirdiği maddelerle oluşturulan testlerin iç tutarlık güvenirlilikleri bakımından farklılaşmadığı, ölçme sonuçlarına karışan hata miktarları ve maddelerin birlikte Türkçe başarısını ölçme konusundaki tutarlılıkları bakımından benzer düzeyde olduğu söylenebilir.

Yaratıcılığı yüksek ve düşük öğretmenler tarafından geliştirilen maddelerle oluşturulan testlerin varyansları arasında yapılan F testi sonucu ($S_1^2/S_2^2=11,05/12,73=0,868$; $F_{0,01;205;205}=0,868 < F_U=1,17$) anlamlı fark bulunmamıştır. Türkçe alanında yaratıcılığı yüksek ve düşük grupların geliştirdiği maddelerle oluşturulan testlerin, Türkçe başarısı yoklanan benzer gruplarda

meydana getirdiđi ayrışma, başka bir deyişle homojenlik/heterojenlik bakımından farklılaşmadığı, benzer düzeyde olduğu söylenebilir.

Normal dağılımlı Türkçe test puanlarının aritmetik ortalamaları arasındaki ilişkisiz iki örneklem t testi sonucu yaratıcılığı yüksek grup lehine anlamlı fark bulunmuştur ($t=3,392$; $p=0,001<0,01$). Bu karşılaştırmaya ait etki büyüklüğü değeri 0,33 olarak elde edilmiştir. Buna göre düşük derecede etki büyüklüğüne sahip olduğu ve bu farkın örneklem büyüklüklerinden kaynaklandığı söylenebilir. Yaratıcılığı yüksek grubun geliştirdiđi maddelerle oluşturulan testin, yaratıcılığı düşük gruba göre Türkçe başarısı yoklanan öğrencilere daha kolay geldiđi söylenebilir.

Yaratıcılığı yüksek ve düşük öğretmenlerin geliştirdikleri maddelerle oluşturulan testlerin kapsam geçerliklerini incelemek amacıyla 12 alan uzmanının görüşlerinden elde edilen seçme frekansları ve hesaplanan KGO değerleri Tablo 4.9'da sunulmuştur.

Tablo 4.9: Farklı Yaratıcılık Düzeyindeki Türkçe Öğretmenlerinin Geliştirdiği Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları

<i>Madde No</i>	<i>Form</i>	<i>Gerekli</i>	<i>Yararlı/ Yetersiz</i>	<i>Gereksiz</i>	<i>KGO</i>	<i>z</i>	<i>p</i>
1	Form-I	12	0	0	1,000*	0,00	1,00
	Form-II	12	0	0	1,000*		
2	Form-I	10	2	0	0,667*	0,00	1,00
	Form-II	10	2	0	0,667*		
3	Form-I	12	0	0	1,000*	0,00	1,00
	Form-II	12	0	0	1,000*		
4	Form-I	11	1	0	0,833*	-1,48	0,14
	Form-II	12	0	0	1,000*		
5	Form-I	10	2	0	0,667*	0,83	0,41
	Form-II	9	3	0	0,500		
6	Form-I	8	2	2	0,333	-1,64	0,10
	Form-II	10	2	0	0,667*		
7	Form-I	10	2	0	0,667*	0,00	1,00
	Form-II	10	2	0	0,667*		
8	Form-I	12	0	0	1,000*	1,48	0,14
	Form-II	11	1	0	0,833*		
9	Form-I	12	0	0	1,000*	0,00	1,00
	Form-II	12	0	0	1,000*		
10	Form-I	11	1	0	0,833*	0,00	1,00
	Form-II	11	1	0	0,833*		
11	Form-I	11	1	0	0,833*	0,00	1,00
	Form-II	11	1	0	0,833*		
12	Form-I	12	0	0	1,000*	0,00	1,00
	Form-II	12	0	0	1,000*		
13	Form-I	8	4	0	0,333	-0,83	0,41
	Form-II	9	3	0	0,500		
14	Form-I	12	0	0	1,000*	0,00	1,00
	Form-II	12	0	0	1,000*		
15	Form-I	11	1	0	0,833*	0,00	1,00
	Form-II	11	1	0	0,833*		
16	Form-I	12	0	0	1,000*	0,00	1,00
	Form-II	12	0	0	1,000*		
17	Form-I	10	2	0	0,667*	0,83	0,41
	Form-II	12	0	0	1,000*		
18	Form-I	11	1	0	0,833*	-1,48	0,14
	Form-II	12	0	0	1,000*		
19	Form-I	11	1	0	0,833*	0,00	1,00
	Form-II	11	1	0	0,833*		
20	Form-I	11	1	0	0,833*	0,00	1,00
	Form-II	11	1	0	0,833*		

* $p < 0,05$

Tablo 4.9 incelendiğinde 12 uzmanın görüşleri doğrultusunda KGO'na $\alpha=0,05$ anlamlılık düzeyinde katkı yapan maddeler Form-I için 6. ve 13. maddeler haricindeki; Form-II için 5. ve 13. maddeler haricindeki diğer 18 maddedir. Aynı kazanımı yoklayan maddelerin KGO değerlerinin z testi sonuçlarına göre 3 tanesi pozitif yani yaratıcılığı yüksek grup lehine; 4 tanesi ise negatif olarak yaratıcılığı düşük grup lehinedir. İki gruba ait aynı kazanımı yoklayan maddelerin KGO değerlerinin z testlerinin hiçbiri anlamlı çıkmamıştır.

Formlara ait KGİ, kritik değer üzerindeki 18 KGO ortalamalarından hesaplanmış ve Form-I için 0,86; Form-II için ise 0,89 değerleri bulunmuştur. Elde edilen KGİ değerlerine ilişkisiz iki oranı test etmeye yarayan z testi uygulanmış; 0,01 anlamlılık düzeyinde iki test arasında anlamlı fark bulunmamıştır ($z=-0,29$; $p=0,77>0,01$). Gerek aynı kazanımı yoklayan maddelerin KGO değerleri, gerekse tüm teste ait KGİ değerlerinin anlamlı derecede farklılaştığına dair kanıt elde edilmemiştir. Kapsam geçerliğine yaratıcılığın etkisi bakımından matematik alanı ile paralel sonuç elde edilmiştir.

Öğrencilerin MOAS Türkçe alt testi puanları ile eş formlardan aldıkları puanların korelasyon analizi yapılarak, eş formların uyum geçerliğine kanıt aranmıştır. Yapılan PMÇK analizi sonuçları ve iki korelasyonun farkına ait z testi sonuçları Tablo 4.10'da verilmiştir.

Tablo 4.10: Öğrencilerin Merkezî Sınav Matematik Alt Testi Puanları ile Eş Formlardan Aldıkları Türkçe Puanlarının Korelasyon Analizi Sonuçları

	<i>MOAS Puanları</i>	<i>z</i>	<i>p</i>
Form I	0,786**	1,223	0,22
Form II	0,735**		

** $p<0,01$

Tablo 4.10 incelendiğinde uyum geçerliğine kanıt olarak sunulan korelasyon değerlerinin 0,80 değerinin altında kaldığı ve yetersiz olduğu görülmektedir (Ebel & Frisbie, 1991, s.106). Gerek Form-I gerekse Form-II puanlarının MOAS ile yüksek düzeyde, pozitif ve anlamlı ($p_{\text{Form-I,MOAS}}=0,00<0,01$; $p_{\text{Form-II,MOAS}}=0,00<0,01$) ilişkisinin olduğu görülmektedir. İki korelasyon katsayısı arasındaki z testi sonucu ($z=1,22$; $p>0,01$) anlamlı fark bulunmamıştır. Yaratıcılığı yüksek ve düşük madde yazarı Türkçe öğretmenlerinin geliştirdiği maddelerle oluşturulan testlerin uyum geçerlikleri arasında anlamlı bir fark bulunmamıştır.

Bu bulgular ışığında matematik ve Türkçe öğretmenlerinin geliştirdikleri maddelerle oluşturulan testlerin iç tutarlık güvenilirlikleri, varyansları, kapsam, uyum geçerlikleri yaratıcılık düzeylerine göre anlamlı bir şekilde farklılaşmamış; aritmetik ortalamaları matematik alanında yaratıcılığı düşük grup lehine; Türkçe alanında ise yaratıcılığı yüksek grup lehine anlamlı bir şekilde farklılaşmıştır.

Testlerin psikometrik özellikleri birlikte değerlendirildiğinde yaratıcılığın aritmetik ortalamalar haricinde herhangi bir etkisinin olduğunu ileri sürebilecek kanıt bulunmamıştır. Aritmetik ortalamalardaki etkisi ise iki alanda farklı yönde ortaya çıkmıştır. Etki büyüklüklerinin küçük olması, bu farklılıkların örneklem büyüklüğünden kaynaklanmış olabileceğine işarettir. Ayrıca bu farkın oluşmasında, yaratıcılık gruplarındaki öğretmenlerin farklı özellikleri ya da geliştirilen maddelerin öğrencilerin anlayabilecekleri ve cevaplayabilecekleri nitelikte olmaması da etkili olmuş olabilir.

TDK (2017) yaratmak kelimesini “zekâ, düşünce ve hayal gücünden yararlanarak o zamana kadar görülmemiş yeni bir şey ortaya koymak, yapmak” olarak tanımlamaktadır. Yaratıcılığın, ihtiyaç duyulan bir alanda gereksinimleri karşılayabilecek, önceki çözüm ve yöntemlerden farklı ve yeni ürünleri ortaya koyabilmesi beklenir. Madde geliştirme açısından yaratıcılığın özellikle uygulama, analiz gibi üst bilişsel basamaklarda ölçme yapabilecek ölçme araçlarının geliştirilmesinde etkili olması beklenirken, çeldirici üretme haricinde, araştırma bulguları bu yönde bir yorumu destekleyecek nitelikte değildir. Araştırmada en üst bilişsel süreç olarak uygulama basamağında madde geliştirilmiştir. Analiz, değerlendirme ve yaratma basamaklarında daha karmaşık zihinsel süreçler söz konusu olduğundan yaratıcılığın da bu basamaklarda etkili olabileceği düşünülebilir.

Geniş ölçekli testler için bir maddenin ortalama geliştirilme süresinin iki yıl olduğu düşünüldüğünde (Koon & Kamata, 2013) kısa bir sürede madde geliştirilmesinin istenmesi madde yazarının yaratıcılığını ortaya koyabileceği imkânı bulamamasına neden olmuş olabilir. Ayrıca yaratıcılığın ölçülmesinde kullanılan TYDT'nin öğretmenlerin yaratıcılıkları arasındaki farklılığı ortaya koymada yetersiz kalmış olabileceği de öne sürülebilir.

4.1.3. Yaratıcılık Düzeyleri ile Yapı Geçerlikleri

Matematik

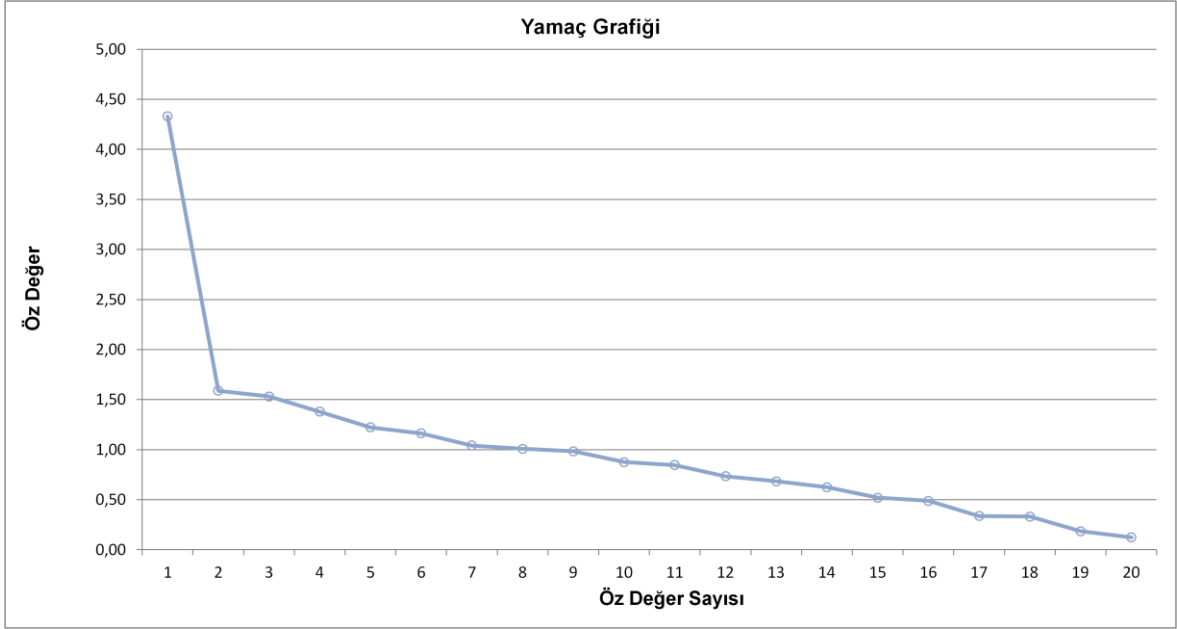
Yaratıcılığı yüksek madde yazarı öğretmenlerin geliştirdikleri maddelerle oluşturulan testin maddelerine faktör analizi yapıp yapılamayacağını kontrol etmek amacıyla yapılan analiz sonucu KMO değeri 0,69 ile orta düzeyde ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=352,1$; $sd=190$; $p=0,00$). Bu değerler, verilerin faktör analizine uygun olduğunu göstermektedir. Tetrakorik korelasyon matrisine dayalı uygulanan Paralel Analizler Yöntemi (PAY) sonucu testin tek faktörlü olduğu görülmüştür. Temel Bileşenler Analizi (TBA) sonucunda, döndürme uygulanmadan elde edilen faktör sayısı, faktör öz değerleri, açıklanan varyans oranları ve maddelerin tek (ilk) faktöre verdiği yükleri gösteren değerler Tablo 4.11'de verilmiştir.

Tablo 4.11: Matematik Form-I'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplamlı Varyans (%)</i>	<i>Madde No</i>	<i>Faktör Yük Değerleri</i>
1	4,33	21,65	21,65	1	0,446
2	1,59	7,95	29,59	2	0,501
3	1,53	7,66	37,25	3	0,492
4	1,38	6,90	44,15	4	0,374
5	1,22	6,11	50,26	5	0,508
6	1,16	5,82	56,08	6	0,523
7	1,04	5,21	61,29	7	0,497
8	1,01	5,05	66,33	8	0,382
				9	0,493
				10	0,430
				11	0,431
				12	0,465
				13	0,415
				14	0,440
				15	0,427
				16	0,588
				17	0,501
				18	0,490
				19	0,445
				20	0,403

Toplam açıklanan varyansa bakıldığında özdeğeri 1'den büyük olan 8 faktör görülmektedir. İlk faktör (özdeğeri 4,33) varyansın %21,65'ni açıklarken ikinci faktör (özdeğeri 1,59) ise varyansın %7,95'ni açıklamaktadır. Faktör sayısına karar

vermek için özdeğeri 1'den büyük olan faktörler dikkate alındığında test 8 faktörlü görülmektedir. Şekil 4.1'de verilen Form-I'e ait yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir.



Şekil 4.1: Matematik Form-I'e Ait Yamaç Grafiği

PAY uygulanarak elde edilen faktör sayısı, TBA sonucu elde edilen döndürülmemiş sonuçlar için özdeğerler ve faktör yükleri bir arada dikkate alınırsa birinci ve ikinci özdeğer arasındaki farka dayanarak ve yamaç grafiğinde birinci faktörden sonra ani düşüşün olması ve diğer faktörlerden sonra grafiğin nispeten yatay seyretmesinden dolayı testin tek boyutlu olduğu söylenebilir. Maddelerin birinci faktör üzerindeki yükleri 0,374 ile 0,588 arasındadır. Maddelerin tek faktöre verdikleri yük değerlerine göre, yaratıcılığı yüksek madde yazarlarının ölçülmek istenen yapıyı birlikte uygun şekilde ölçebilecek maddeler geliştirdiği söylenebilir.

Yaratıcılığı düşük madde yazarı öğretmenlerin geliştirdikleri maddelerle oluşturulan test puanlarına yönelik analiz sonucu KMO değeri 0,74 ile iyi düzeyde ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=429,1$; $sd=190$; $p=0,00$). Bu değerler, verilerin faktör analizine uygun olduğunu göstermektedir. Uygulanan PAY sonucu tek faktör bulunmuştur. TBA sonucunda,

döndürme uygulanmadan elde edilen faktör sayısı, faktör öz değerleri, açıklanan varyans oranları ve maddelerin tek (ilk) faktöre verdiği yükleri gösteren değerler Tablo 4.12’de verilmiştir.

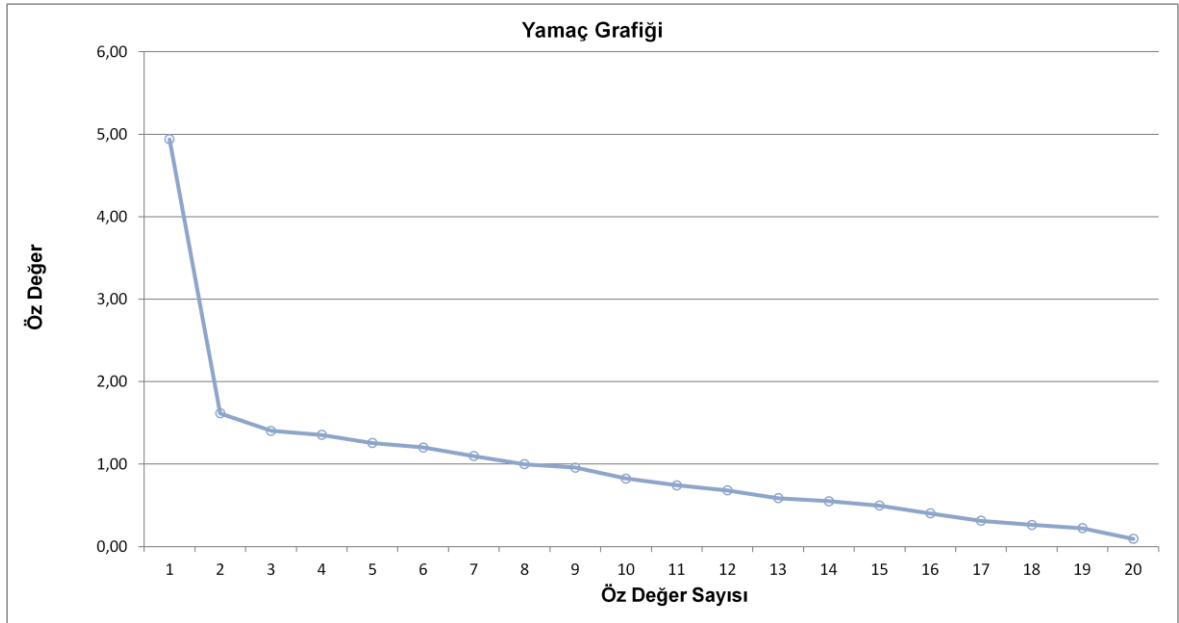
Tablo 4.12: Matematik Form-II’e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplamlı Varyans (%)</i>	<i>Madde No</i>	<i>Faktör Yük Değerleri</i>
1	4,94	24,69	24,69	1	0,521
2	1,62	8,07	32,77	2	0,562
3	1,40	7,02	39,79	3	0,432
4	1,35	6,77	46,57	4	0,650
5	1,26	6,28	52,85	5	0,344
6	1,20	6,01	58,85	6	0,242
7	1,10	5,48	64,34	7	0,509
				8	0,397
				9	0,309
				10	0,484
				11	0,570
				12	0,438
				13	0,473
				14	0,479
				15	0,650
				16	0,580
				17	0,674
				18	0,606
				19	0,480
				20	0,198

Toplam açıklanan varyansa bakıldığında ölçekte özdeğeri 1’den büyük olan 7 faktör görülmektedir. İlk faktör (özdeğeri 4,94) varyansın %24,69’nu açıklarken ikinci faktör (özdeğeri 1,62) ise varyansın %8,07’ni açıklamaktadır. Faktör sayısına karar vermek için özdeğeri 1’den büyük olan faktörler dikkate alındığında test 7 faktörlü görülmektedir. Şekil 4.2’de verilen Form-II’e ait yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir.

PAY sonucu elde edilen tek faktörlü yapı, TBA sonucu elde edilen döndürülmemiş sonuçlar için özdeğerler ve faktör yükleri bir arada dikkate alınırsa birinci ve ikinci özdeğer arasındaki farka dayanarak ve yamaç grafiğinde birinci faktörden sonra

ani düşüşün olması fakat sonraki faktörlerden sonra nispeten yatay seyretmesinden dolayı testin tek boyutlu olduğu söylenebilir. Maddelerin birinci faktör üzerindeki yükleri, 6. ve 20. maddeler hariç 0,309 ile 0,650 arasındadır. 2 maddenin tek faktöre 0,30 altında yük vermesi, yaratıcılığı düşük madde yazarlarınca geliştirilen maddelerle oluşturulan testin yapı geçerliğine kanıt sunulmasında olumsuz etkide bulunmaktadır. Bu maddelerin öğrenciler tarafından yeterince anlaşılmadığı diğer bir deyişle yazarlarca yapıya uygun madde geliştirememiş olabileceği söylenebilir.



Şekil 4.2: Matematik Form-II'e Ait Yamaç Grafiği

İki teste ait AFA sonuçları birlikte değerlendirildiğinde, testlerin tek boyutlu yapıda oldukları ve benzeştikleri söylenebilir. Açıklanan varyansların (Form I için 21,65; Form II için 24,69) tek faktörlü ölçme araçlarında açıklanan varyansın %20 ve daha fazla olması ölçütüne (Reckase, 1979) uyması, PAY sonucu tek faktörlü yapı elde edilmesi ve öz değerler, faktör yükleri, yamaç grafikleri birlikte değerlendirildiğinde testlerin tek boyutlu olduğu iddia edilebilir. Ayrıca araştırmacı tarafından farklı faktör sayılarında yapılan analizlerde, maddelerin faktör yükleri esas alınarak yapılabilecek gruplandırmalar anlamlı sonuçlanmamıştır. Alanyazında kimi çalışmalarda da (Anıl ve Güzeller, 2011; Anıl, Güzeller, Çokluk & Şekercioğlu, 2010; Örs, 2010) bu tür testlerde tek faktörlü yapılar elde edilmiştir. Fakat bu sonuç, Kılıç ve Kelecioğlu'nun (2016) 2013-2014 Eğitim ve Öğretim Yılı

MOAS matematik alt testinin 2 boyutlu olduđu yönündeki bulgusu ile uyuşmamaktadır.

Testlerin yapı geçerliklerini istatistiksel olarak karşılaştırabilmek için aynı kazanıma yönelik maddelerin döndürülmemiş faktör yüklerinin karşılaştırılabileceđi düşünölmüştür. Maddelerin faktör yükleri tetrakorik korelasyon matrisinden elde edilen değerler olduğundan korelasyonların test edilmesinde kullanılan bağımsız gruplar için z testi ile karşılaştırılmış ve sonuçları Tablo 4.13'de sunulmuştur.

Tablo 4.13: Matematik Form I ve Form II'e Ait AFA Sonucu Elde Edilen Aynı Kazanımdaki Maddelerin Döndürülmemiş Faktör Yüklerinin z Testi Sonuçları

<i>Madde No</i>	<i>Form</i>	<i>Faktör Yüğü</i>	<i>z</i>	<i>p</i>
1	Form I	0,446	-0,98	0,33
	Form II	0.521		
2	Form I	0,501	-0,85	0,20
	Form II	0.562		
3	Form I	0,492	0,76	0,45
	Form II	0.432		
4	Form I	0,374	-3,82	0,00
	Form II	0.650		
5	Form I	0,508	2,01	0,04
	Form II	0.344		
6	Form I	0,523	3,34	0,00
	Form II	0.242		
7	Form I	0,497	-0,16	0,87
	Form II	0.509		
8	Form I	0,382	-0,18	0,86
	Form II	0.397		
9	Form I	0,493	2,21	0,03
	Form II	0.309		
10	Form I	0,430	-0,68	0,50
	Form II	0.484		
11	Form I	0,431	-1,86	0,06
	Form II	0.570		
12	Form I	0,465	0,34	0,73
	Form II	0.438		
13	Form I	0,415	-0,72	0,47
	Form II	0.473		
14	Form I	0,440	-0,50	0,62
	Form II	0.479		
15	Form I	0,427	-3,19	0,00
	Form II	0.650		
16	Form I	0,588	0,12	0,90
	Form II	0.580		
17	Form I	0,501	-2,67	0,01
	Form II	0.674		
18	Form I	0,490	-1,67	0,10
	Form II	0.606		
19	Form I	0,445	-0,45	0,66
	Form II	0.480		
20	Form I	0,403	2,27	0,02
	Form II	0.198		

Tablo 4.13 incelendiğinde iki testin aynı kazanımı yoklayan maddelerinin tek faktöre verdikleri faktör yükleri kıyaslandığında; 0,01 anlamlılık düzeyinde 3 madde için anlamlı, diğer 17 madde için farklılaşmadığı anlaşılmıştır. İki testin de tek faktörlü yapıda olmaları, açıklanan varyans ve yamaç grafiklerinin benzeşmeleri, tek faktöre iki madde haricinde 0,30 ve üzeri yük vermeleri nedenleri ile yapı geçerliklerine sunulan kanıtların benzeştiği söylenebilir.

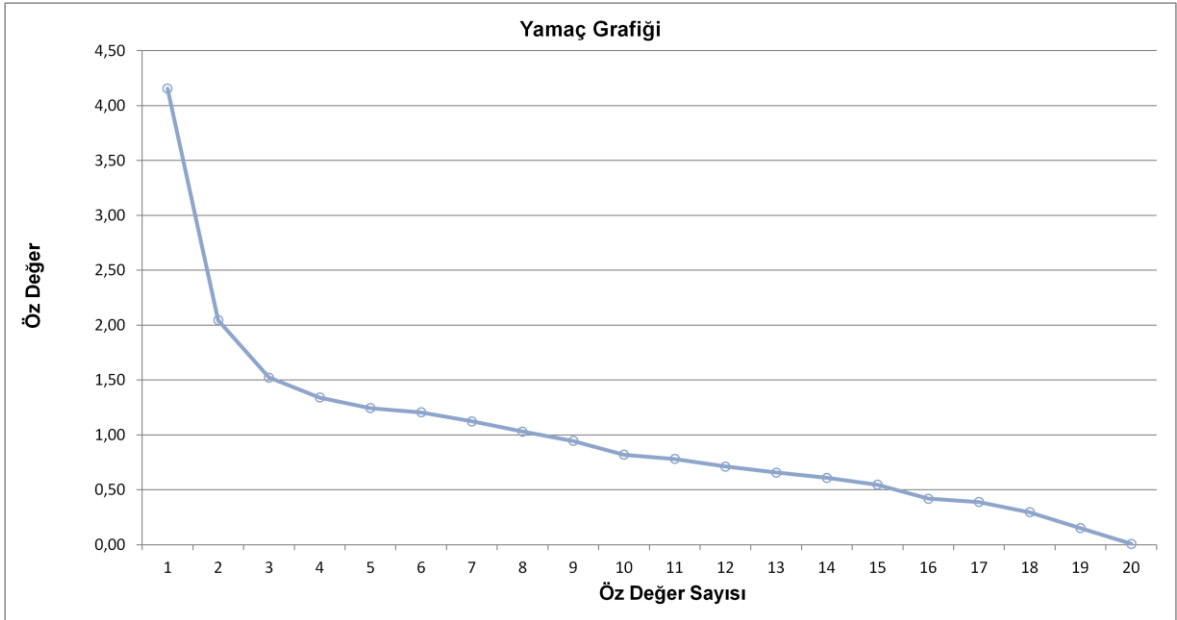
Türkçe

Yaratıcılığı yüksek madde yazarı Türkçe öğretmenlerin geliştirdikleri maddelerle oluşturulan testin maddelerine faktör analizi yapıp yapılamayacağını kontrolü için uygulanan analiz sonucu KMO değeri 0,67 ile orta düzeyde ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=381,9$; $sd=190$; $p=0,00$). Test puanlarına yapılan PAY sonucu tek faktör elde edilmiş, TBA sonuçları Tablo 4.14'de verilmiştir.

Tablo 4.14: Türkçe Form-I'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplam Varyans (%)</i>	<i>Madde No</i>	<i>Faktör Yük Değerleri</i>
1	4,15	20,77	20,77	1	0,604
2	2,05	10,23	31,00	2	0,429
3	1,52	7,61	38,61	3	0,456
4	1,34	6,70	45,32	4	0,466
5	1,24	6,22	51,54	5	0,476
6	1,21	6,03	57,57	6	0,381
7	1,12	5,62	63,19	7	0,342
8	1,03	5,15	68,34	8	0,282
				9	0,403
				10	0,401
				11	0,135
				12	0,573
				13	0,646
				14	0,388
				15	0,515
				16	0,331
				17	0,502
				18	0,402
				19	0,533
				20	0,550

Toplam açıklanan varyansa bakıldığında özdeğeri 1'den büyük olan 8 faktör görülmektedir. İlk faktör (özdeğeri 4,15) varyansın %20,77'sini açıklarken ikinci faktör (özdeğeri 2,05) ise varyansın %10,23'ünü açıklamaktadır. Faktör sayısına karar vermek için özdeğeri 1'den büyük olan faktörler dikkate alındığında test 8 faktörlü görülmektedir. Şekil 4.3'de verilen Form-1'e ait yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir. PAY sonucu elde edilen tek faktör, TBA sonucu ulaşılan döndürülmemiş sonuçlar için özdeğerler ve faktör yükleri bir arada dikkate alınırsa birinci ve ikinci özdeğer arasındaki farka da dayanarak ve yamaç grafiğinde birinci faktörden sonra ani düşüşün olması, müteakip faktörlerde ise nispeten yatay seyretmesinden dolayı testin tek boyutlu olduğu söylenebilir. Maddelerin, birinci faktör üzerindeki yükleri (8 ve 11. maddeler hariç) 0,331 ile 0,646 arasındadır. Söz konusu iki maddenin yapıyı ölçmede yetersiz olduğu, bu maddelerin ölçtüğü davranışlarla ilgili olarak madde yazarlarının uygun düzeyde madde geliştiremedikleri söylenebilir.



Şekil 4.3: Türkçe Form-1'e Ait Yamaç Grafiği

Yaratıcılığı düşük madde yazarı öğretmenlerin geliştirdikleri maddelerle oluşturulan testin maddelerine faktör analizi yapıp yapılamayacağını kontrol etmek amacıyla hesaplanan KMO değeri 0,74 ile iyi düzeyde ve Bartlett küresellik

testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=394,5$; $s_d=190$; $p=0,00$). Verilerin faktör analizine uygun olduğu görüldükten sonra uygulanan PAY sonucu test puanlarının tek faktörlü yapıda olduğu görülmüştür. TBA sonucunda, döndürme uygulanmadan elde edilen faktör sayısı, faktör öz değerleri, açıklanan varyans oranları ve maddelerin tek faktöre verdiği yükleri gösteren değerler Tablo 4.15'de verilmiştir.

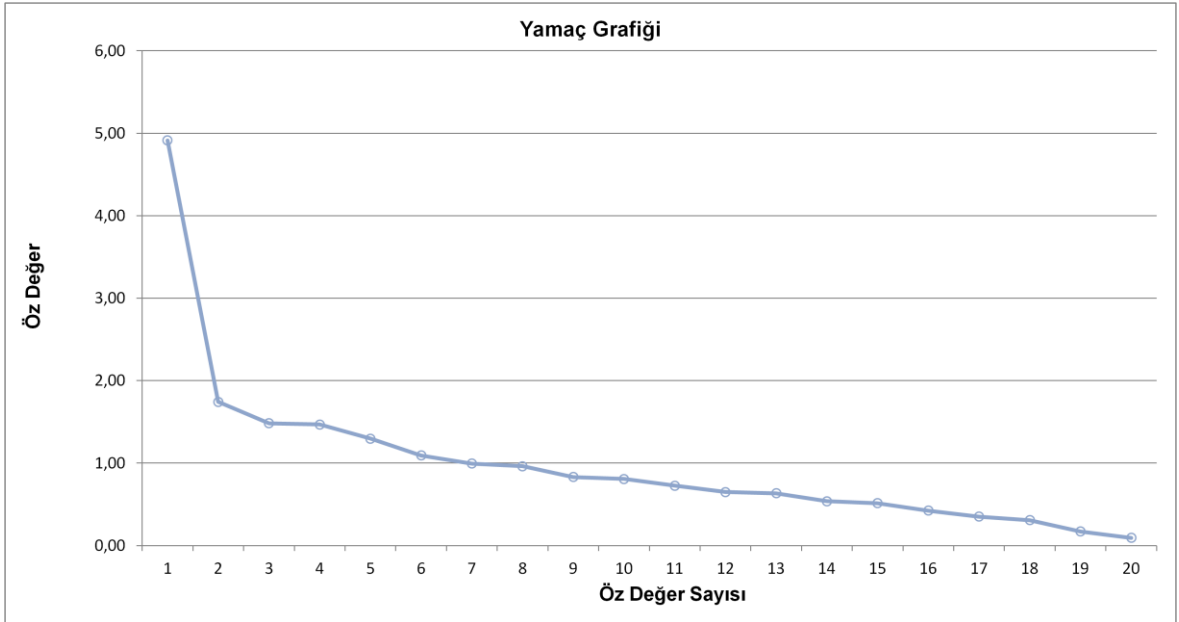
Tablo 4.15: Türkçe Form-II'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplam Varyans (%)</i>	<i>Madde No</i>	<i>Faktör Yük Değerleri</i>
1	4,91	24,57	24,57	1	0,481
2	1,74	8,71	33,28	2	0,378
3	1,48	7,41	40,69	3	0,438
4	1,47	7,34	48,04	4	0,533
5	1,30	6,48	54,52	5	0,412
6	1,09	5,47	59,98	6	0,668
				7	0,474
				8	0,564
				9	0,564
				10	0,758
				11	0,337
				12	0,585
				13	0,421
				14	0,483
				15	0,496
				16	0,395
				17	0,478
				18	0,304
				19	0,466
				20	0,452

Toplam açıklanan varyansa bakıldığında ölçekte özdeğeri 1'den büyük olan 6 faktör görülmektedir. İlk faktör (özdeğeri 4,91) varyansın %24,57'sini açıklarken ikinci faktör (özdeğeri 1,74) ise varyansın %8,71'ini açıklamaktadır. Faktör sayısına karar vermek için özdeğeri 1'den büyük olan faktörler dikkate alındığında test 6 faktörlü görülmektedir. Şekil 4.4'de verilen Form-II'e ait yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir.

PAY sonucu elde edilen tek faktörlü yapı ile TBA sonucunda elde edilen döndürülmemiş sonuçlar için özdeğerler ve faktör yükleri bir arada dikkate alınırsa birinci ve ikinci özdeğer arasındaki farka dayanarak ve yamaç grafiğinde birinci faktörden sonra ani düşüşün olması ve fakat sonraki faktörlerden sonra görece yatay seyretmesine dayanarak testin tek boyutlu olduğu söylenebilir. Maddelerin birinci faktör üzerindeki yükleri 0,304 ile 0,758 arasındadır. Maddelerin tek faktöre verdikleri yük değerlerine göre, yaratıcılığı düşük madde yazarlarının ölçülmek istenen yapıyı birlikte uygun şekilde ölçebilecek maddeler geliştirdiği şeklinde yorumlanabilir.

İki teste ait AFA sonuçları birlikte değerlendirildiğinde, testlerin tek boyutlu yapıda oldukları ve benzeştikleri söylenebilir. Bu bulgu Kılıç ve Kelecioğlu'nun (2016) MOAS'nın Türkçe alt testine uyguladıkları faktör analizi sonucu elde ettikleri tek boyutlu yapı ile de uyumaktadır.



Şekil 4.4: Türkçe Form-II'e Ait Yamaç Grafiği

Tablo 4.16'da verilen iki testin aynı kazanımı yoklayan maddelerinin tek faktörlü yapıda döndürülmemiş faktör yükleri kıyaslandığında; 0,01 anlamlılık düzeyinde 3 maddenin yaratıcılığı düşük; 1 maddenin ise yaratıcılığı yüksek grup lehine farklılaştığı görülmektedir. Bu farklılaşma 20 madde üzerinden değerlendirildiğinde, iki testin anlamlı derecede farklılaşmadığı söylenebilir.

Tablo 4.16: Türkçe Form I ve Form II'e Ait AFA Sonucu Elde Edilen Aynı Kazanımdaki Maddelerin Döndürülmemiş Faktör Yüklerinin z Testi Sonuçları

<i>Madde No</i>	<i>Form</i>	<i>Faktör Yükü</i>	<i>z</i>	<i>p</i>
1	Form I	0,604	1,76	0,08
	Form II	0,481		
2	Form I	0,429	0,61	0,54
	Form II	0,378		
3	Form I	0,456	0,23	0,83
	Form II	0,438		
4	Form I	0,466	-0,90	0,37
	Form II	0,533		
5	Form I	0,476	0,80	0,42
	Form II	0,412		
6	Form I	0,381	-4,09	0,00
	Form II	0,668		
7	Form I	0,342	-1,60	0,11
	Form II	0,474		
8	Form I	0,282	-3,51	0,00
	Form II	0,564		
9	Form I	0,403	-2,13	0,03
	Form II	0,564		
10	Form I	0,401	-5,71	0,00
	Form II	0,758		
11	Form I	0,135	-2,17	0,03
	Form II	0,337		
12	Form I	0,573	-0,18	0,86
	Form II	0,585		
13	Form I	0,646	3,22	0,00
	Form II	0,421		
14	Form I	0,388	-1,18	0,24
	Form II	0,483		
15	Form I	0,515	0,26	0,80
	Form II	0,496		
16	Form I	0,331	-0,74	0,46
	Form II	0,395		
17	Form I	0,502	0,32	0,75
	Form II	0,478		
18	Form I	0,402	1,13	0,26
	Form II	0,304		
19	Form I	0,533	0,90	0,37
	Form II	0,466		
20	Form I	0,550	1,32	0,19
	Form II	0,452		

İki alandaki test puanlarına yapılan AFA sonucu, yapı geçerlik kanıtlarının benzeştiği ve madde yazarlarının ölçülen yapıyı benzer düzeyde yoklayabilen, öğrencilerin eğitim düzeylerine uygunlukta benzer düzeyde anlaşılabilir maddeler geliştirdiği söylenebilir.

Bu bulgular ışığında yaratıcılığı yüksek ve düşük madde yazarlarının geliştirdikleri maddelerle oluşturulan testlerin yapı geçerlikleri, hem matematik ve hem de Türkçe alanında benzer düzeydedir.

4.2. Madde Yazarlığı Eğitiminin Madde ve Test Psikometrik Özelliklerine Etkisine Yönelik Bulgular

4.2.1. Eğitimin Madde Psikometrik Özelliklerine Etkisi

Matematik

Madde yazarlığı eğitiminin, eğitim alan öğretmenlerin geliştirdikleri maddelerin psikometrik özelliklerine etki edip etmediği, eğitim öncesi ve sonrası geliştirdikleri maddelerin psikometrik özelliklerinde farklılık olup olmadığının test edilmesi ile anlaşılabilen değeri değerlendirilmiş ve Form III ile Form V maddelerine ait psikometrik özellikler z testi ile karşılaştırılmıştır. Madde güçlük, madde ayırıcılık gücü ve madde çeldiricilik indeksleri arasında yapılan z testi sonuçları Tablo 4.17'de sunulmuştur.

Tablo 4.17: Matematik Öğretmenlerinin Madde Yazarlığı Eğitimi Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük, Ayırıcılık Gücü ve Çeldiricilik İndeksleri Arasındaki z Testi Sonuçları

Madde No	Form	Güçlük			Ayırıcılık			Çeldiricilik		
		p_i	z	p	r_{jx}	z	p	ζ	z	p
1	V	0,73	4,17	0,00	0,22	-0,64	0,52	0,61*	-7,66	0,00
	III	0,53			0,28			0,93		
2	V	0,43	0,00	1,00	0,20	-2,21	0,03	0,70	-7,69	0,00
	III	0,43			0,40			0,98		
3	V	0,67	3,48	0,00	0,22	-0,11	0,91	0,81	-3,24	0,00
	III	0,50			0,23			0,92		
4	V	0,75	5,78	0,00	0,10	0,10	0,92	0,71*	-3,40	0,00
	III	0,47			0,09			0,85		
5	V	0,35	-2,46	0,01	0,15	-2,14	0,03	0,74	-0,94	0,35
	III	0,47			0,35			0,78		
6	V	0,78	6,26	0,00	0,09	-1,23	0,22	0,71*	-2,11	0,03
	III	0,48			0,21			0,80		
7	V	0,79	6,30	0,00	0,06	-2,83	0,00	0,62*	-7,78	0,00
	III	0,49			0,33			0,94		
8	V	0,80	6,15	0,00	0,03	-2,04	0,04	0,73*	-6,77	0,00
	III	0,51			0,23			0,97		
9	V	0,79	6,30	0,00	0,12	-3,15	0,00	0,44*	-6,36	0,00
	III	0,49			0,41			0,75		
10	V	0,42	-1,01	0,31	0,17	-1,38	0,17	0,66	-2,93	0,00
	III	0,47			0,30			0,79		
11	V	0,29	-4,52	0,00	0,19	-1,96	0,05	0,72	-4,32	0,00
	III	0,51			0,37			0,89		
12	V	0,62	2,24	0,03	0,15	-2,14	0,03	0,67	-5,94	0,00
	III	0,51			0,35			0,91		
13	V	0,70	3,52	0,00	0,14	-1,91	0,06	0,78	-2,10	0,04
	III	0,53			0,32			0,86		
14	V	0,51	0,40	0,69	0,07	-2,28	0,02	0,64	-5,39	0,00
	III	0,49			0,29			0,87		
15	V	0,78	5,68	0,00	0,21	-0,21	0,83	0,67*	-6,55	0,00
	III	0,51			0,23			0,93		
16	V	0,48	0,40	0,69	0,26	-1,34	0,18	0,75	-3,67	0,00
	III	0,46			0,38			0,89		
17	V	0,75	5,59	0,00	0,17	-2,28	0,02	0,68*	-6,04	0,00
	III	0,48			0,38			0,92		
18	V	0,52	0,20	0,84	0,19	-1,17	0,24	0,67	-7,19	0,00
	III	0,51			0,30			0,95		
19	V	0,38	-1,63	0,10	0,16	-2,27	0,02	0,60	-6,98	0,00
	III	0,46			0,37			0,90		
20	V	0,46	-0,40	0,69	0,04	-4,83	0,00	0,69	-3,56	0,00
	III	0,48			0,48			0,84		

* $\zeta < \zeta_{min}$

Tablo 4.17 incelendiğinde madde güçlük indeksleri arasındaki z testi sonuçlarının 14 tanesi pozitif yani eğitim öncesi lehine, 5 tanesinin ise negatif olarak eğitim sonrası lehine olduğu görülmektedir. Anlamlı farklılıkların 10 tanesi eğitim öncesi lehine, 1 tanesi ise eğitim sonrası lehinedir. Bu bulgular eğitim öncesi kolay madde sayısının, eğitim sonrasında fazla olduğuna işaret etmektedir. Eğitim öncesi,

madde sayısının yarısının madde güçlük indeksinin yüksek; güçlük düzeyinin anlamlı şekilde kolay olduğu görülmüştür. Madde güçlük indeksleri eğitim sonrası anlamlı şekilde düşmüş, yani daha zor maddeler geliştirilmiştir.

Madde yazarlarının geliştirdiği maddelerin madde güçlük indekslerinin madde yazarlığı eğitim durumuna göre anlamlı bir farklılık gösterip göstermediği yoklanmıştır. Bu yoklamaya ait normal dağılımlı eğitim sonrası elde edilen madde güçlük indeksleri ($S-W_{FORM III} = 0,957$, $p = 0,483 > 0,05$) ile normal dağılımlı eğitim öncesi madde güçlükleri ($S-W_{FORM V} = 0,905$, $p = 0,052 > 0,05$) arasında yapılan t testi sonuçları Tablo 4.18'de sunulmuştur.

Tablo 4.18: Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları

<i>Grup</i>	<i>n</i>	\bar{x}	<i>Sx</i>	<i>Sd</i>	<i>t</i>	<i>p</i>	<i>Cohen's d</i>
Eğitim Sonrası	20	0,489	0,0253	38	-2,833	0,007	0,90
Eğitim Öncesi	20	0,600	0,1734				

Tablo 4.18'de görüldüğü üzere madde güçlük indeksleri eğitim sonrası, öncesine göre anlamlı bir farklılık göstermiştir ($t_{38}=-2,833$; $p<0,01$). Bu karşılaştırmaya ait etki büyüklüğü değeri 0,90 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu ve bu farkın örneklem büyüklüğünden kaynaklanmadığı söylenebilir. Eğitim öncesi ortalama madde güçlüğü 0,60 iken eğitim sonrası 0,49'a düşmüştür. Madde güçlük indeksi 0,50 veya bu civarda olan maddelerin, bu maddenin ölçtüğü nitelik yönünden cevaplayıcılar arasındaki farkları en iyi ortaya koyan maddeler olduğu söylenebilir (Baykul, 2015, s.226). Bu sebeple eğitim sonrası madde güçlüklerinin istendik düzeye yaklaştığı, madde yazarlığı eğitilmiş yazarların ideal orta güçlükte madde geliştirdikleri söylenebilir. Bu bulgu madde yazarlığı eğitiminin madde güçlüğüne olumlu yönde katkı sağladığı bulgusuna ulaşan diğer araştırmalarla uyumaktadır (Hingorjo & Jaleel, 2012; Iramaneerat, 2012).

Madde ayırıcılık gücü indeksleri arasındaki z testi sonuçlarınının 19 tanesi negatif yani eğitim sonrası lehine, 1 tanesinin ise pozitif olarak eğitim öncesi lehine olduğu görülmektedir. 10 anlamlı farklılığın tümü eğitim sonrası lehinedir. Bu bulgular eğitim sonrası geliştirilen maddelerin ayırıcılık gücü yüksek madde sayısının,

eğitim öncesine göre fazla olduğuna işaret etmektedir. Veriler 1 madde haricinde 19 maddede ayırıcılık gücünün arttığını ve bu artışın 10 tanesinin de anlamlı olduğuna işaret etmektedir. Madde yazarlarının geliştirdiği maddelerin madde ayırıcılık gücü indekslerinin madde yazarlığı eğitim durumuna göre anlamlı bir farklılık gösterip göstermediği yoklanmıştır. Bu yoklamaya ait normal dağılımlı eğitim sonrası elde edilen madde ayırıcılık gücü indeksleri ($S-W_{FORM III}=0,961$; $p=0,561>0,05$) ile normal dağılımlı eğitim öncesi madde güçlükleri ($S-W_{FORM V}=0,963$; $p=0,614>0,05$) arasında yapılan t testi sonuçları Tablo 4.19'da sunulmuştur.

Tablo 4.19: Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Ayırıcılık Güçlük İndekslerinin Farklılaşım Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları

<i>Grup</i>	<i>n</i>	\bar{x}	<i>Sx</i>	<i>Sd</i>	<i>t</i>	<i>p</i>	<i>Cohen's d</i>
Eğitim Sonrası	20	0,315	0,088	38	6,884	0,00	2,18
Eğitim Öncesi	20	0,147	0,065				

Tablo 4.19'da görüldüğü üzere madde ayırıcılık gücü indeksleri eğitim sonrası, öncesine göre anlamlı bir farklılık göstermiştir ($t_{38}=6,884$; $p<0,01$). Bu karşılaştırmaya ait etki büyüklüğü değeri 2,18 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu ve bu farkın örneklem büyüklüğünden kaynaklanmadığı söylenebilir. Eğitim öncesi ortalama madde ayırıcılık gücü 0,147 iken eğitim sonrası 0,315'e yükselmiştir. Bir maddenin teste alınabilmesi için 0,30 ve üzerinde ayırıcılık gücü indeksine sahip olması gerektiğinden eğitim sonrası madde ayırıcılık güçlüklerinin ortalamalarının bu düzeye ulaştığı, madde yazarlığı eğitilmiş yazarların kazanıma sahip öğrenciler ile sahip olmayanları daha iyi ayırabilen maddeler geliştirdiği söylenebilir. Bu bulgu Hingorjo ve Jaleel (2012) ile Iramaneerat'in (2012) bulgularıyla paralellik göstermektedir.

Tablo 4.17'de görüleceği üzere eğitim sonrası geliştirilen maddelerin tümü kabul edilebilir çeldiricilik indeksi üzerinde çeldiricilik indeksine sahiptirler. Eğitim öncesi ise 8 madde kabul edilebilir çeldiricilik indeksinin altında çeldiricilik indeksindedir. Madde çeldiricilik indeksleri arasındaki z testi sonuçlarının tümünün negatif yani eğitim sonrası lehine olduğu görülmektedir. Toplam 17 anlamlı farklılık eğitimin madde çeldiriciliğini arttırdığı başka bir deyişle madde yazarlığı eğitiminin,

matematik öğretmenlerinin daha dengeli işleyen çeldirici geliştirmesini sağladığı söylenebilir.

Madde yazarlarının geliştirdiği maddelerin madde çeldiricilik indekslerinin madde yazarlığı eğitim durumuna göre anlamlı bir farklılık gösterip göstermediği yoklanmıştır. Bu yoklamaya ait normal dağılımlı eğitim sonrası elde edilen madde çeldiricilik indeksleri ($S-W_{FORM III}=0,951$; $p=0,379>0,05$) ile normal dağılımlı eğitim öncesi madde güçlükleri ($S-W_{FORM V}=0,908$; $p=0,058>0,05$) arasında yapılan t testi sonuçları Tablo 4.20'de sunulmuştur.

Tablo 4.20: Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Çeldiricilik İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları

<i>Grup</i>	<i>n</i>	\bar{x}	<i>Sx</i>	<i>Sd</i>	<i>t</i>	<i>p</i>	<i>Cohen's d</i>
Eğitim Sonrası	20	0,884	0,065	38	8,936	0,00	2,83
Eğitim Öncesi	20	0,680	0,078				

Tablo 4.20'de görüldüğü üzere madde çeldiricilik indeksleri eğitim sonrası, öncesine göre anlamlı bir farklılık göstermiştir ($t_{38}=8,936$; $p<0,01$). Bu karşılaştırmaya ait etki büyüklüğü değeri 2,83 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu ve bu farkın örneklem büyüklüğünden kaynaklanmadığı söylenebilir. Eğitim öncesi ortalama çeldiriciliği 0,680 iken eğitim sonrası 0,884'e yükselmiştir. Eğitim sonrası madde çeldiriciliklerinin artması; madde yazarlığı eğitiminin yazarlarda daha dengeli işleyen çeldirici üretmesini sağladığı söylenebilir.

Türkçe

Türkçe öğretmenlerinin eğitim sonrası ve aynı öğretmenlerin eğitim öncesi geliştirdikleri maddelere ait güçlük, madde ayırıcılık gücü ve madde çeldiricilik indeksleri arasında yapılan z testi sonuçları Tablo 4.21'de sunulmuştur.

Tablo 4.21: Türkçe Öğretmenlerinin Madde Yazarlığı Eğitimi Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük, Ayırcılık Gücü, Çeldiricilik İndekslerine Arasındaki z Testi Sonuçları

Madde No	Form	p_i	Güçlük		r_{jx}	Ayırcılık		Ç	Çeldiricilik	
			z	p		z	p		z	p
1	Form III	0,37	0,00	1,00	0,29	-0,91	0,36	0,93	7,90	0,00
	Form V	0,37			0,37			0,60*		
2	Form III	0,50	-3,71	0,00	0,24	0,42	0,67	0,83	5,56	0,00
	Form V	0,68			0,20			0,58*		
3	Form III	0,35	-7,53	0,00	0,39	2,11	0,03	0,98	9,30	0,00
	Form V	0,72			0,20			0,61*		
4	Form III	0,61	-3,05	0,00	0,36	0,57	0,57	0,93	5,20	0,00
	Form V	0,75			0,31			0,74		
5	Form III	0,44	-6,41	0,00	0,40	0,81	0,42	0,95	6,29	0,00
	Form V	0,75			0,33			0,72		
6	Form III	0,45	-1,92	0,00	0,28	0,75	0,45	0,93	6,92	0,00
	Form V	0,98			0,21			0,70*		
7	Form III	0,53	5,61	0,00	0,56	2,92	0,00	0,90	4,87	0,00
	Form V	0,26			0,33			0,71		
8	Form III	0,42	-2,24	0,03	0,37	0,35	0,73	0,77	0,00	1,00
	Form V	0,53			0,34			0,77*		
9	Form III	0,57	-5,27	0,00	0,31	0,44	0,66	0,78	-0,25	0,80
	Form V	0,81			0,27			0,79*		
10	Form III	0,44	-1,63	0,10	0,31	-1,40	0,16	0,81	4,07	0,00
	Form V	0,52			0,43			0,63*		
11	Form III	0,55	-3,58	0,00	0,43	3,52	0,00	0,65	-1,98	0,05
	Form V	0,72			0,11			0,74*		
12	Form III	0,56	3,25	0,00	0,29	-0,11	0,91	0,68	-1,57	0,12
	Form V	0,40			0,30			0,75*		
13	Form III	0,50	-3,08	0,00	0,18	-0,85	0,40	0,89*	6,84	0,00
	Form V	0,65			0,26			0,71		
14	Form III	0,47	-1,83	0,07	0,26	-0,33	0,74	0,94	6,73	0,00
	Form V	0,56			0,29			0,68*		
15	Form III	0,61	3,45	0,00	0,31	-0,68	0,50	0,82	4,52	0,00
	Form V	0,44			0,37			0,62		
16	Form III	0,48	-3,48	0,00	0,35	0,78	0,44	0,92	5,89	0,00
	Form V	0,65			0,28			0,69		
17	Form III	0,61	2,25	0,02	0,28	-0,90	0,37	0,74	1,77	0,08
	Form V	0,50			0,36			0,66		
18	Form III	0,41	-3,86	0,00	0,28	1,06	0,29	0,74	1,56	0,12
	Form V	0,60			0,18			0,67*		
19	Form III	0,70	4,74	0,00	0,38	1,24	0,22	0,70	-0,67	0,50
	Form V	0,47			0,27			0,73		
20	Form III	0,37	1,73	0,08	0,28	-0,22	0,83	0,66	-0,43	0,67
	Form V	0,29			0,30			0,68*		

*Ç<Çmin

Tablo 4.21 incelendiğinde madde güçlük indeksleri arasındaki z testi sonuçlarının 6 tanesi pozitif yani eğitim sonrası lehine, 13 tanesinin ise negatif olarak eğitim öncesi lehine olduğu görülmektedir. Anlamlı farklılıkların 3 tanesi eğitim sonrası lehine, 11 tanesi ise eğitim öncesi lehinedir. Bu sonuçlar eğitim öncesi kolay madde sayısının, eğitim sonrasında fazla olduğuna işaret etmektedir. Eğitim

öncesi, madde sayısının yarısının madde güçlük indeksinin yüksek; güçlük düzeyinin anlamlı şekilde kolay olduğu görülmüştür. Madde güçlük indeksleri eğitim sonrası anlamlı şekilde düşmüş, yani daha zor maddeler geliştirilmiştir.

Madde yazarlarının geliştirdiği maddelerin madde güçlük indekslerinin madde yazarlığı eğitim durumuna göre anlamlı bir farklılık gösterip göstermediği yoklanmıştır. Bu yoklamaya ait normal dağılımlı eğitim sonrası elde edilen madde güçlük indeksleri ($S-W_{FORM III}=0,962$; $p=0,591>0,05$) ile normal dağılımlı eğitim öncesi madde güçlükleri ($S-W_{FORM V}=0,889$; $p=0,206>0,05$) arasında yapılan t testi sonuçları Tablo 4.22'de sunulmuştur.

Tablo 4.22: Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Güçlük İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları

<i>Grup</i>	<i>n</i>	\bar{x}	<i>Sx</i>	<i>Sd</i>	<i>t</i>	<i>p</i>	<i>Cohen's d</i>
Eğitim Sonrası	20	0,456	0,1695	38	-5,522	0,00	1,75
Eğitim Öncesi	20	0,724	0,1356				

Tablo 4.22'de görüldüğü üzere madde güçlük indeksleri ortalaması eğitim sonrası, öncesine göre anlamlı bir farklılık göstermiştir ($t_{38}=-5,522$; $p<0,01$). Bu karşılaştırmaya ait etki büyüklüğü değeri 1,75 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu ve bu farkın örneklem büyüklüğünden kaynaklanmadığı söylenebilir. Eğitim öncesi ortalama madde güçlüğü 0,72 iken eğitim sonrası 0,46'ya düşmüştür. Eğitim sonrası madde güçlüklerinin istendik düzeye yaklaştığı, madde yazarlığı eğitilmiş yazarların daha beklendiği düzeyde güçlükte madde geliştirdikleri söylenebilir. Bu bulgu matematik alanındaki bulgular ile uyuşmaktadır.

Tablo 4.21 incelendiğinde madde ayırıcılık gücü indeksleri arasındaki z testi sonuçlarının 12 tanesi pozitif yani eğitim sonrası lehine, 8 tanesinin ise negatif olarak eğitim öncesi lehine olduğu görülmektedir. Bu değerler eğitim sonrası geliştirilen maddelerin ayırıcılık gücü yüksek madde sayısının, eğitim öncesine göre fazla olduğuna işaret etmektedir. Analiz sonuçları 8 maddenin ayırıcılık gücünün düştüğünü 12 maddede ise arttığını ve bu artışın 2 tanesinin de $\alpha=0,01$ düzeyinde anlamlı olduğuna işaret etmektedir. Madde yazarlarının geliştirdiği maddelerin madde ayırıcılık gücü indekslerinin madde yazarlığı eğitim durumuna göre anlamlı bir farklılık gösterip göstermediği yoklanmıştır. Bu yoklamaya ait

normal dağılımlı eğitim sonrası elde edilen madde ayırıcılık gücü indeksleri ($S-W_{FORM III}=0,923$; $p=0,114>0,05$) ile normal dağılımlı eğitim öncesi madde ayırıcılık gücü indeksleri ($S-W_{FORM V}=0,975$; $p=0,856>0,05$) arasında yapılan t testi sonuçları Tablo 4.23'de sunulmuştur.

Tablo 4.23: Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Ayırıcılık Gücü İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları

<i>Grup</i>	<i>n</i>	\bar{x}	<i>Sx</i>	<i>Sd</i>	<i>t</i>	<i>p</i>
Eğitim Sonrası	20	0,328	0,082	38	1,676	0,102
Eğitim Öncesi	20	0,286	0,077			

Tablo 4.23'de görüldüğü üzere madde ayırıcılık gücü indeksleri eğitim sonrası, öncesine göre anlamlı bir farklılık göstermemektedir, $t(38)=1,676$; $p>0,01$. Eğitim öncesi ortalama madde ayırıcılık gücü 0,286 iken eğitim sonrası 0,328'e yükselmiştir. Eğitim sonrası madde ayırıcılık gücü indekslerinin ortalaması 0,30 değerinin üzerine çıksa da, eğitim öncesine göre anlamlı farklılık yaratmamıştır. Türkçe alanında madde yazarlığı eğitimi gören yazarların kazanıma sahip öğrenciler ile sahip olmayanları daha iyi ayırabilen maddeler geliştirdiği ortalama madde güçlük indeksleri değerine dayanarak söylenebilir fakat bu istatistiksel olarak anlamlı bir farklılaşma değildir.

Tablo 4.21'de görüleceği üzere eğitim öncesi kabul edilebilir çeldiricilik indeksi altında 12 madde bulunurken bu sayı eğitim sonrası 1'e düşmüştür. Madde çeldiricilik indeksleri arasındaki z testi sonuçlarından 14 tanesi pozitif yani eğitim sonrası lehine; 5 tanesi negatif olarak eğitim öncesi lehinedir. Anlamlı 12 farklılığın tümü eğitim sonrası lehinedir. Bu sonuçlar eğitimin madde çeldiriciliğini arttırdığı başka bir deyişle madde yazarlığı eğitiminin, Türkçe öğretmenlerinin daha dengeli işler çeldiricili maddeler geliştirmesini sağladığı söylenebilir. Madde yazarlarının geliştirdiği maddelerin madde çeldiricilik indekslerinin ortalamalarının madde yazarlığı eğitim durumuna göre anlamlı bir farklılık gösterip göstermediği yoklanmıştır. Bu yoklamaya ait normal dağılımlı eğitim sonrası elde edilen madde çeldiricilik indeksleri ($S-W_{FORM III}=0,875$; $p=0,014>0,01$) ile normal dağılımlı eğitim öncesi madde çeldiricilik indeksleri ($S-W_{FORM V}=0,937$; $p=0,210>0,01$) arasında yapılan t testi sonuçları Tablo 4.24'de sunulmuştur.

Tablo 4.24: Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdikleri Maddelerin Çeldiricilik İndekslerinin Farklılaşp Farklılaşmadığına Yönelik Yapılan Bağımsız Gruplar t Testi Sonuçları

<i>Grup</i>	<i>n</i>	\bar{x}	<i>Sx</i>	<i>Sd</i>	<i>t</i>	<i>p</i>	<i>Cohen's d</i>
Eğitim Sonrası	20	0,858	0,082	38	8,389	0,00	2,65
Eğitim Öncesi	20	0,664	0,062				

Tablo 4.24'de görüldüğü üzere madde çeldiricilik indeksleri eğitim sonrası, öncesine göre anlamlı bir farklılık göstermiştir, $t(38)=8,389$; $p<0,01$. Bu karşılaştırmaya ait etki büyüklüğü değeri 2,65 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu ve bu farkın örneklem büyüklüğünden kaynaklanmadığı söylenebilir. Eğitim öncesi ortalama çeldiricilik 0,664 iken eğitim sonrası 0,858'e yükselmiştir. Eğitim sonrası madde çeldiriciliklerinin artması; madde yazarlığı eğitiminin yazarlarda daha dengeli işleyen çeldirici üretmesini sağladığı şeklinde yorumlanabilir.

Bu bulgular ışığında, madde yazarı matematik ve Türkçe öğretmenlerinin geliştirdiği maddelerin güçlüğü ve çeldiriciliği eğitim sonrası lehine, eğitim öncesine göre anlamlı bir şekilde farklılaşmıştır. Matematik öğretmenlerinin geliştirdiği maddelerin ayırıcılık güçleri eğitim sonrası lehine anlamlı bir şekilde farklılaşırken, Türkçe alanında farklılaşmamıştır.

Matematik ve Türkçe alanlarında, madde yazarlığı eğitimi sonrası geliştirilen maddelerin güçlük ve çeldiricilik indeksleri eğitim sonrası lehine anlamlı şekilde farklılaşmıştır. Bu bulgular Abdulghani vd. (2015) ile Vinay vd.'nin (2016) araştırma bulgularına uyumludur. Maddelerin istendik güçlük düzeyi olan 0,50'ye yaklaşımlarında ve dengeli işleyen çeldirici bulabilmede madde yazarlığı eğitiminin etkili olduğunu öne sürebilecek bulgular elde edilmiştir. Madde ayırıcılık gücü indeksleri matematik alanında anlamlı şekilde eğitim sonrası lehine farklılaşmış, Türkçe alanında ortalama madde ayırıcılık gücü indeksi artmasına rağmen bu artış anlamlı olmamıştır. Bu bulgulara göre madde yazarlığı eğitiminin maddenin ayırıcılık gücüne etkisi bakımından alanlara göre farklılaştığı şeklinde yorumlanabilir. Bu bulgu Pate & Caldwell'in (2014) araştırma bulguları ile de uyumludur. Madde yazarlığı eğitimi içerisinde ağırlıklı üzerinde durulan madde yazım kurallarına uygun madde geliştirmenin de elde edilen bulgularda etkili olduğu söylenebilir.

4.2.2. Eğitimin Testin Psikometrik Özelliklerine Etkisi

Matematik

Matematik öğretmenlerinin eğitim sonrası geliştirdiği maddelerle oluşturulan Form III ve aynı öğretmenlerin eğitim öncesi geliştirdikleri maddelerle oluşturulan Form V'in psikometrik özelliklerinin farklılık gösterip göstermediği analiz edilmiş ve analiz sonuçları Tablo 4.25'de gösterilmiştir.

Tablo 4.25: Matematik Öğretmenlerinin Eğitim Sonrası ve Öncesi Geliştirdiği Maddelerle Oluşturulan Testlerin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları

Formlar	KR-20	z	Cohen's q	S_x^2	F	Cohen's d	\bar{X}	U	Cohen's d
Form III	0,753	4,48**	0,45	17,53	2,22**	0,15	9,79	14013**	0,58
Form V	0,489			7,89			12,01		

** $p < 0,01$

Tablo 4.25'de görüldüğü üzere eğitim sonrası geliştirilen maddelerle oluşturulan testin iç tutarlık güvenilirliği 0,75 değeri ile uygun düzeyde olduğu söylenebilir (Davidshofer & Murphy, 2005, s.150). Testlerin KR-20 iç tutarlık güvenilirlikleri arasında yapılan z testi sonucu eğitim sonrası geliştirilen maddelerle oluşturulan test lehine anlamlı bulunmuştur ($z=4,48$; $p=0,00 < 0,01$). Bu karşılaştırmaya ait etki büyüklüğü değeri ise 0,45 olarak elde edilmiştir. Buna göre orta derecede etki büyüklüğüne sahip olduğu ve bu farkın sadece örneklem büyüklüğünden kaynaklanmadığı söylenebilir. Bu bulgu madde yazarlığı eğitiminin birbirleriyle daha tutarlı madde geliştirmede etken olduğu şeklinde yorumlanabilir.

Testlerin varyansları arasında yapılan F testi sonucu, eğitim sonrası geliştirilen maddelerle oluşturulan test lehine manidar düzeyde fark bulunmuştur ($S_1^2/S_2^2=17,53/7,89=2,222$; $F_{0,005;202;202}=2,222 < F_U=1.439$). Bu durum eğitim sonrası oluşturulan testi yanıtlayan öğrenci puanlarının öncesi gruba göre daha heterojen olduğuna işaret eder. Benzer başarı düzeyindeki öğrencilerde bu farklılığın oluşmasının altında yatan sebebin, eğitim sonrası oluşturulan testin, öncesine göre daha ayrıştırıcı başka bir ifade ile öğrenciler arası farklılığı daha belirgin ortaya koyabilmesinden kaynaklandığı ileri sürülebilir. Bu karşılaştırmaya ait etki büyüklüğü değeri 0,15 olarak elde edilmiştir. Buna göre varyanslar arasındaki farkın küçük bir etki büyüklüğüne sahip olduğu, yani bu farkın örneklem büyüklüğünden kaynaklanmış olabileceği söylenebilir.

Testlerin aritmetik ortalamalarının karşılaştırılmasında puan dağılımlarının normallikleri test edilmiş ve hem eğitim sonrası (S-W=0,977; $p=0,002<0,05$) hem de eğitim öncesi (S-W=0,986; $p=0,046<0,05$) puan dağılımlarının normal dağılımlı olmadığı görülmüştür. Bu sebeple aritmetik ortalamaların eğitim öncesi ve sonrası arasında farklılaşp farklılaşmadığının yoklanması için yapılan ilişkisiz ölçümler için Mann Whitney U testi sonucu eğitim öncesi geliştirilen maddelerle oluşturulan test lehine manidar düzeyde fark bulunmuştur ($U=14013$; $p=0,00<0,01$; $SO_{Form III}=171,03$; $SO_{Form IV}=235,97$). Bu sonuç öğrencilerin eğitim öncesi geliştirilen maddelerle oluşturulan testten, eğitim sonrası oluşturulan teste göre daha yüksek puan aldıklarını, maddelerin daha kolay geldiğini göstermektedir. Bu karşılaştırmaya ait etki büyüklüğü değeri 0,58 olarak elde edilmiştir. Buna göre orta derecede etki büyüklüğüne sahiptir ve bu farklılaşma yalnızca örneklem büyüklüğünden kaynaklanmamaktadır denebilir.

Madde yazarlığı eğitimine tâbi tutulan öğretmenlerin geliştirdikleri maddelerle oluşturulan testlerin kapsam geçerliklerini incelemek amacıyla 14 alan uzmanının görüşlerinden elde edilen veriler kullanılmış ve hesaplanan KGO değerleri ve aynı kazanımı yoklayan maddeler için yapılan iki oranı test etmede kullanılan z testi sonuçları Tablo 4.26'da sunulmuştur.

Tablo 4.26: Madde Yazarığı Eğitimi Alan Matematik Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdiğı Maddelerle Oluşturulan Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları

<i>Madde No</i>	<i>Form</i>	<i>Gerekli</i>	<i>Yararlı/ Yetersiz</i>	<i>Gereksiz</i>	<i>KGO</i>	<i>z</i>	<i>p</i>
1	III	14	0	0	1,000*	1,47	0,14
	V	13	1	0	0,857*		
2	III	14	0	0	1,000*	8,37	0,00
	V	4	10	0	-0,429		
3	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
4	III	14	0	0	1,000*	1,47	0,14
	V	13	1	0	0,857*		
5	III	12	2	0	0,714*	3,94	0,00
	V	7	4	3	0,000		
6	III	12	2	0	0,714*	0,00	1,00
	V	12	2	0	0,714*		
7	III	12	2	0	0,714*	3,94	0,00
	V	7	6	1	0,000		
8	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
9	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
10	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
11	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
12	III	14	0	0	1,000*	1,47	0,14
	V	13	1	0	0,857*		
13	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
14	III	14	0	0	1,000*	2,16	0,03
	V	12	2	0	0,714*		
15	III	13	1	0	0,857*	-1,47	0,14
	V	14	0	0	1,000*		
16	III	14	0	0	1,000*	3,94	0,00
	V	9	5	0	0,286		
17	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
18	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		
19	III	14	0	0	1,000*	7,10	0,00
	V	5	6	3	-0,286		
20	III	14	0	0	1,000*	0,00	1,00
	V	14	0	0	1,000*		

* $p < 0,05$

Tablo 4.26 incelendiğinde 14 uzmanının görüşleri doğrultusunda KGO'na $\alpha=0,05$ anlamlılık düzeyinde katkı yapan maddeler Form-III için tümü; Form-IV için 2, 5, 7, 16, 19. maddeler haricindeki diğer 15 maddedir. Aynı kazanımı yoklayan maddelerin KGO değerlerinin z testi sonuçlarına göre 9 tanesi pozitif yani eğitim sonrası lehine; 1 tanesi ise negatif olarak eğitim öncesi lehinedir. 5 anlamlı farklılıkların tümü eğitim sonrası lehinedir. 10 madde için ise KGO değerleri eşittir. KGİ, kritik değer üzerindeki maddelerin KGO ortalamalarından hesaplanmış ve Form III için 0,950; Form V için ise 0,933 değeri bulunmuştur. Elde edilen KGİ değerlerine ilişkisiz iki oranı test etmeye yarayan z testi uygulanmış; 0,01 anlamlılık düzeyinde iki test arasında anlamlı fark bulunmamıştır ($z=0,21$; $p=0,83>0,01$). Gerek aynı kazanımı yoklayan maddelerin KGO değerleri, gerekse tüm teste ait KGİ değerlerinin anlamlı derecede farklılaştığına dair kanıt elde edilmemiştir.

8. Sınıf öğrencilerinin MOAS sonuçları ile eş formlardan aldıkları puanların korelasyon analizi yapılarak, eş formların uyum geçerliğine kanıt aranmıştır. Form III ve V'e ait puan dağılımlarının normallikten sapma göstermesi sebebiyle korelasyon analizinde parametrik olmayan korelasyon tekniklerinden SBSFK kullanılmış ve sonuçları Tablo 4.27'de verilmiştir.

Tablo 4.27: Öğrencilerin Merkezî Sınav Matematik Alt Testi Puanları ile Eş Formlardan Aldıkları Matematik Puanlarının Korelasyon Analizi Sonuçları

<i>Formlar</i>	<i>MOAS</i>	<i>z</i>	<i>Cohen's q</i>
Form III	0,910**	7,058**	0,71
Form IV	0,676**		

** $p<0,01$

Tablo 4.27 incelendiğinde 8.sınıf öğrencilerinin, Form III puanları ile MOAS matematik puanları arasında yüksek düzeyde, pozitif ve anlamlı ($\rho_{\text{FormIII,MOASIII}}=0,910$; $p=0,00<0,01$); Form V puanları ile MOAS matematik puanları arasında orta düzeyde, pozitif ve anlamlı ($\rho_{\text{FormV,MOASV}}=0,676$; $p=0,00<0,01$) bir ilişki olduğu görülmektedir. Eğitim sonrası geliştirilen maddelerle oluşturulan testin korelasyon katsayısının Ebel & Frisbie'nin (1991, s.106) uyum geçerliği için önerdiği değerden yüksek çıkması eğitimin uyum geçerliğine olumlu etki yaptığını işaretler. İki korelasyon katsayısı arasındaki z testi sonucu ($z=7,058$; $p<0,01$) eğitim

sonrası lehine anlamlı fark bulunmuştur. Bu karşılaştırmaya ait etki büyüklüğü değeri 0,71 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahiptir ve örneklem büyüklüğünden kaynaklanmamaktadır yorumu yapılabilir. Matematik öğretmenlerinin madde yazarlığı eğitimi sonrası geliştirdikleri maddelerden oluşturulan testin, eğitim öncesine göre MOAS ile daha iyi uyum gösterdiği ve geçerliğine kanıt olarak sunulabileceği söylenebilir.

Türkçe

Türkçe öğretmenlerinin eğitim sonrası geliştirdiği maddelerle oluşturulan Form III ve aynı öğretmenlerin eğitim öncesi geliştirdikleri maddelerle oluşturulan Form V test psikometrik özelliklerinin farklılık gösterip göstermediği incelenmiş ve Tablo 4.28'de gösterilmiştir.

Tablo 4.28: Türkçe Öğretmenlerinin Eğitim Sonrası ve Öncesi Geliştirdiği Testlerin İstatistiklerin Farklılaşmasına Yönelik Yapılan Analiz Sonuçları

<i>Formlar</i>	<i>KR-20</i>	<i>z</i>	<i>p</i>	<i>S_x²</i>	<i>F</i>	<i>Cohen's d</i>	<i>\bar{X}</i>	<i>t</i>	<i>Cohen's d</i>
Form III	0,765	1,034	0,30	16,182	1,404**	0,12	9,131	-13,612**	1,34
Form V	0,719			11,521			14,471		

** $p < 0,01$

Tablo 4.28 incelendiğinde testlerin KR-20 iç tutarlık güvenilirlikleri arasında yapılan z testi sonuçlarına göre anlamlı fark bulunmamıştır ($z=1,034$; $p > 0,01$). Madde yazarlığı eğitiminin, her ne kadar eğitim sonrası geliştirilen testin iç tutarlık güvenilirliği 0,75 üzeri değere sahip olsa da, Türkçe alanında birbirleriyle tutarlı ölçme yapan madde geliştirmede etkili olmadığı söylenebilir.

Varyansları arasında yapılan F testi sonucu, eğitim sonrası geliştirilen maddelerle oluşturulan test lehine manidar düzeyde fark bulunmuştur ($S_1^2/S_2^2=16,182/11,521=1,404$; $F_{0,005;205;205}=1,404 < F_U=1.439$). Bu durum eğitim sonrası oluşturulan testi yanıtlayan öğrenci puanlarının öncesi gruba göre daha heterojen olduğuna işaretler. Benzer başarı düzeyindeki öğrencilerde bu farklılığın oluşmasının altında yatan sebebin, eğitim sonrası oluşturulan testin, öncesine göre daha ayrıştırıcı başka bir ifade ile öğrenciler arası farklılığı daha belirgin ortaya koyabilmesinden kaynaklandığı ileri sürülebilir. Bu karşılaştırmaya ait etki büyüklüğü değeri 0,12 olarak elde edilmiştir. Buna göre varyanslar arasındaki

farkın küçük bir etki büyüklüğüne sahip olduğu, yani bu farkın örneklem büyüklüğünden kaynaklanmış olabileceği söylenebilir.

Puan dağılımları normallik sınırları içerisinde olan testlerin aritmetik ortalamalar arasındaki ilişkisiz iki örneklem t testi sonucu eğitim öncesi geliştirilen maddelerle oluşturulan test lehine manidar düzeyde fark bulunmuştur ($t=-13,612$; $p<0,01$). Bu sonuç eğitim öncesi oluşturulan testin öğrencilere, eğitim sonrasına göre daha kolay geldiğini, maddelerinin güçlük indekslerinin daha büyük olduğu tespitini destekler niteliktedir. Bu karşılaştırmaya ait etki büyüklüğü değeri 1,34 olarak elde edilmiştir. Buna göre yüksek derecede etki büyüklüğüne sahip olduğu ve örneklem büyüklüğünden kaynaklanmadığı söylenebilir.

Madde yazarlığı eğitimine tâbi tutulan öğretmenlerin geliştirdikleri maddelerle oluşturulan testlerin kapsam geçerliklerini incelemek amacıyla 12 alan uzmanının görüşlerinden elde edilen seçme frekansları, hesaplanan KGO değerleri ve aynı kazanımı yoklayan maddeler için yapılan iki oranı test etmede kullanılan z testi sonuçları Tablo 4.29'da sunulmuştur.

Tablo 4.29: Madde Yazarlığı Eğitimi Alan Türkçe Öğretmenlerinin Eğitim Öncesi ve Sonrası Geliştirdiği Maddelerle Oluşturulan Testlerin Kapsam Geçerlik Çalışmasına Yönelik Uzmanlarca Seçme Frekansları, KGO Değerleri, z Testi Sonuçları

<i>Madde No</i>	<i>Form</i>	<i>Gerekli</i>	<i>Yararlı/ Yetersiz</i>	<i>Gereksiz</i>	<i>KGO</i>	<i>z</i>	<i>p</i>
1	III	11	1	0	0,833*	-1,48	0,14
	V	12	0	0	1,000*		
2	III	11	1	0	0,833*	0,94	0,35
	V	10	2	0	0,667*		
3	III	11	1	0	0,833*	-1,48	0,14
	V	12	0	0	1,000*		
4	III	11	1	0	0,833*	0,00	1,00
	V	11	1	0	0,833*		
5	III	12	0	0	1,000*	2,19	0,03
	V	10	2	0	0,667*		
6	III	12	0	0	1,000*	3,47	0,00
	V	8	2	2	0,333		
7	III	12	0	0	1,000*	2,19	0,03
	V	10	2	0	0,667*		
8	III	11	1	0	0,833*	0,00	1,00
	V	11	1	0	0,833*		
9	III	11	0	0	0,833*	-1,48	0,14
	V	12	0	0	1,000*		
10	III	11	1	0	0,833*	0,00	1,00
	V	11	1	0	0,833*		
11	III	11	1	0	0,833*	0,00	1,00
	V	11	1	0	0,833*		
12	III	12	0	0	1,000*	0,00	1,00
	V	12	0	0	1,000*		
13	III	12	0	0	1,000*	2,83	0,00
	V	9	3	0	0,500		
14	III	12	0	0	1,000*	0,00	1,00
	V	12	0	0	1,000*		
15	III	11	1	0	0,833*	0,00	1,00
	V	11	1	0	0,833*		
16	III	12	0	0	1,000*	0,00	1,00
	V	12	0	0	1,000*		
17	III	11	0	0	0,833*	-1,48	0,14
	V	12	2	0	1,000*		
18	III	12	0	0	1,000*	0,00	1,00
	V	12	0	0	1,000*		
19	III	11	1	0	0,833*	0,00	1,00
	V	11	1	0	0,833*		
20	III	11	1	0	0,833*	0,00	1,00
	V	11	1	0	0,833*		

* $p < 0,05$

Tablo 4.29 incelendiğinde 12 uzmanın görüşleri doğrultusunda KGO'na $\alpha=0,05$ anlamlılık düzeyinde katkı yapan maddeler Form-III için tümü; Form-IV için 6. ve

13. maddeler haricindeki diğer 18 maddedir. Aynı kazanımı yoklayan maddelerin KGO değerlerinin z testi sonuçlarına göre 5 tanesi pozitif yani eğitim sonrası lehine; 4 tanesi ise negatif olarak eğitim öncesi lehinedir. 2 anlamlı farkın ikisi de eğitim sonrası lehinedir. 11 madde için ise KGO değerleri eşittir.

Test bazında KGİ, kritik değer üzerindeki maddelerin KGO ortalamalarından Form-III için 0,900; Form-V için ise 0,880 hesaplanmıştır. Elde edilen KGİ değerlerine ilişkisiz iki oranı test etmeye yarayan z testi uygulanmış; iki test arasında anlamlı fark bulunmamıştır ($z=0,20$; $p=0,84>0,01$). Gerek aynı kazanımı yoklayan maddelerin KGO değerleri, gerekse tüm teste ait KGİ değerlerinin anlamlı derecede farklılaştığına dair kanıt elde edilmemiştir.

8. Sınıf öğrencilerinin MOAS sonuçları ile eş formlardan aldıkları puanların korelasyon analizi yapılarak, eş formların uyum geçerliğine kanıt aranmasında kullanılan PMÇK sonuçları Tablo 4.30'da verilmiştir.

Tablo 4.30: Öğrencilerin Merkezî Sınav Türkçe Alt Testi Puanları ile Eş Formlardan Aldıkları Türkçe Puanlarının Korelasyon Analizi Sonuçları

<i>Formlar</i>	<i>MOAS</i>	<i>z</i>	<i>Cohen's q</i>
Form III	0,941**	7,76**	0,78
Form IV	0,749**		

** $p<0,01$

Tablo 4.30 incelendiğinde 8.sınıf öğrencilerinin, Form III puanları ile MOAS Türkçe puanları arasında yüksek düzeyde, pozitif ($r=0,941$) ve anlamlı ($p=0,00<0,01$); Form V puanları ile MOAS Türkçe puanları arasında yüksek düzeyde, pozitif ($r=0,749$) ve anlamlı ($p=0,00<0,01$) bir ilişki olduğu görülmektedir. Matematik alanında olduğu gibi Türkçe alanında da madde yazarlığı eğitiminin uyum geçerliğine olumlu katkı yaptığı söylenebilir. İki korelasyon katsayısı arasında yapılan z testi sonucu ($z=7,76$; $p<0,01$) eğitim sonrası lehine anlamlı fark bulunmuştur. Bu karşılaştırmaya ait etki büyüklüğü 0,78 ile büyük etki düzeyinde bulunmuş ve farklılığın örneklem büyüklüğünden kaynaklanmadığı düşünülmüştür. Türkçe öğretmenlerinin madde yazarlığı eğitimi sonrası geliştirdikleri maddelerden oluşturulan testin, eğitim öncesine göre MOAS ile daha iyi uyum gösterdiği söylenebilir.

Bu bulgular ışığında, madde yazarı matematik ve Türkçe öğretmenlerinin geliştirdiği maddelerle oluşturulan testlerin varyans ve uyum geçerlikleri eğitim sonrası lehine; aritmetik ortalamaları eğitim öncesi lehine anlamlı farklılaşmış; kapsam geçerliği ise farklılaşmamıştır. Matematik alanında güvenilirlik eğitim sonrası lehine anlamlı şekilde farklılaşırken, Türkçe alanında ise anlamlı şekilde farklılaşmamıştır.

Eğitim sonrası geliştirilen maddelerle oluşturulan testlerin özellikleri, eğitim öncesi oluşturulan testlerin özellikleri ile karşılaştırılması sonucu, iki alanda da; varyans, aritmetik ortalama ve uyum geçerliklerine madde yazarlığı eğitiminin anlamlı etkide bulunduğu görülmüştür. Madde yazarlığı eğitiminin ölçülen kapsamı daha iyi veya daha yetersiz düzeyde birlikte ölçebilen madde geliştirmeye etki etmediği anlaşılmıştır. Hâlbuki madde yazarlığı eğitiminde kazanımı ölçebilecek madde geliştirmenin önemi öğretmenlere vurgulanmasına rağmen her iki alanda da kapsam geçerliğinde farklılaşmama bulgusuna ulaşılması, eğitimin bu konuda yetersiz olmuş olabileceğini akla getirmektedir. İç tutarlık güvenilirliği ile ilgili bulgular dikkate alındığında, madde yazarlığı eğitimi matematik alanında etkili olmuş, Türkçe alanında ise anlamlı farklılaşmaya dönük kanıt elde edilmemiştir. Madde yazarlığı eğitiminin birbirleriyle daha tutarlı ölçme yapabilecek madde geliştirmede yani testlerin güvenilirliğini daha üst düzeylere çıkarmada alanlara bağlı olduğu, her alanda benzer etkiyi yapamadığı şeklinde yorumlanabilir. Araştırmada iki alandaki tüm öğretmenlere aynı kapsamda madde yazarlığı eğitimi verilmiştir. Bu şekildeki bir bulgunun altında yatan neden, madde yazarlığı eğitiminin farklı alanlarda farklı kapsamlarda verilmesi gerekliliği olabilir.

4.2.3. Eğitimin Yapı Geçerliğine Etkisi

Matematik

Eğitim sonrası matematik öğretmenlerinin geliştirdikleri maddelerle oluşturulan testin maddelerine faktör analizi yapıp yapılamayacağını kontrol etmek amacıyla yapılan analiz sonucu KMO değeri 0,80 ile çok iyi düzeyde ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=2067,5$; $sd=190$; $p=0,00$). PAY sonucu üç faktörlü yapı görülmüş, test puanlarına uygulanan TBA sonuçlarını gösteren değerler Tablo 4.31'de verilmiştir.

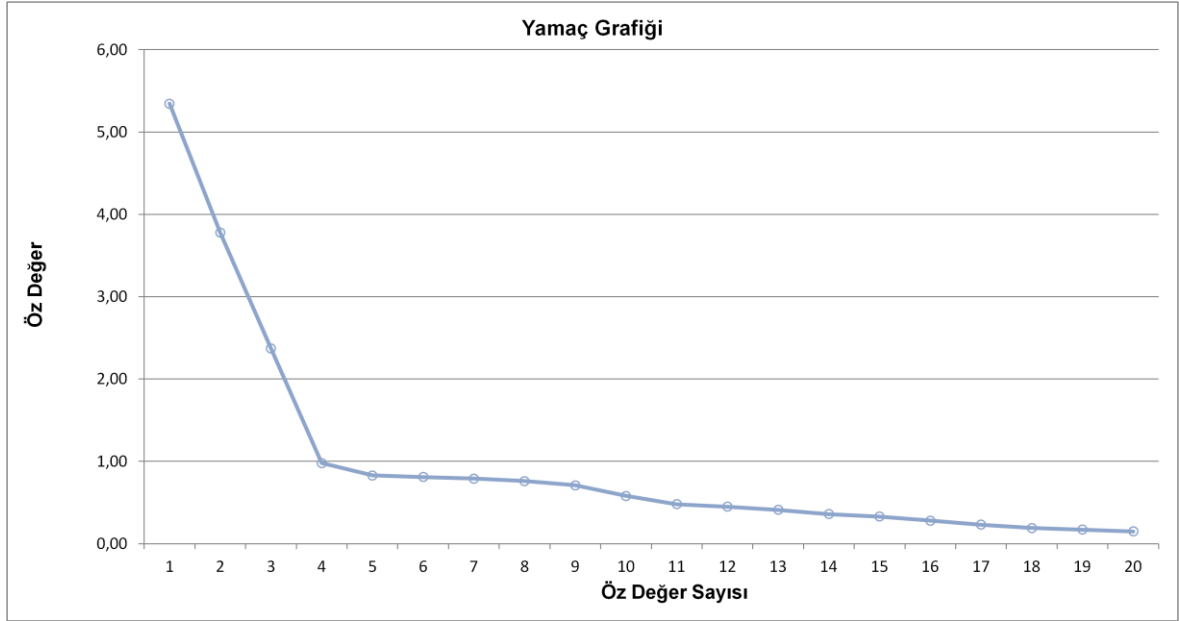
Tablo 4.31: Matematik Form-III'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplam Varyans (%)</i>	<i>Madde No</i>	<i>Faktör-1 Yüğü</i>	<i>Faktör-2 Yüğü</i>	<i>Faktör-3 Yüğü</i>
1	5,35	26,72	26,72	1		0,456	
2	3,78	18,91	45,63	2		0,529	
3	2,37	11,86	57,49	3			0,734
				4			0,718
				5		0,533	
				6			0,782
				7		0,492	
				8			0,798
				9		0,577	
				10		0,473	
				11	0,686		
				12		0,548	
				13	0,538		
				14	0,599		
				15	0,643		
				16	0,595		
				17	0,604		
				18		0,455	
				19	0,639		
				20	0,568		

Tablo 4.31 incelendiğinde, analize alınan 20 maddenin öz değeri 1'den büyük olan üç faktör altında toplandığı görülmektedir. Bu üç faktörün test puanlarına ilişkin açıkladıkları varyans %57,49'dur. Maddelerle ilgili olarak tanımlanan iki faktörün ortak varyanslarının (communalities) ise 0,473 ile 0,685 arasında değiştiği gözlenmektedir. Buna göre, analizde önemli faktör olarak ortaya çıkan üç faktörün birlikte, maddelerdeki toplam varyansın ve teste ilişkin varyansın çoğunluğunu açıkladıkları anlaşılmaktadır.

Analizde önemli faktör sayısı, özdeğer ölçütüne göre üç olarak tanımlanmıştır. Bu durum, özdeğerlere göre çizilen Şekil 4.5'deki Form-III'e ait çizgi grafiğinde de açıkça görülmektedir. Grafikte, dördüncü faktöre kadar ivmeli bir düşüş, bu faktörden itibaren ise yatay devam eden, önemli bir düşüş eğilimi göstermeyen bir seyir gözlemlenmektedir. Diğer bir deyişle, dördüncü ve sonraki faktörlerin varyansa olan katkıları birbirlerine yakındır denebilir. Gerek özdeğer gerekse çizgi grafiği birlikte yorumlandığında matematik öğretmenlerinin eğitim sonrası geliştirdikleri maddelerle oluşturulan teste öğrencilerin verdiği yanıtların puanlanması ile edilen test puanlarının üç boyutlu olduğu söylenebilir. Bu bulgu

alanyazındaki benzer testlere yönelik yapılan çalışmalardan farklılaşmaktadır (Anıl ve Güzeller, 2011; Anıl vd., 2010; Kılıç ve Kelecioğlu, 2016; Örs, 2010).



Şekil 4.5: Matematik Form-III'e Ait Yamaç Grafiği

Faktör yükleri incelendiğinde, birinci faktöre 11, 13, 14, 15, 16, 17, 19, 20. maddeler 0,538-0,686 arasında; ikinci faktöre 1, 2, 5, 7, 9, 10, 12, 18. maddeler 0,455-0,577; üçüncü faktöre 3, 4, 6, 8. maddeler 0,718-0,798 arasında yük vermektedir. Maddelerin yük verdikleri faktörler dikkate alındığında, birinci faktörün “Kareköklü ifadeler” alt öğrenme alanına ait kazanımları yokladığı; ikinci faktörün 5’inci madde hariç “Üslü ifadeler” alt öğrenme alanına ait kazanımları yokladığı; üçüncü faktörün “Çarpanlar ve katlar” alt öğrenme alanına ait kazanımları yokladığı görülmektedir. 5’inci madde “Çarpanlar ve katlar” alt öğrenme alanındaki bir kazanımı yoklamak üzere geliştirilmesine rağmen “Üslü ifadeler” alt öğrenme alanına daha fazla yük vermektedir. Bu maddenin ölçülmek istenen alt öğrenme alanı yerine diğer bir alt öğrenme alanına daha fazla yük vermesi, madde yazarının uygun düzeyde madde geliştiremediğine bağlanabilir. Fakat bu durum, alan uzmanlarınca birbirlerinin tamamlayıcısı olarak görülen, “üslü ifadeler” ile “çarpanlar ve katlar” alt öğrenme alanları için anlamsız görülmemektedir. Ayrıca 16. madde “Kareköklü ifadeler” alt öğrenme alanı için geliştirilen bir madde

olmasına rağmen “Üslü ifadeler” alt öğrenme alanına da önemli derecede yük (0,492) vermektedir.

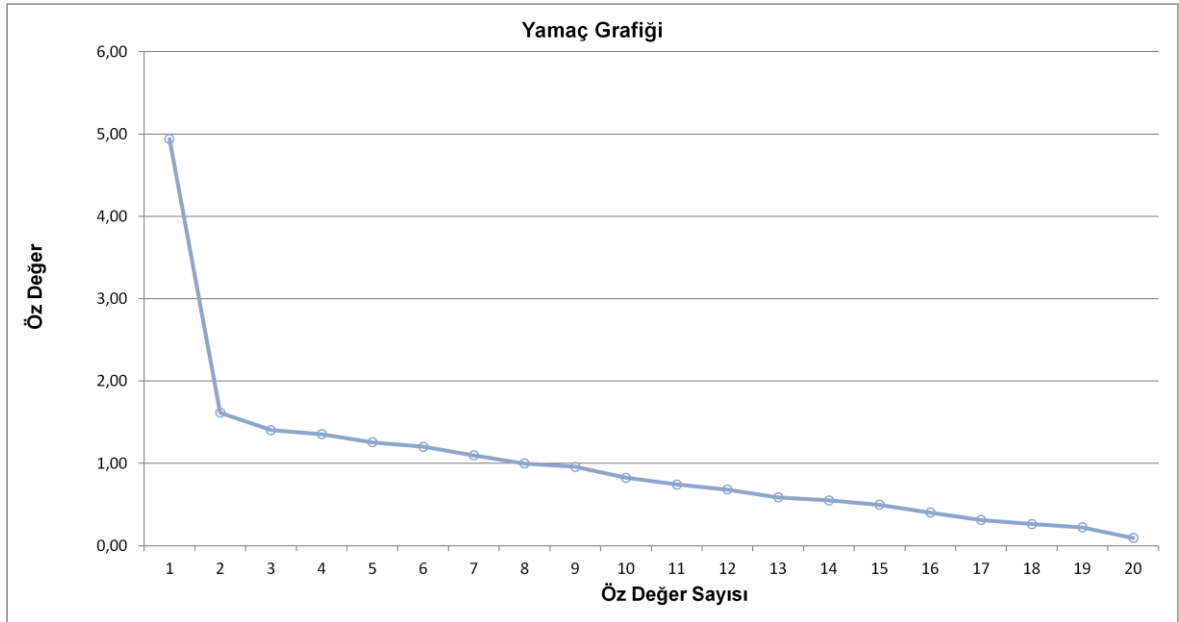
Madde yazarlığı eğitimi öncesi, madde yazarı matematik öğretmenlerinin geliştirdikleri maddelerle oluşturulan testin maddelerine faktör analizi yapıp yapılamayacağını kontrol etmek amacıyla hesaplanan KMO değeri 0,71 ile iyi düzeyde ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=386,705$; $sd=190$; $p=0,00$). Verilerin faktör analizine uygun olduğu görüldükten sonra uygulanan PAY sonucu test puanlarının tek faktörlü olduğu görülmüştür. Test puanlarının tetrakorik korelasyon matrisine dayalı, TBA sonucunda, döndürme uygulanmadan elde edilen faktör sayısı, faktör öz değerleri, açıklanan varyans oranları ve maddelerin tek(ilk) faktöre verdiği yükleri gösteren değerler Tablo 4.32’de verilmiştir.

Tablo 4.32: Matematik Form-V’e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplamlı Varyans (%)</i>	<i>Madde No</i>	<i>Faktör Yük Değerleri</i>
1	4,60	21,17	21,17	1	0,490
2	1,60	7,36	28,53	2	0,484
3	1,47	6,78	35,31	3	0,376
4	1,37	6,29	41,60	4	0,494
5	1,24	5,69	47,29	5	0,410
6	1,18	5,43	52,72	6	0,547
7	1,07	4,90	57,62	7	0,427
8	1,00	4,62	62,24	8	0,591
				9	0,378
				10	0,418
				11	0,414
				12	0,357
				13	0,384
				14	0,445
				15	0,452
				16	0,537
				17	0,473
				18	0,416
				19	0,483
				20	0,418

Toplam açıklanan varyansa bakıldığında ölçekte özdeğeri 1’den büyük olan 8 faktör görülmektedir. İlk faktör (özdeğeri 4,60) varyansın %21.17’sini açıklarken

ikinci faktör (özdeğeri 1,60) ise varyansın %7,36'sını açıklamaktadır. Faktör sayısına karar vermek için özdeğeri 1'den büyük olan faktörler dikkate alındığında test 8 faktörlü görülmektedir. Şekil 4.6'da verilen Form-V'e ait yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir.



Şekil 4.6: Matematik Form-V'e Ait Yamaç Grafiği

Döndürülmemiş sonuçlar için özdeğerler ve faktör yükleri bir arada dikkate alınırsa birinci ve ikinci özdeğer arasındaki farka dayanarak; yamaç grafiğinde birinci faktörden sonra ani düşüşün olması ve ikinci faktörden itibaren nispeten yatay seyretmesinden; PAY sonucu tek faktör elde edilmesinden dolayı testin tek boyutlu olduğu söylenebilir. Maddelerin birinci faktör üzerindeki yükleri 0,357 ile 0,591 arasındadır. Bu bulgu, Kılıç ve Kelecioğlu'nun (2016) MOAS matematik alt testinin iki boyutlu olduğuna dair bulgusu ile uyuşmamaktadır.

İki teste ait AFA sonuçları birlikte değerlendirildiğinde, eğitim öncesi geliştirilen testin tek; sonrası geliştirilenin ise üç boyutlu yapıda olduğu ve farklılaştıkları söylenebilir. Eğitim sonrası geliştirilen testin, alt öğrenme alanlarını daha iyi açıklayarak yapı geçerliliğine kanıt sunduğu ileri sürülebilir.

Türkçe

Eğitim sonrası öğretmenlerin geliştirdikleri maddelerle oluşturulan testin maddelerine faktör analizi yapıp yapılamayacağını kontrol etmek amacıyla KMO değerinden ve Bartlett testinden yararlanılmıştır. Veri analizi sonucu KMO değeri 0,78 ile iyi düzeyde ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=535,1$; $sd=190$; $p=0,00$). Verilerin faktör analizine uygun olduğu görülmüş ve uygulanan PAY sonucu tek faktörlü yapı gözlemlenmiştir.

Test puanlarının tetrakorik korelasyon matrisine dayalı, TBA sonucunda, döndürme uygulanmadan elde edilen faktör sayısı, faktör öz değerleri, açıklanan varyans oranları ve maddelerin tek faktöre verdiği yükleri gösteren değerler Tablo 4.33'de verilmiştir.

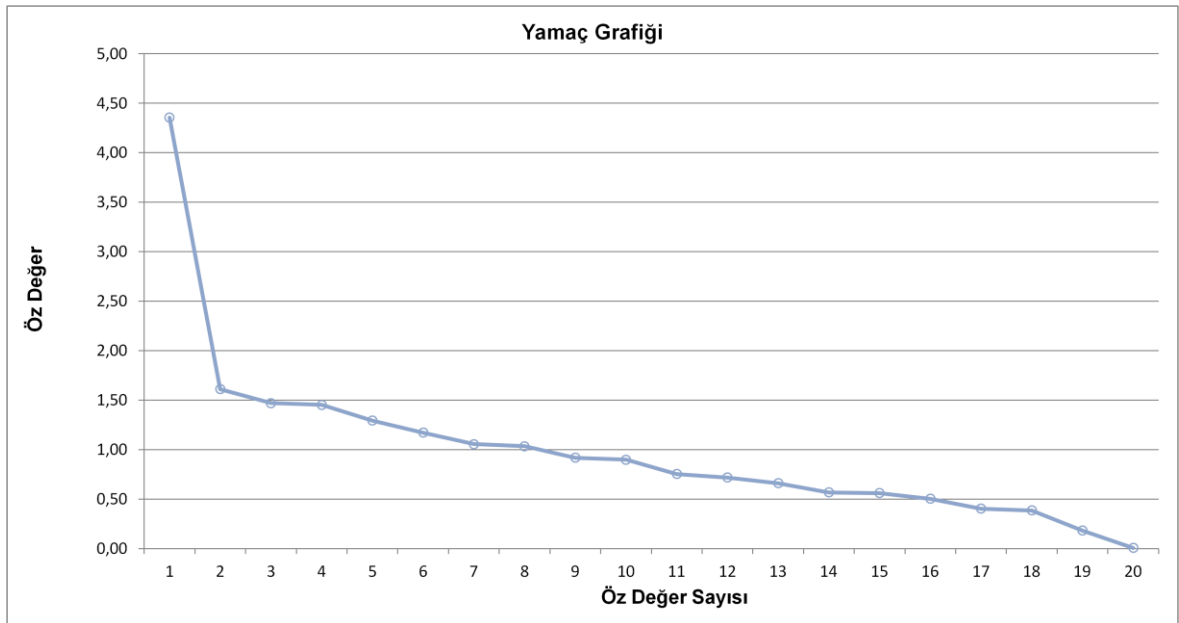
Tablo 4.33: Türkçe Form-III'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplam Varyans (%)</i>	<i>Madde No</i>	<i>Faktör Yüğü</i>
1	4,36	21,78	21,78	1	0,331
2	1,61	8,06	29,84	2	0,266
3	1,47	7,35	37,19	3	0,529
4	1,45	7,26	44,45	4	0,550
5	1,29	6,46	50,92	5	0,520
6	1,17	5,86	56,77	6	0,364
7	1,06	5,28	62,05	7	0,885
8	1,03	5,17	67,22	8	0,529
				9	0,394
				10	0,396
				11	0,602
				12	0,404
				13	0,243
				14	0,377
				15	0,392
				16	0,495
				17	0,420
				18	0,378
				19	0,470
				20	0,387

Toplam açıklanan varyansa bakıldığında ölçekte özdeğeri 1'den büyük olan 8 faktör görülmektedir. İlk faktör (özdeğeri 4,36) varyansın %21,78'sini açıklarken ikinci faktör (özdeğeri 1,61) ise varyansın %8,06'sını açıklamaktadır. Faktör sayısına karar vermek için özdeğeri 1'den büyük olan faktörler dikkate alındığında test 8 faktörlü görülmektedir. Şekil 4.7'de verilen Form-III'e ait yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat

diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir.

Analizde önemli faktör sayısı, özdeğer ölçütüne göre sekiz olarak tanımlanmıştır. Şekil 4.7'deki yamaç grafiğinde, ikinci faktöre kadar ivmeli bir düşüş, bu faktörden itibaren ise yatay devam eden, önemli bir düşüş eğilimi göstermeyen bir seyir gözlemlenmektedir. Diğer bir deyişle, ikinci ve sonraki faktörlerin varyansa olan katkıları birbirlerine yakındır denebilir. PAY sonucu elde edilen tek faktör, gerek özdeğer gerekse çizgi grafiği birlikte yorumlandığında Türkçe öğretmenlerinin eğitim sonrası geliştirdikleri maddelerle oluşturulan teste öğrencilerin verdiği yanıtların puanlanması ile edilen test puanlarının tek boyutlu olduğu söylenebilir. Bu bulgu Kılıç ve Kelecioğlu'nun (2016) MOAS Türkçe alt testinin tek boyutlu olduğu yönündeki bulgusu ile de uyuşmaktadır.



Şekil 4.7: Türkçe Form-III'e Ait Yamaç Grafiği

Maddelerin birinci faktör üzerindeki yükleri 2. ve 13. maddeler hariç 0,331 ile 0,885 arasındadır. Söz konusu iki madde ise yapı geçerliğine olumsuz yönde etkide bulunmaktadır.

Madde yazarlığı eğitimi öncesi, madde yazarı Türkçe öğretmenlerinin geliştirdikleri maddelerle oluşturulan testin maddelerine faktör analizi yapılıp yapılamayacağını

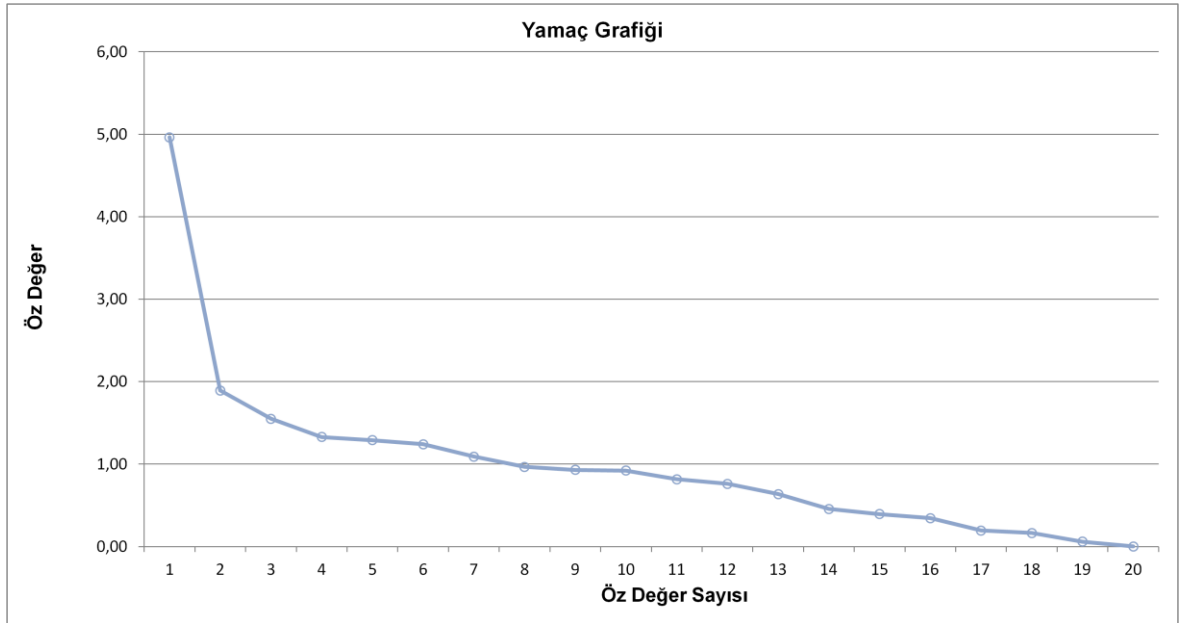
kontrol etmek amacıyla hesaplanan KMO değeri 0,68 ile orta düzeyde ve Bartlett küresellik testi sonucu ise $\alpha=0,01$ düzeyinde manidar bulunmuştur ($\chi^2=478,8$; $s_d=190$; $p=0,00$). Faktörleşmeye uygunluğu görülen test puanlarına uygulanan PAY sonucu tek faktörlü yapı gözlemlenmiştir. Test puanlarının tetrakorik korelasyon matrisine dayalı, TBA sonucunda, döndürme uygulanmadan elde edilen faktör sayısı, faktör öz değerleri, açıklanan varyans oranları ve maddelerin tek faktöre verdiği yükleri gösteren değerler Tablo 4.34'de verilmiştir.

Tablo 4.34: Türkçe Form-V'e Ait AFA Sonucu Elde Edilen Faktör Öz Değerleri, Açıklanan Varyans ve Faktör Yük Değerleri

<i>Faktörler</i>	<i>Faktör Öz Değerleri</i>	<i>Açıklanan Varyans (%)</i>	<i>Toplam Varyans (%)</i>	<i>Madde No</i>	<i>Faktör Yük Değerleri</i>
1	4,96	24,81	24,81	1	0,583
2	1,89	9,46	34,27	2	0,344
3	1,55	7,75	42,02	3	0,357
4	1,33	6,65	48,67	4	0,488
5	1,29	6,45	55,12	5	0,528
6	1,24	6,21	61,33	6	0,395
7	1,09	5,45	66,78	7	0,490
				8	0,558
				9	0,464
				10	0,773
				11	0,197
				12	0,549
				13	0,465
				14	0,490
				15	0,601
				16	0,477
				17	0,596
				18	0,333
				19	0,452
				20	0,533

Toplam açıklanan varyansa bakıldığında ölçekte özdeğeri 1'den büyük olan 7 faktör görülmektedir. İlk faktör (özdeğeri 4,96) varyansın %24,81'ini açıklarken ikinci faktör (özdeğeri 1,89) ise varyansın %9,46'sını açıklamaktadır. Faktör sayısına karar vermek için özdeğeri 1'den büyük olan faktörler dikkate alındığında test 7 faktörlü görülmektedir. Şekil 4.8'de verilen Form-V'e ait yamaç grafiği dikkate alındığında, birinci faktörden sonra ani ve ivmeli bir düşüş olduğu fakat diğer faktörlerden sonra grafiğin nispeten yatay olarak devam ettiği ve ani bir düşüş olmadığı görülmektedir.

Döndürülmemiş sonuçlar için özdeğerler ve faktör yükleri bir arada dikkate alınırsa birinci ve ikinci özdeğer arasındaki farka dayanarak ve yamaç grafiğinde birinci faktörden sonra ani düşüşün olmasından dolayı testin tek boyutlu olduğu söylenebilir. Bu bulgu Kılıç ve Kelecioğlu'nun (2016) MOAS Türkçe alt testinin tek boyutlu olduğu yönündeki bulgusu ile de uyuşmaktadır. Maddelerin birinci faktör üzerindeki yükleri (11. madde hariç) 0,333 ile 0,773 arasındadır. Söz konusu maddenin yapıyı ölçme konusundaki yetersizliği madde yazarının uygun şekilde geliştirememesine bağlanabilir.



Şekil 4.8: Türkçe Form-V'e Ait Yamaç Grafiği

İki teste ait AFA sonuçları birlikte değerlendirildiğinde, eğitim öncesi ve sonrası geliştirilen testlerin tek boyutlu yapıda oldukları ve farklılaşmadıkları söylenebilir. Tablo 4.35'de iki testin aynı kazanımı yoklayan maddelerinin tek faktörlü yapıda döndürülmemiş faktör yükleri kıyaslandığında; 0,01 anlamlılık düzeyinde, 2 maddenin eğitim sonrası; 3 maddenin ise eğitim öncesi lehine farklılaştığı görülmüştür. Açıklanan varyans, boyutlanma ve faktör yükleri birlikte değerlendirildiğinde anlamlı farklılık gözlemlenmemiştir.

Tablo 4.35: Türkçe Form III ve Form V'e Ait AFA Sonucu Elde Edilen Aynı Kazanımdaki Maddelerin Döndürülmemiş Faktör Yüklerinin z Testi Sonuçları

<i>Madde No</i>	<i>Form</i>	<i>Faktör Yükü</i>	<i>z</i>	<i>p</i>
1	III	0,331	-3,26	0,00
	V	0,583		
2	III	0,266	-0,87	0,39
	V	0,344		
3	III	0,529	2,17	0,03
	V	0,357		
4	III	0,550	0,86	0,39
	V	0,488		
5	III	0,520	-0,11	0,91
	V	0,528		
6	III	0,364	-0,37	0,72
	V	0,395		
7	III	0,885	8,69	0,00
	V	0,490		
8	III	0,529	-0,42	0,68
	V	0,558		
9	III	0,394	-0,87	0,39
	V	0,464		
10	III	0,396	-6,13	0,00
	V	0,773		
11	III	0,602	5,00	0,00
	V	0,197		
12	III	0,404	-1,90	0,06
	V	0,549		
13	III	0,243	-2,58	0,01
	V	0,465		
14	III	0,377	-1,41	0,16
	V	0,490		
15	III	0,392	-2,83	0,00
	V	0,601		
16	III	0,495	0,24	0,81
	V	0,477		
17	III	0,420	-2,41	0,02
	V	0,596		
18	III	0,378	0,52	0,60
	V	0,333		
19	III	0,470	0,23	0,82
	V	0,452		
20	III	0,387	-1,87	0,06
	V	0,533		

Bu bulgular ışığında madde yazarlığı eğitimi sonrası geliştirilen maddelerle oluşturulan testin yapı geçerliği matematik alanında, eğitim öncesine göre daha iyi düzeyde iken Türkçe alanında ise benzer düzeydedir.

Testlerin yapı geçerlikleri dikkate alındığında, matematik alanında ölçülen yapıyı daha iyi ortaya koyabilecek ölçme aracının geliştirilmesinde madde yazarlığı eğitiminin etkili olduğu ve fakat Türkçe alanında aynı etkinin görülemediği söylenebilir.

4.3. Yaratıcılık ve Madde Yazarlığı Eğitiminin Madde Psikometrik Özelliklerini Yordama Gücüne Yönelik Bulgular

4.3.1. Madde Güçlüğü

Matematik

Deney ve kontrol grubundaki (Form III, IV) matematik öğretmenlerinin yaratıcılık ve madde yazarlığı eğitim durumlarının, geliştirdikleri maddelerin madde güçlüklerini yordamasında anlamlı katkı sağlayanların ve bu değişkenlerin her birinin madde güçlüklerini yordamasında açıklanan toplam varyansa katkısı, standart çoklu regresyon analizi ile SPSS 24.0 programı kullanılarak belirlenmiş ve Tablo 4.36'da sonuçları sunulmuştur.

Tablo 4.36: Matematik Maddelerinin Güçlüğü'nün Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları

Değişken	B	Standart Hata B	β	t	p	İkili r	Kısmi R
(Sabit)	0,754	0,110		6,831	0,000	-	-
Eğitim	-0,120	0,031	-0,529	-3,889	0,000	-0,531	-0,539
Yaratıcılık	0,000	0,000	-0,182	-1,338	0,189	-0,187	-0,215
R=0,561	R²=0,315						
F_(2,37)=8,501	p=0,001						

MADDE GÜÇLÜĞÜ=0,754-0,120 EĞİTİM

Tablo 4.36 incelendiğinde, eğitim ile madde güçlüğü arasında negatif ve orta düzeyde bir ilişkinin ($r=-0,53$) olduğu, yaratıcılık değişkeni kontrol edildiğinde korelasyonun $r=-0,54$ olarak hesaplandığı görülmektedir. Yaratıcılık ile madde güçlüğü arasında negatif ve düşük düzeyde ($r=-0,19$) bir ilişki vardır. Eğitim değişkeni kontrol edildiğinde ise korelasyonun $r=-0,22$ olarak hesaplandığı görülmektedir.

Eğitim ve yaratıcılık değişkenleri birlikte, madde güçlüğü ile orta düzeyde ve anlamlı bir ilişki vermektedir ($R=0,561$, $R^2=0,315$, $p<0,01$). Söz konusu iki değişken birlikte, madde güçlüğündeki toplam varyansın yaklaşık %32'ni açıklamaktadır.

Standardize edilmiş regresyon katsayısına (β) göre, yordayıcı değişkenlerin madde güçlüğü üzerindeki göreceli önem sırası eğitim ve yaratıcılıktır. Regresyon katsayılarının anlamlılığına ilişkin t testi sonuçları incelendiğinde ise, sadece eğitim

değişkeninin madde güçlüğü üzerinde önemli (anamlı) bir yordayıcı olduğu görülmektedir ($p < 0,01$). Yaratıcılık önemli bir etkiye sahip değildir.

Eğitim ve yaratıcılık değişkenlerinin madde güçlüğündeki varyansı açıklama yüzdelerini belirlemek amacıyla yapılan aşamalı çoklu regresyon analizi sonucu eğitimin madde güçlüğündeki varyansın %28'ni, yaratıcılığın ise %4'nü açıkladığı görülmüştür.

Türkçe

Deney ve kontrol grubundaki Türkçe öğretmenlerinin yaratıcılık ve madde yazarlığı eğitim durumlarının, geliştirdikleri maddelerin madde güçlüklerini yordamasında anlamlı katkı sağlayanların ve bu değişkenlerin her birinin madde güçlüklerini yordamasında açıklanan toplam varyansa katkısı, standart çoklu regresyon analizi ile belirlenmiş ve Tablo 4.37'de sonuçları gösterilmiştir.

Tablo 4.37: Türkçe Maddelerinin Güçlüğüne Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları

<i>Değişken</i>	<i>B</i>	<i>Standart Hata B</i>	β	<i>t</i>	<i>p</i>	<i>İkili r</i>	<i>Kısmi R</i>
(Sabit)	0,575	0,188		3,051	0,004	-	-
Eğitim	-0,158	0,051	-0,462	-3,126	0,003	-0,468	-0,457
Yaratıcılık	0,000	0,000	-0,033	0,226	0,822	0,118	0,037
$R=0,469$	$R^2=0,220$						
$F_{(2,37)}=5,215$	$p=0,010$						

MADDE GÜÇLÜĞÜ=0,575-0,158 EĞİTİM

Tablo 4.37 incelendiğinde, eğitim ile madde güçlüğü arasında negatif ve orta düzeyde bir ilişkinin ($r=-0,47$) olduğu, yaratıcılık değişkeni kontrol edildiğinde korelasyonun $r=-0,46$ olarak hesaplandığı görülmektedir. Yaratıcılık ile madde güçlüğü arasında pozitif ve düşük düzeyde ($r=0,12$) bir ilişki vardır. Eğitim değişkeni kontrol edildiğinde ise korelasyonun $r=0,04$ olarak hesaplandığı görülmektedir.

Eğitim ve yaratıcılık değişkenleri birlikte, madde güçlüğü ile orta düzeyde ve anlamlı olmayan bir ilişki vermektedir ($R=0,469$, $R^2=0,220$, $p > 0,01$). Söz konusu iki değişken birlikte, madde güçlüğündeki toplam varyansın yaklaşık %22'ni açıklamaktadır.

Standardize edilmiş regresyon katsayısına (β) göre, yordayıcı değişkenlerin madde güçlüğü üzerindeki görece önem sırası eğitim ve yaratıcılıktır. Regresyon katsayılarının anlamlılığına ilişkin t testi sonuçları incelendiğinde ise, sadece eğitim değişkeninin madde güçlüğü üzerinde önemli (anlamlı) bir yordayıcı olduğu görülmektedir ($p < 0,01$). Yaratıcılık önemli bir etkiye sahip değildir.

Eğitim ve yaratıcılık değişkenlerinin madde güçlüğündeki varyansı açıklama yüzdelerini belirlemek amacıyla yapılan aşamalı çoklu regresyon analizi sonucu eğitimin madde güçlüğündeki varyansın %22'ni, yaratıcılığın ise %0,1'ni açıkladığı görülmüştür.

Bu bulgular ışığında matematik ve Türkçe öğretmenlerinin madde yazarlığı eğitim durumları madde güçlüğüne anlamlı bir yordayıcısı iken; yaratıcılıkları anlamlı bir yordayıcısı değildir.

4.3.2. Madde Ayırıcılık Gücü

Matematik

Deney ve kontrol grubundaki matematik öğretmenlerinin yaratıcılık ve madde yazarlığı eğitim durumlarının, geliştirdikleri maddelerin madde ayırıcılık güçlerini yordamasında anlamlı katkı sağlayanların ve bu değişkenlerin her birinin madde güçlüklerini yordamasında açıklanan toplam varyansa katkısı, standart çoklu regresyon analizi ile belirlenmiş ve Tablo 4.38'de sonuçları gösterilmiştir.

Tablo 4.38: Matematik Maddelerinin Ayırıcılık Gücünün Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları

Değişken	B	Standart Hata B	β	t	p	İkili r	Kısmi R
(Sabit)	0,392	0,097	-	4,040	0,000	-	-
Eğitim	0,115	0,027	0,555	4,269	0,000	0,553	0,574
Yaratıcılık	0,000	0,000	-0,262	-2,018	0,051	-0,258	-0,315
$R=0,612$	$R^2=0,374$						
$F_{(2,37)}=11,073$	$p=0,000$						

MADDE GÜÇLÜĞÜ=0,392+0,115 EĞİTİM

Tablo 4.38 incelendiğinde, eğitim ile madde ayırıcılık gücü arasında pozitif ve orta düzeyde bir ilişkinin ($r=0,55$) olduğu, yaratıcılık değişkeni kontrol edildiğinde korelasyonun $r=0,57$ olarak hesaplandığı görülmektedir. Yaratıcılık ile madde

ayırıcılık gücü arasında negatif ve düşük düzeyde ($r=-0,26$) bir ilişki vardır. Eğitim değişkeni kontrol edildiğinde ise korelasyonun $r=-0,32$ olarak hesaplandığı görülmektedir.

Eğitim ve yaratıcılık değişkenleri birlikte, madde ayırıcılık gücü ile orta düzeyde ve anlamlı bir ilişki vermektedir ($R=0,612$, $R^2=0,374$, $p<0,01$). Söz konusu iki değişken birlikte, madde ayırıcılık gücündeki toplam varyansın yaklaşık %37'ni açıklamaktadır.

Standardize edilmiş regresyon katsayısına (β) göre, yordayıcı değişkenlerin madde ayırıcılık gücü üzerindeki göreceli önem sırası eğitim ve yaratıcılıktır. Regresyon katsayılarının anlamlılığına ilişkin t testi sonuçları incelendiğinde ise, sadece eğitim değişkeninin madde ayırıcılık gücü üzerinde önemli (anlamlı) bir yordayıcı olduğu görülmektedir ($p<0,01$). Yaratıcılık önemli bir etkiye sahip değildir.

Eğitim ve yaratıcılık değişkenlerinin madde ayırıcılık gücündeki varyansı açıklama yüzdelerini belirlemek amacıyla yapılan aşamalı çoklu regresyon analizi sonucu eğitimin madde ayırıcılık gücündeki varyansın %30'nu, yaratıcılığın ise %7'ni açıkladığı görülmüştür.

Türkçe

Deney ve kontrol grubundaki Türkçe öğretmenlerinin yaratıcılık indeksleri ve madde yazarlığı eğitim durumlarına göre geliştirdikleri maddelerin ayırıcılık güçlerini yordamasına ilişkin çoklu regresyon analizi sonuçları Tablo 4.39'da verilmiştir.

Tablo 4.39: Türkçe Maddelerinin Ayırıcılık Gücünün Yordanmasına İlişkin Standart Çoklu Regresyon Analizi Sonuçları

<i>Değişken</i>	<i>B</i>	<i>Standart Hata B</i>	β	<i>t</i>	<i>p</i>	<i>İkili r</i>	<i>Kısmi R</i>
(Sabit)	0,187	0,098	-	1,903	0,065	-	-
Eğitim	0,183	0,026	0,750	6,920	0,000	0,760	0,751
Yaratıcılık	0,000	0,000	-0,050	-0,459	0,649	-0,187	-0,075
$R=0,761$	$R^2=0,579$						
$F_{(2,37)}=25,480$	$p=0,000$						

MADDE GÜÇLÜĞÜ=0,187+0,183 EĞİTİM

Tablo 4.39 incelendiğinde, eğitim ile madde ayırıcılık gücü arasında pozitif ve yüksek düzeyde bir ilişkinin ($r=0,76$) olduğu, yaratıcılık değişkeni kontrol edildiğinde korelasyonun $r=0,75$ olarak hesaplandığı görülmektedir. Yaratıcılık ile madde ayırıcılık gücü arasında negatif ve düşük düzeyde ($r=-0,19$) bir ilişki vardır. Eğitim değişkeni kontrol edildiğinde ise korelasyonun $r=-0,08$ olarak hesaplandığı görülmektedir.

Eğitim ve yaratıcılık değişkenleri birlikte, madde ayırıcılık gücü ile yüksek düzeyde ve anlamlı bir ilişki vermektedir ($R=0,761$, $R^2=0,579$, $p<0,01$). Söz konusu iki değişken birlikte, madde ayırıcılık gücündeki toplam varyansın yaklaşık %58'ni açıklamaktadır. Standardize edilmiş regresyon katsayısına (β) göre, yordayıcı değişkenlerin madde ayırıcılık gücü üzerindeki görece önem sırası eğitim ve yaratıcılıktır. Regresyon katsayılarının anlamlılığına ilişkin t testi sonuçları incelendiğinde ise, sadece eğitim değişkeninin madde güçlüğü üzerinde önemli (anlamlı) bir yordayıcı olduğu görülmektedir ($p<0,01$). Yaratıcılık önemli bir etkiye sahip değildir. Eğitim ve yaratıcılık değişkenlerinin madde ayırıcılık gücündeki varyansı açıklama yüzdelerini belirlemek amacıyla yapılan aşamalı çoklu regresyon analizi sonucu eğitimin madde ayırıcılık gücündeki varyansın %57'ni, yaratıcılığın ise %1'ni açıkladığı görülmüştür.

Bu bulgular ışığında matematik ve Türkçe öğretmenlerinin madde yazarlığı eğitim durumları madde ayırıcılık gücünün anlamlı bir yordayıcısı iken; yaratıcılıkları anlamlı bir yordayıcısı değildir.

Madde güçlük indekslerinin her iki alanda da madde yazarlığı eğitiminin %22 ilâ %28 varyans açıklama yüzdesi ile anlamlı yordayıcı olduğu, yaratıcılığın ise anlamlı bir yordayıcı olmadığı bulgusuna ulaşılmıştır. Madde yazarlığı eğitimi, madde ayırıcılık gücü indekslerinin her iki alanda da %30 ilâ %57 varyans açıklama yüzdesi ile anlamlı yordayıcısı olduğu, yaratıcılığın ise anlamlı bir yordayıcı olmadığı bulgusuna ulaşılmıştır. Madde psikometrik özelliklerinde açıklanamayan kısımların, kendilerinden madde ve test puanı elde edilen öğrenci grubunun özelliklerden kaynaklanabileceği gibi alana ve dile hâkimiyet, maddeyi yanıtlayanların eğitimsel ve düşünsel özelliklerini tanıma, madde yazarlığında tecrübeli olma, vb. madde yazarında bulunması gereken başka özelliklerden de kaynaklanmış olabilir.

5. SONUÇ ve ÖNERİLER

Bu bölümde araştırmının bulgu ve yorumlarına dayalı olarak ulaşılan sonuçların özetine ve bu sonuçlardan yola çıkarak geliştirilen önerilere yer verilmiştir.

5.1. Sonuçlar

Bu araştırmada madde yazarlarının yaratıcılık düzeyleri ile madde yazarlığı eğitimi olup olmamalarının, geliştirdikleri maddeler ve bu maddelerle oluşturulan testlerin psikometrik özelliklerine ne derecede etki ettiğinin belirlenmesi amaçlanmıştır. Bu amaçla matematik ve Türkçe öğretmenlerinin yaratıcılık düzeyleri tespit edilmiş, müteakiben aynı öğretmenlerden seçkisiz bir gruba ise madde yazarlığı eğitimi verilmiştir. Madde ve test psikometrik özelliklerinin önce geliştiricilerinin yaratıcılıkları ile olan ilişkisine dair analizler yapılmış sonra madde yazarlığı eğitiminden etkilenip etkilenmedikleri sınanmıştır. Araştırma ile elde edilen bulguların özeti Tablo 5.1'de gösterilmiştir.

Tablo 5.1: Araştırma Bulguları Özeti

<i>Madde ve Test Psikometrik Özellikleri</i>	<i>1.Alt Problem</i>		<i>2.Alt Problem</i>		<i>3.Alt Problem</i>			
	<i>Yaratıcılık</i>		<i>Madde Yazarlığı Eğitimi</i>		<i>Yaratıcılık</i>		<i>Madde Yazarlığı Eğitimi</i>	
	<i>Matematik</i>	<i>Türkçe</i>	<i>Matematik</i>	<i>Türkçe</i>	<i>Matematik</i>	<i>Türkçe</i>	<i>Matematik</i>	<i>Türkçe</i>
Madde Güçlüğü	-	-	+	+	-	-	+	+
Madde Ayırıcılık Gücü	-	-	+	-	-	-	+	+
Madde Çeldiriciliği	+	+	+	+				
İç Tutarlık Güvenirliği	-	-	+	-				
Varyans	-	-	+	+				
Aritmetik Ortalama	+	+	+	+				
Kapsam Geçerliliği	-	-	-	-				
Uyum Geçerliliği	-	-	+	+				
Yapı Geçerliliği	-	-	+	-				

Tablo 5.1 incelendiğinde, yaratıcılığın, madde ve test psikometrik özellikleri ile çeldiricilik dışında ilişkili olmadığı anlaşılmaktadır. Fakat kısa bir sürede uygulanan madde yazarlığı eğitiminin ise madde ve test psikometrik özelliklerine anlamlı katkı sağladığı görülmüştür. Madde yazarlığı eğitiminin farklı alanlarda ve farklı madde/test psikometrik özelliklerinde farklı etkilere sebebiyet verdiği ancak genel itibari ile ölçme araçlarının yapısal özelliklerine olumlu katkıda bulunduğu söylenebilir. Bu durum yaratıcılık için ise geçerli görülmemektedir.

Araştırma bulguları dikkate alındığında, bir teste madde geliştirecek madde yazarı seçiminde yaratıcılık gibi bir özellik sahibi bireyleri aramaktansa; üzerine madde geliştirilecek alana hakim, cevaplayıcı grubun eğitimsel özelliklerini bilen ve madde yazma konusunda tecrübeli uzmanların tespit edilip, bu uzmanlara madde yazarlığı eğitiminin verilmesinin uygun olacağı söylenebilir. Bu eğitim, madde psikometrik özelliklerindeki açıklanan varyanslar dikkate alındığında bir yere kadar etkili görünmektedir. Madde psikometrik özelliklerinde açıklanamayan kısımlar ise madde yazarının sahip olduğu diğer özelliklere bir işarettir.

Sonuç olarak hedef birey/gruptaki ölçülmek istenen özelliğin durumunu en hatasız şekilde ortaya koyan ölçme araçlarının geliştirilebilmesi için maddeyi geliştiren madde yazarının sahip olması gereken özellikleri ve bu özelliklerin ağırlığını açıklamaya yönelik ileri araştırmalara ihtiyacın devam edeceği söylenebilir. Bu özelliklerin belirlenmesi, istendik düzeyde ölçme aracı geliştirebilecek uzmanların seçimi ve bu uzmanlara verilecek eğitimin kapsamını da ortaya koyabilecektir.

5.2. Öneriler

Bu araştırmada elde edilen bulgular doğrultusunda aşağıdaki önerilerde bulunulabilir.

5.2.1. Araştırmaya Dönük Öneriler

1. Bu araştırmada madde yazarının yaratıcılığı ve madde yazarlığı konusunda eğitimi olması değişkenlerinin geliştirdikleri maddelerin psikometrik özelliklerine etkisi incelenmiştir. Madde geliştirmeye etkisi olduğu öne sürülen alana ve dile hâkimiyet, hedef grubu tanıma, madde yazarlığında tecrübeli olma ve madde yazarında bulunması gereken diğer özelliklerin kullanıldığı maddelerin ve/veya testin psikometrik özelliklerini açıklamaya dönük modellemeli bir araştırma yapılabilir.

2. Arařtırmada maddelerin psikometrik zellikleri 203 ve 206'řarlık gruplardan elde edilen veriler zerinden hesaplanmıřtır. Daha byk rneklem grubuyla alıřılarak yaratıcılıđın madde ve test psikometrik zellikleriyle olan iliřkisini yoklayabilmek iin, lkemizde uygulanan geniř lekli ALES, YDS, KPSS, YGS, LYS gibi sınavların hesaplanmış istatistikleri ile bu sınavların madde geliřtiricilerin yaratıcılıkları llerek, bulguları daha genellenebilir arařtırmalar yapılabilir.

3. Arařtırmada matematik ve Trke alanlarındaki maddeler zerinde inceleme yapılmıřtır. Yapılan bu arařtırmanın benzeri alıřmaların diđer alanlarda yapılmasının uygun olacađı deđerlendirilmektedir.

4. Madde yazarlarının yaratıcılıkları TYDT ile llmřtr. Yaratıcılıđın, bařka aralarla llerek benzer bir arařtırmanın yapılması nerilebilir.

5. Bu arařtırmada madde yazarlarına eđitim verilmiř fakat bu eđitimin ne derecede madde yazarınca edinildiđi llmemiřtir. Madde yazarının bu konudaki eđitim dzeyinin llerek geliřtirdiđi maddelerin psikometrik zellikleri ile iliřkisini ortaya koyan arařtırmaların yapılabileceđi deđerlendirilmektedir.

6. Arařtırmada madde eldiricilerinin ne derecede dengeli iřler olduđunun bir istatistiđi olan madde eldiricilik indeksi hesaplanarak kullanılmıřtır. SM'nin glđ ve ayırıcılık gc ile iliřkili olması ve maddenin lme gcn daha ayrıntılı ortaya koymadaki katkısı nedeniyle mteakip madde analizi yapılan arařtırmalarda hesaplanarak yorumlanması nerilebilir.

7. Bu arařtırmada alan uzmanlarına madde yazarlıđı eđitimi verilerek madde ve test psikometrisine etkisi arařtırılmıřtır. lme uzmanlarına belirli bir alan eđitimi verilerek, bu alanda geliřtirdikleri madde ve testlerin psikometrik zelliklerini inceleyen bir arařtırmanın yapılması nerilebilir.

8. SM'lerle yapılan bu arařtırmanın farklı madde tipleriyle yapılmasının uygun olacađı deđerlendirilmektedir.

9. KTK esaslarınca yapılan bu alıřmanın bir benzeri, madde tepki kuramı erevesinde yapılabilir.

5.2.2. Uygulamaya Dönük Öneriler

1. Bu araştırma ile madde yazarlığı eğitiminin madde psikometrik özelliklerine olumlu etkide bulunduğu bulgusuna ulaşılmıştır. Bu nedenle, sonucu bireylerin yaşantılarını önemli derecede etkileyen testlere madde geliştirmek üzere seçilen madde yazarlarının, madde yazarlığı eğitimine tâbi tutulmasının uygun olacağı değerlendirilmektedir.

2. Hâlihazırda ilk ve orta öğretimde görev yapan öğretmenlere verilen ölçme ve değerlendirme eğitiminin detaylandırılarak madde yazarlığı eğitiminin ağırlıklı verilmesinin uygun olacağı değerlendirilmektedir.

3. Eğitimde Ölçme ve Değerlendirme Ana Bilim Dalı Programları içerisine Bilişsel/Duyuşsal Özelliklerin Ölçülmesi ve Test/Ölçek Geliştirme derslerinden ayrı olarak madde yazma/geliştirme vb. derslerinin ilave edilmesinin uygun olacağı değerlendirilmektedir.

KAYNAKÇA

- Abdulghani, H. M., Ahmad, F., Irshad, M., Khalil, M. S., Al-Shaikh, G. K., Syed, S., & Haque, S. (2015). Faculty development programs improve the quality of multiple choice questions items' writing. *Scientific reports*, 5, doi: 10.1038/srep09556
- Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement: Issues and Practice*, 19(3), 16-26. doi: 10.1111/j.1745-3992.2000.tb00034.x
- ACRP (The Association of Clinical Research Professionals) (2017). *Academy item writers*. <https://www.acrpnet.org/about/the-academy/academy-item-writers/>, Eriřim tarihi: 03 Şubat 2017.]
- AERA, APA, NCME. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association. ISBN:978-0-935302-35-6
- Ahlbom, A. (2017). Significance testing: Why does it prevail?. *European Journal of Epidemiology*, 1-2, doi: 10.1007/s10654-016-0220-0
- Akkuş, O. ve Baykul, Y. (2000). Çoktan seçmeli test maddelerini puanlamada, seçenekleri farklı biçimlerde ağırlıklandırmanın madde ve test istatistiklerine olan etkisinin incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* 20, 9-15
- Alpar, R. (2014). *Spor, sağlık ve eğitim bilimlerinden örneklerle uygulamalı istatistik ve geçerlik-güvenirlik-SPSS'de çözümleme adımları ile birlikte (3.Baskı)*. Ankara: Detay Yayıncılık. ISBN: 978-6-055-68187-6
- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Allyn & Bacon. ISBN: 978-0801319037
- Angoff, W. (1953). Test reliability and effective test length. *Psychometrika*, 18, 1-14, doi: 10.1007/BF02289023
- Anıl, D. ve Acar, M. (2008). Sınıf öğretmenlerinin ölçme değerlendirme sürecinde karşılaştıkları sorunlara ilişkin görüşleri. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 5(2), 44-61
- Anıl, D. ve Güzeller, C. O. (2011). Seviye belirleme sınavı fen ve teknoloji alt testi ile diğer alt testler arasındaki ilişkinin yol analizi ile incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 11(1), 1-10
- Anıl, D., Güzeller, C. O., Çokluk, Ö. ve Şekercioğlu, G. (2010). Level determination exam (SBS-2008) the determination of the validity and reliability of 7th grade mathematics sub-test. *Procedia - Social and Behavioral Sciences*, 2(2), 5292-5298. doi: 10.1016/j.sbspro.2010.03.863
- Aranguren, M. (2014). Validez de constructo del Test de Pensamiento Creativo de Torrance en una muestra de jóvenes argentines. *Anuario de Psicología/The UB Journal of Psychology*, 44, 55-70

- Aslan, E. (2001). Torrance yaratıcı düşünce testinin Türkçe versiyonu. *M.Ü. Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 14, 19-40
- Aşık, İ. (2009). *Matematik öğretmenlerinin ölçme değerlendirme araçlarını kullanabilme düzeyleri ve yaklaşımları*. (Yayımlanmamış Yüksek Lisans Tezi). İstanbul: Marmara Üniversitesi EBE. <http://tez2.yok.gov.tr/>
- Atılğan, H., Kan, A., ve Doğan, N. (2009). *Eğitimde ölçme ve değerlendirme*. Anı Yayıncılık.ISBN:9786051701288
- Attali, Y., & Bar-Hillel, M. (2003). Guess where: The position of correct answers in multiple-choice test items as a psychometric variable. *Journal of Educational Measurement*, 40, 109-128. doi: 10.1111/j.1745-3984.2003.tb01099.x
- Aydın, E. (2016). *Ortaokul matematik 8.sınıf ders kitabı*. Ankara: Sevgi Yayınları. ISBN:978-975-8270-45-3
- Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79-86. doi: 10.1177/0748175613513808
- Bart, W. M., Hokanson, B., & Can, I. (2017). An Investigation of the Factor Structure of the Torrance Tests of Creative Thinking. *Educational Sciences: Theory & Practice*, 17(2). doi: 10.12738/estp.2017.2.0051
- Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Mathematical and Statistical Psychology*, 3(2), 77-85. doi:10.1111/j.2044-8317.1950.tb00285.x
- Baş, G. ve Beyhan, Ö. (2016). Öğretmenlerin eğitimde ölçme ve değerlendirmeye yönelik özyeterlik algılarının bazı değişkenler açısından incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 7(1).doi: 10.21031/epod.34199
- Baykul, Y. (1999). *İstatistik: Metodlar ve uygulamalar (3.Baskı)*. Ankara: Anı Yayıncılık. ISBN: 975-6956-12-7.
- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması (3. baskı)*. Ankara: Pegem Akademi.ISBN: 978-605-364-088-2
- Birgin, O. ve Gürbüz, R. (2008). Sınıf öğretmeni adaylarının ölçme ve değerlendirme konusundaki bilgi düzeylerinin incelenmesi. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 20, 163-179
- Brennan, R. L. (2001). *Generalizability theory*. Iowa City, IA: ACT Publications.ISBN: 978-1-4419-2938-9
- Buchanan, J. R. (1885). *Manual of psychometry: The dawn of a new civilization* (Second Ed.). Dudley M.Holman, Press of The Boxbury Advocate
- Büyüköztürk, Ş. (2015). *Veri analizi el kitabı (21.Baskı)*. Ankara: Pegem A Yayıncılık.ISBN: 978-975-6802-74-8
- Büyüköztürk,Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., ve Demirel, F. (2016). *Bilimsel araştırma yöntemleri (20. baskı)*. Ankara: Pegem Akademi.ISBN: 978-9944-919-28-9

- Cash, P., Stankovic, T., & Storga, M. (2016). *Experimental design research*. Cham: Springer International Publishing. ISBN: 978-3-319-33779-1
- Cheng, Y. L., Kim, K. H., Hull, & M. F. (2010). Comparisons of creative styles and personality types between American and Taiwanese college student and the relationship between creative potential and personality types. *Psychology of Aesthetics, Creativity, and the Arts*, 4, 103-112. doi: 10.1037/a0017430
- Cizek, G.J. (1994). The effects of altering the position of options in a multiple-choice examination. *Educational and Psychological Measurement*, 54, 8-20. doi: 10.1177/0013164494054001002
- Clapham, M. M. (1998). Structure of figural Forms A and B of the Torrance Tests of Creative Thinking. *Educational & Psychological Measurement*, 58, 275–283. doi: 10.1177/0013164498058002010
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed). Hillsdale, NJ: Erlbaum. ISBN:0-8058-0283-5
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155. doi: 10.1037/0033-2909.112.1.155
- Cohen, R. J., & Swerdlik, M. (2009). *Psychological testing and assessment: An introduction to tests and measurement (7th Ed.)*. USA: The McGraw–Hill Companies. ISBN:978-0-39-011360-3
- Cramond, B., Matthews-Morgan, J., Bandalos, D., & Zuo, L. (2005). A report on the 40-year follow-up of the Torrance Tests of Creative Thinking: Alive and well in the new millennium. *Gifted Child Quarterly*, 49(4), 283-291. doi: 10.1177/001698620504900402
- Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. Cengage Learning: 5191 Natorp Boulevard Mason, Ohio 45040. ISBN:978-0-495-39591-1
- Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 37(2), 99-114
- Davidshofer, K. R., & Murphy, C. O. (2005). *Psychological testing: principles and applications*. Pearson/Prentice-Upper Saddle River, NJ. ISBN:0-13-129383-4
- Dickinson, J. R. (2014). The Uniformity of Distractor Response Distributions in Multiple-Choice Questions. *Developments in Business Simulation and Experiential Learning*, 41
- Dickinson, J. R. (2015). A New Statistic for Item Analysis. In *Revolution in Marketing: Market Driving Changes* (pp. 206-206). Springer International Publishing. doi: 10.1007/978-3-319-11761-4_97
- Doğan, N., ve Tezbaşaran, A. A. (2003). Klasik test kuramı ve örtük özellikler kuramının örneklem bağlamında karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 25(25)
- Donnelly, R., & Fitzmaurice, M. (2005). Designing modules for learning. In G. O'Neill, S. Moore & B. McMullin(eds.) *Emerging issues in the practice of University Learning and Teaching*, Dublin, All Ireland Society for Higher Education (AISHE)

- Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. *Advances in health sciences education, 10*(2), 133-143. doi: 10.1007/s10459-004-4019-5
- Downing, S. M., & Haladyna, T. M. (1997). Test Item Development: Validity Evidence from Quality Assurance Procedures. *Applied Measurement in Education vol 10*(1), 61-82. doi: 10.1207/s15324818ame1001_4
- Downing, S. M., & Haladyna, T. M. (2006). *Handbook of test development*. Lawrence Erlbaum Associates Publishers. ISBN:0-203-87477-3
- DPB (Devlet Personel Başkanlığı) (2017). *2004 - 2012 yılları arasında yapılan genel kpss'lere katılan aday sayıları*. [Çevrim-içi:<http://www.dpb.gov.tr/tr-tr/istatistik/2004-2012-yillari-arasinda-yapilan-genel-kpsslere-katilan-aday-sayilari> Erişim tarihi: 16 Şubat 2017.]
- Dufresne, R. J., Leonard, W. J., & Gerace, W. J. (2002). Marking sense of students' answers to multiple-choice questions. *The Physics Teacher, 40*(3), 174-180. doi: 10.1119/1.1466554
- Ebel, R. L. (1951). Writing the test item. *Educational measurement, 185-245*. American Council on Education: Washington D.C.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurements* (5th Ed.). Prentice-Hall of India. ISBN: 0-87692-700-2
- Eid, M., Gollwitzer, M., Schmitt, M. (2011). *Statistik und Forschungsmethoden Lehrbuch*. Weinheim: Beltz.
- Erdemir, A. Z. (2007). *İlköğretim ikinci kademe öğretmenlerinin ölçme değerlendirme tekniklerini etkin kullanabilme yeterliklerinin araştırılması*. (Yayımlanmamış yüksek lisans tezi). Kahramanmaraş Sütçü İmam Üniversitesi Sosyal Bilimler Enstitüsü, Kahramanmaraş. <http://tez2.yok.gov.tr/>
- Erdoğdu, M. Y. ve Kurt, F. (2012). Öğretmenlerin ölçme ve değerlendirme yeterlik algılarının bazı değişkenler açısından incelenmesi. *Elektronik Eğitim Bilimleri Dergisi, 1*(2), 23-36
- Ergen, Z. G. ve Akyol, A. K. (2012). Anaokuluna devam eden çocukların yaratıcılıklarının incelenmesi. *Kuramsal Eğitimbilim Dergisi, 5*(2), 156-170
- Erkuş, A. (2003). *Psikometri üzerine yazılar*. Ankara: Türk Psikologlar Derneği Yayınları. ISBN: 975-6761-04-0
- Erkuş, A. (2010). Psikometrik terimlerin Türkçe karşılıklarının anlamları ile yapılan işlemlerin uyumsuzluğu. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 1*(2), 72-77
- ETS (Educational Testing Service) (2017). *Course Catalog*. [Çevrim-içi: <https://www.ets.org/Media/Global/pdf/globalInstCatalog.pdf>, Erişim tarihi: 03 Mart 2017.]

- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299. doi: 10.1037/1082-989X.4.3.272
- Fishkin, A. S., Cramond, B., Olszewski-Kubilius, P., & Rejskind, G. (2000). Investigating Creativity in Youth: Research and Methods. *Book Reviews, High Ability Studies, 11*:2, 221-234. doi: 10.1080/13598130020001232
- Fraenkel, J. R., Wallen, N. E., & Hyun, H.H. (2011). *How to design and evaluate research in education* (8th Ed.). New York: McGraw-Hill International Edition. ISBN:978-0-07-809785-0
- Frary, R. B. (1982). A simulation study of reliability and validity of multiple-choice test scores under six response-scoring modes. *Journal of Educational and Behavioral Statistics, 7*(4), 333-351. doi: 10.3102/10769986007004333
- Frey, B. B., Petersen, S., Edwards, L. M., Pedrotti, J. T., & Peyton, V. (2005). Item-writing rules: Collective wisdom. *Teaching and Teacher Education, 21*(4), 357-364. doi: 10.1016/j.tate.2005.01.008
- Fuhrman, M. (1996). Developing good multiple-choice tests and test questions. *Journal of Geoscience Education, 44*(4), 379-384. doi: 10.5408/1089-9995-44.4.379
- Fulkerson, D., Mittelholtz, D. J., & Nichols, P. D. (2009). The psychology of writing items: Improving figural response item writing. In *annual meeting of the American Educational Research Association, San Diego, CA*.
- Fulkerson, D., Nichols, P., & Mittelholtz, D. (2010, May 3). What item writers think when writing items: Toward a theory of item writing expertise. Paper presented at the *Annual Meeting of the American Educational Research Association, Denver, CO*.
- Furr, R. M. & Bacharach, V. R. (2013). *Psychometrics: an introduction*. Sage Publications, Inc. 2455 Teller Road Thousand Oaks, California 91320. ISBN:978-1-4522-5680-1
- Galton, F. R. S. (1879). Psychometric experiments. *Brain, 2*, 149–162
- Geisinger, K. F., Bracken, B. A., Carlson, J. F., Hansen, J. I. C., Kuncel, N. R., Reise, S. P., & Rodriguez, M. C. (2013). *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology*. American Psychological Association. doi: 10.1037/14047-000
- Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Psychological Education and Exercises Science, 5*(1), 13-14. doi: 10.1207/S15327841MPEE0501_2
- Gök, B. ve Erdoğan, T. (2011). Sınıf öğretmeni adaylarının yaratıcı düşünme düzeyleri ve eleştirel düşünme eğilimlerinin incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 44*(2), 29-51
- Gregory, R. J. (2011). *Psychological testing: history, principles, and applications* (6th ed.). Pearson. ISBN: 9780205782147
- Guilford, J. P. (1956). Structure of intellect. *Psychological Review, 53*, 267–293. doi: 10.1037/h0040755

- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley & Sons Inc.
- Güleç, S. ve Alkış, S. (2003). İlköğretim Birinci Kademe Öğrencilerinin Derslerdeki Başarı Düzeylerinin Birbiri ile İlişkisi. *İlköğretim Online*, 2(2)
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3th Ed.). Routledge. ISBN:0-8058-4661-1
- Haladyna, T. M., & Downing, S. M., (1989a). A taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education* vol 2(1), 37-50. doi: 10.1207/s15324818ame0201_3
- Haladyna, T. M., & Downing, S. M. (1989b). Validity of a taxonomy of multiple-choice item-writing rules. *Applied Measurement in Education* vol 2(1), 51-78. doi: 10.1207/s15324818ame0201_4
- Haladyna, T. M., & Downing, S. M. (1993). How many options is enough for a multiple-choice test item?. *Educational and Psychological Measurement*, 53(4), 999-1010. doi: 10.1177/0013164493053004013
- Haladyna, T. M., Downing, S. M., & Rodriguez M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Appl Meas Educ.* 15(3):309–334. doi: 10.1207/S15324818AME1503_5
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge. ISBN:978-0-415-87604-9
- Haladyna, T. M., & Shindoll, R. R. (1989). Item shells: A method for writing effective multiple-choice test items. *Evaluation & the Health Professions*, 12(1), 97-106. doi: 10.1177/016327878901200106
- Hambleton, R. K., & Eignor, D. R. (1979). A Practitioner's Guide to Criterion-Referenced Test Development, Validation, and Test Score Usage. *Laboratory of Psychometric and Evaluation Research Report No. 70*. [Çevrim-içi: <http://files.eric.ed.gov/fulltext/ED249269.pdf>, Erişim tarihi: 19 Kasım 2016.]
- Hancock, G. R. (1994). Cognitive complexity and the comparability of multiple-choice and constructed-response test formats. *The Journal of experimental education*, 62(2), 143-157. doi: 10.1080/00220973.1994.9943836
- Hansen, J. D., & Dexter, L. (1997). Quality multiple-choice test questions: Item-writing guidelines and an analysis of auditing testbanks. *Journal of Education for Business*, 73(2), 94-97. doi: 10.1080/08832329709601623
- Hawkey, R., & Green, A. (2012). *An empirical investigation of the process of writing Academic Reading test items for the international English Language Testing System*. In *Studies in Language Testing*. Cambridge University Press
- Heausler, N. L., & Thompson, B. (1988). Structure of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement*, 48, 463–468. doi: 10.1177/0013164488482021
- Hingorjo, M. R., & Jaleel, F. (2012). Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. *JPMA-Journal of the Pakistan Medical Association*, 62(2), 142

- Hocevar, D. (1979). Ideational fluency as a confounding factor in the measurement of originality. *Journal of Educational Psychology*, 71, 191–196. doi: 10.1037/0022-0663.71.2.191
- Hocevar, D., & Michael, W. (1979). The effects of scoring formulas on the discriminant validity of tests of divergent thinking. *Educational and Psychological Measurement*, 39, 917–921. doi: 10.1177/001316447903900427
- Hohensinn, C., & Baghaei, P. (2017). Does the Position of Response Options in Multiple-Choice Tests Matter?. *Psicológica*, 38, 93-109
- Horst, P. (1933). The difficulty of a multiple choice test item. *Journal of educational psychology*, Vol. 24, Issue 3, 229-232. doi: 10.1037/h0073588
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55. doi: 10.1080/10705519909540118
- Ingham, K. A. T. E. (2008). The Cambridge ESOL approach to item writer training: the case of ICFE Listening. *Research Notes*, 32, 5-9. [Çevrim-içi: <http://www.cambridgeenglish.org/images/23151-research-notes-32.pdf>, Erişim tarihi: 11 Kasım 2016.]
- Iramaneerat, C. (2012). The impact of item writer training on item statistics of multiple-choice items for medical student examination. *Siriraj Med J*, 64(6), 178-182
- Irvine, S. H., & Kyllonen, P. C. (2002). *Item Generation for Test Development*. Lawrence Erlbaum Associates, New York. ISBN:0-8058-3441-9
- Işık, İ. (2014). Yokluk hipotezi anlamlılık testi ve etki büyüklüğü tartışmalarının psikoloji araştırmalarına yansımaları. *Eleştirel Psikoloji Bülteni*, 5, 55-80
- ITEA (2012). *Item Writer Workshops*. [Çevrim-içi: <http://www.itea.org/professional-certification/22-advance/236-item-writing-workshops.html#wave>, Erişim tarihi: 20 Kasım 2016.]
- Jozefowicz, R. F., Koeppen, B. M., Case, S., Galbraith, R., Swanson, D., & Glew, R. (2002). The quality of in-house medical school examinations. *Academic Medicine*, 77(2), 156–161
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. doi: 10.1007/BF02291575
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38(4), 319-342. doi: 10.1111/j.1745-3984.2001.tb01130.x
- Kaplan, R., & Saccuzzo, D. (2012). *Psychological testing: Principles, applications, and issues*. (8th Ed.) Nelson Education. ISBN:978-1-133-49201-6
- Karataş, S. ve Özcan, S. (2010).Yaratıcı düşünme etkinliklerinin öğrencilerin yaratıcı düşüncelerine ve proje geliştirmelerine etkisi. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi*, 11(1), 225-243
- Kılıç, A. F. ve Kelecioğlu, H. (2016). TEOG Ortak ve Mazeret Sınavındaki Türkçe ve Matematik Alt Testlerinin Psikometrik Özelliklerinin Karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 7(1).doi: 10.21031/epod.14532

- Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance tests of creative thinking (TTCT). *Creativity Research Journal*, 18, 3-14.doi: 10.1207/s15326934crj1801_2
- Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *Journal of Creative Behavior*, 42, 106-130.doi: 10.1002/j.2162-6057.2008.tb01290.x
- Klein, S.P. (1998). Standards for teacher tests. *Journal of Personnel Evaluation in Education*, 12(2), 123-138.doi: 10.1023/A:1008029010483
- Kline, P. (2015a). *A handbook of test construction (psychology revivals): introduction to psychometric design*. Routledge.ISBN:978-1-138-90511-5
- Kline, R. B. (2015b). *Principles and practise of structural equating modeling*. (4th Ed). Guilford Press.New York.ISBN:978-1-4625-2334-4
- Knowles, S. L., & Welch, C. A. (1992). A meta-analytic review of item discrimination and difficulty in multiple-choice items using none-of-the-above. *Educational and Psychological Measurement*, 52, 571-577.doi: 10.1177/0013164492052003006
- Koon, S., & Kamata, A. (2013). An applied examination of methods for detecting differential distractor functioning. *International Journal of Quantitative Research in Education*, 1(4), 364-382.doi: 10.1504/IJQRE.2013.058306
- Krumm, G., Lemos, V., & Filippetti, V. A. (2014). Factor structure of the Torrance tests of creative thinking figural Form B in Spanish-speaking children: Measurement invariance across gender. *Creativity Research Journal*, 26, 72–81.doi: 10.1002/jocb.76
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151-160.doi: 10.1007/BF02288391
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.doi: 10.2307/2529310
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*,28(4), 563–575.doi: 10.1111/j.1744-6570.1975.tb01393.x
- Lee, Y. J., Kim, M., Jin, Q., Yoon, H. G., & Matsubara, K. (2017). Primary Science Curricula: Past and Present Realities. In *East-Asian Primary Science Curricula* (pp. 1-10). Springer Singapore
- Leikin, R., & Sriraman, B. (Eds.). (2016). *Creativity and Giftedness: Interdisciplinary perspectives from mathematics and beyond*. Springer.ISBN:978-3-319-38838-0
- Levy, R., & Mislevy, R. J. (2016). *Bayesian psychometric modeling*. CRC Press.ISBN: 978-1-4398-8468-3
- Lord, F. M. (1977). Optimal number of choices per item—A comparison of four approaches. *Journal of Educational Measurement*, 14, 33-38.doi:10.1111/j.1745-3984.1977.tb00026.x
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, Mass. : Addison-Wesley

- McMorris, R. F., Boothroyd, R. A., & Pietrangelo, D. J. (1997). Humor in educational testing: A review and discussion. *Applied Measurement in Education*, 10, 269-297. doi: 10.1207/s15324818ame1003_5
- MEB (Millî Eğitim Bakanlığı) (2008). *Öğretmen Yeterlikleri: Öğretmenlik Mesleği Genel ve Özel Alan Yeterlikleri*. Ankara: Devlet Kitapları. [Çevrim-içi: http://otmg.meb.gov.tr/belgeler/ogretmen_yeterlikleri_kitabi/%C3%96%C4%9Fretmen_Yeterlikleri_Kitab%C4%B1_genel_yeterlikler_par%C3%A7a_2.pdf, Erişim tarihi: 01 Ekim 2016.]
- MEB (Millî Eğitim Bakanlığı) (2013). *Ortaokul matematik (5, 6, 7 ve 8. sınıflar) öğretim programı*. Ankara: Talim ve Terbiye Kurulu Başkanlığı. [Çevrim-içi: <http://ttkb.meb.gov.tr/program2.aspx/program2.aspx?islem=1&kno=215>, Erişim tarihi: 11 Ekim 2016.]
- MEB (Millî Eğitim Bakanlığı) (2015a). *Türkçe dersi (1-8. sınıflar) öğretim programı*. Ankara: Talim ve Terbiye Kurulu Başkanlığı. [Çevrim-içi: <http://ttkb.meb.gov.tr/program2.aspx>, Erişim tarihi: 11 Ekim 2016.]
- MEB (Millî Eğitim Bakanlığı) (2015b). *2014-2015 eğitim öğretim yılı I. dönem ortak sınavlar test ve madde istatistikleri*. [Çevrim-içi: <http://odsgm.meb.gov.tr/test/analizler/docs/2014-2015-1-Donem-Ortak-Sinavlar-Genel-Bilgiler.pdf>, Erişim tarihi: 29 Kasım 2016.]
- MEB (Millî Eğitim Bakanlığı) (2016a). *2015-2016 yılı eğitim istatistikleri*. [Çevrim-içi: http://ankara.meb.gov.tr/meb_iys_dosyalar/2016_06/15050140_20152016_egitim_ogretim_yili_istatistikler.rar, Erişim tarihi: 16 Kasım 2016.]
- MEB (Millî Eğitim Bakanlığı) (2016b). *2015-2016 eğitim öğretim yılı I. dönem ortak sınavlar test ve madde istatistikleri*. [Çevrim-içi: https://odsgm.meb.gov.tr/meb_iys_dosyalar/2016_02/08121811_analiz2627kasm2015ortaksnavraporu.pdf, Erişim tarihi: 29 Kasım 2016.]
- Mehrens, W. A., & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology* (4th Ed.). Wadsworth Thomson Learning. ISBN:0-03-030-407-5
- Mullis, I., V., S. & Martin, M., O. (2013). *Timss 2015 item writing guidelines*. Lynch School of Education, Boston College. [Çevrim-içi: https://timssandpirls.bc.edu/publications/timss/T15_item_writing_guidelines.pdf, Erişim tarihi: 20 Aralık 2016.]
- Nathan, A. T., & Jennifer, P. D. (2015). *Item writing and review manual*. [Çevrim-içi: http://www.assess.com/wp-content/uploads/2015/10/Item_Writing_Review_Manual_2015.10.23.pdf, Erişim tarihi: 21 Ekim 2016.]
- NCTM (National Council of Teachers of Mathematics). (2000). *Principles and standards for school mathematics* (Vol. 1). National Council of Teachers of. ISBN: 0-87353-480-8
- Newell, A., Shaw, J. C., & Simon, H. A. (1959). *The processes of creative thinking* (p. 2). Santa Monica, CA: Rand Corporation
- Newstead, S., & Dennis, I. (1994) The reliability of exam marking in psychology: examiners examined, *Psychologist*, 7(5), 216–219

- Nickerson, R. S. (2000). Null hypothesis significance testing: a review of an old and continuing controversy. *Psychological methods*, 5(2), 241. doi: 10.1037/1082-989X.5.2.241
- Oescher, J., & Kirby, P. C. (1990). Assessing teacher-made tests in secondary math and science classrooms. *Annual Meeting of the National Council on Measurement in Education, Boston, MA*. [Çevrim-içi: <http://files.eric.ed.gov/fulltext/ED322169.pdf>, Erişim tarihi: 21 Ekim 2016.]
- Okur, M. (2008). *4. ve 5. sınıf öğretmenlerinin fen ve teknoloji dersinde kullanılan alternatif ölçme ve değerlendirme tekniklerine ilişkin görüşlerinin belirlenmesi* (Yayımlanmamış yüksek lisans tezi, Zonguldak Karaelmas Üniversitesi Sosyal Bilimler Enstitüsü, Zonguldak). <http://tez2.yok.gov.tr/>
- Örs, S. (2010). *İlköğretim 6., 7. ve 8. sınıf Seviye Belirleme Sınavı fen ve teknoloji alt testlerinin faktör yapılarının belirlenmesi*. (Yayımlanmamış Yüksek Lisans Tezi). Ankara Üniversitesi. <http://tez2.yok.gov.tr/>
- Özgüven, İ. E. (2014). *Psikolojik testler*. 12. Basım. Nobel Akademik Yayıncılık: Ankara. ISBN:978-605-133-317-5
- Öztuna, D., Elhan, & A. H., Tüccar, E. (2006). Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. *Turkish Journal of Medical Sciences*, 36(3), 171-176
- Pais, J., Silva, A., Guimarães, B., Povo, A., Coelho, E., Silva-Pereira, F., & Severo, M. (2016). Do item-writing flaws reduce examinations psychometric quality?. *BMC Research Notes*, 9(1), 399. doi: 10.1186/s13104-016-2202-4
- Pate, A., & Caldwell, D. J. (2014). Effects of multiple-choice item-writing guideline utilization on item and student performance. *Currents in Pharmacy Teaching and Learning*, 6(1), 130-134. doi: 10.1016/j.cptl.2013.09.003
- Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Royal Society Proceedings*, 58, 241. doi: 10.1098/rspl.1895.0041
- Plucker, J. A., & Renzulli, J. S. (1999). Psychometric approaches to the study of human creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 35–61). Cambridge, UK: Cambridge University Press. ISBN:978-0-521-57285-9
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4, 207-230. doi: 10.3102/10769986004003207
- Roberts, D. M. (1993). An empirical study on the nature of trick questions. *Journal of Educational Measurement*, 30, 331-344. doi: 10.1111/j.1745-3984.1993.tb00430.x
- Rodriguez, M. C. (1997). *The art & science of item writing: A meta-analysis of multiple-choice item format effects*. American Education Research Association, Chicago, IL (Vol. 5, p. 2)
- Rudowicz, E., Lok, D., & Kitto, J. (1995). Use of the Torrance Tests of Creative Thinking in an exploratory study of creativity in Hong Kong primary school children: A cross-cultural comparison. *International Journal of Psychology*, 30(4), 417-430. doi: 10.1080/00207599508246577

- Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, 24, 66–75. doi: 10.1080/10400419.2012.652929
- Runco, M. A., & Marz, W. (1992). Scoring divergent thinking tests using total ideational output and a creativity index. *Educational and Psychological Measurement*, 52, 213–221. doi: 10.1177/001316449205200126
- Salisbury, K. (2005). *The edge of expertise?: towards an understanding of listening test item writing as professional practice* (Doctoral dissertation, University of London)
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological assessment*, 8(4), 350. doi: 10.1037/1040-3590.8.4.350
- Simpson, R. M. (1922). Creative imagination. *The American Journal of Psychology*, 234-243
- Sözbilir, M. (2010). *Madde analizi ve test geliştirme*. [Çevrim-içi: <https://olcmevedegerlendirme.files.wordpress.com/2010/09/7-madde-analizi-ve-test-gelistirme.pdf>, Erişim tarihi: 27 Mayıs 2017.]
- Spearman, C. D. (1931). Creative mind. *The Journal of Nervous and Mental Disease*, 74(6), 783
- Spence, L. (2003). *Encyclopedia of Occultism and Parapsychology*. Vol 2, 1257. Kessinger Publishing. ISBN:0-8103-9489-8
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103, 677-680
- Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement*, 22, 271–286. doi: 10.1111/j.1745-3984.1985.tb01064.x
- Stuart-Hamilton, I. (2007). *Dictionary of psychological testing, assessment and treatment*. (2nd Ed.) Jessica Kingsley Publishers. ISBN: 978 1 84310 494 0
- Şenel, T. (2008). *Fen ve teknoloji öğretmenleri için alternatif ölçme ve değerlendirme tekniklerine yönelik bir hizmet içi eğitim programının etkililiğinin araştırılması*. (Yayımlanmamış Yüksek Lisans Tezi). Karadeniz Teknik Üniversitesi, Trabzon. <http://tez2.yok.gov.tr/>
- Şenyurt, S., & Özkan, Y. Ö. (2017). Eğitimde Ölçme ve Değerlendirme Alanında Yapılan Yüksek Lisans Tezlerinin Tematik ve Metodolojik Açıdan İncelenmesi. *İlköğretim Online*, 16(2)
- Tabachnick, B. G., & Fidell, L. S. (2015). *Using Multivariate Statistics* (6th ed.). Baloğlu M. (Çev.) Pearson Education, Inc. ISBN:978-605-320-250-9
- Tamir, R (1993). Positive and negative multiple-choice items: How different are they? *Studies in Educational Evaluation*, 19, 311-325. doi: 10.1016/S0191-491X(05)80013-6

- Tan, Ş. (2009a). Misuses of KR-20 and Cronbach's alpha reliability coefficients. *Eğitim ve Bilim*, 34(152), 101
- Tan, Ş. (2009b). Soru sırasının madde güçlüğü ve ayırıcılık gücüne etkisi. *e-Journal of New World Sciences Academy*, 4(2)
- Tarrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse education in practice*, 6(6), 354-363.doi: 10.1016/j.nepr.2006.07.002
- Tarrant, M., & Ware, J. (2008). Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Medical education*, 42(2), 198-206.doi: 10.1111/j.1365-2923.2007.02957.x
- TDK (Türk Dil Kurumu) (2017). *Güncel Türkçe Sözlük*, [Çevrimiçi: http://www.tdk.gov.tr/index.php?option=com_gts&arama=gts&guid=TDK.GTS.58ef36ba11bf67.50769451, Erişim tarihi: 01 Şubat 2017.]
- Tenzin, K., Dorji, T., & Tenzin, T. (2017). Construction of Multiple Choice Questions Before and After An Educational Intervention. *Journal of the Nepal Medical Association*, 56(205)
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological methods*, 16(2), 209.doi: 10.1037/a0023353
- Toğrol, A. Y. (1999). Yaratıcı Düşünme Testi-Çizim Ürünü'nün Türk Örneklemelerinde Kullanımı. *Eğitim ve Bilim*, 23(112)
- Tok, E. ve Sevinç, M. (2012). Düşünme becerileri eğitiminin okul öncesi öğretmen adaylarının yaratıcı düşünme becerilerine etkisi. *Eğitim ve Bilim*, 37(164), 204-222
- Torrance, E. P. (1966). *The Torrance tests of creative thinking-norms-technical manual research edition-verbal tests, forms A and B- figural tests, forms A and B*. Princeton, NJ: Personnel Press
- Torrance, E. P. (1967). The Minnesota studies of creative behavior: National and international extensions. *The Journal of Creative Behavior*, 1(2), 137-154.doi: 10.1002/j.2162-6057.1967.tb00021.x
- Torrance, E. P. (1974). *The Torrance tests of creative thinking-norms-technical manual research edition-verbal tests, forms A and B;- figural tests, forms A and B*. Princeton, NJ: Personnel Press
- Torrance, E. P. (1990). *The Torrance tests of creative thinking norms-technical manual figural (streamlined) forms A &B*. Bensenville, IL: Scholastic Testing Service, Inc.
- Torrance, E. P. (1998). *The Torrance tests of creative thinking norms-technical manual figural (streamlined) forms A &B*. Bensenville, IL: Scholastic Testing Service, Inc.
- Torrance, E. P. (2003). Developed series of creativity tests. *Los Angeles Times*, p.13
- Torrance, E. P. (2008). *The Torrance tests of creative thinking norms-technical manual figural (streamlined) forms A &B*. Bensenville, IL: Scholastic Testing Service, Inc.

- Torrance, E. P., & Ball, O. E. (1984). *The Torrance tests of creative thinking streamlined (revised) manual, figural A and B*. Bensenville, IL: Scholastic Testing Service, Inc.
- Torrance, E. P., & Wu, T. (1981). A comparative longitudinal study of the adult creative achievements of elementary school children identified as highly intelligent and as highly creative. *Creative Child and Adult Quarterly*, 6, 71–76
- TTKB (Talis Terbiye Kurulu Başkanlığı) (2016). *2016-2017 Eğitim Öğretim Yılı Merkezi Sistem Ortak Sınavları*, [Çevrim-içi: <http://ttkb.meb.gov.tr/www/2016-2017-egitim-ve-ogretim-yili-merkezi-sistem-ortak-sinavlari/icerik/281>, Erişim tarihi: 24 Kasım 2016.]
- Turgut, M. F. (1977). *Eğitimde ölçme ve değerlendirme metotları*. Ankara: Nuve Matbaası
- Turgut, M. F. (2011). Ülkemizde Ölçme ve Değerlendirmenin Dünü–Bugünü ve Yarını–Dünü. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 2
- Ukşul, E. (2016). *Türkiye’de Eğitimde Ölçme Ve Değerlendirme Alanında Yapılmış Bilimsel Yayınların Sosyal Ağ Analizi İle Değerlendirilmesi: Bir Bibliyometrik Çalışma*. (Yayımlanmamış Yüksek Lisans Tezi). Akdeniz Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Anabilim Dalı, Antalya. <http://tez2.yok.gov.tr/>
- Urbina, S. (2014). *Essentials of psychological testing*. John Wiley & Sons. ISBN: 978-1-118-70725-8
- Vinay, P., Sunil, K., Praveen, I., & Yuvaraj, B. (2016). Effect of training on formulation of multiple choice questions: a cross-sectional study amongst faculty in the department of biochemistry of a Medical Institution in India. *International Journal of Research in Medical Sciences*, 4(9), 3694-3701. doi: 10.18203/2320-6012.ijrms20162601
- Wainer, H., & Dorans, N. J. (2000). *Computerized adaptive testing: A primer* (2. Edition). London: Lawrence Erlbaum Associates, Inc. ISBN: 9780585384375
- Wesman, A.G. (1971). Writing the test item. In R.L. Thorndike (Ed.) *Educational Measurement (1st ed, pp 99-111)*. Washington, DC: American Council on Education. <https://archive.org/details/educationalmeasu00thor>
- Widaman, K. F. (1993). Common Factor-Analysis versus Principal Component Analysis – Differential Bias in Representing Model Parameters. *Multivariate Behavioral Research*, 28(3), 263-311. doi: 10.1207/s15327906mbr2803_1
- Wilcox, A. R. (1973). Indices of qualitative variation and political measurement. *Western Political Quarterly*, 26(2), 325-343. doi: 10.2307/446831
- Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education*, 42, 37–42. doi: 10.1177/002248719104200106
- Wood, R. (1977). Multiple choice: A state of the art report. *Evaluation in Education. International Progress*, 1(3), 191-280. doi: 10.1016/0145-9228(77)90007-3
- Yapıcı, M. ve Demirdelen. C. (2007). İlköğretim 4. Sınıf Sosyal Bilgiler Öğretim Programına İlişkin Öğretmen Görüşleri. *İlköğretim Online*, 6 (2): 204–212.

Yıldız, S. (2011). *İlköğretim Programlarındaki Ölçme ve Değerlendirme Yaklaşımları ile İlgili Branş Öğretmenlerinin Görüşleri*. (Yayımlanmamış Yüksek Lisans Tezi). İstanbul: Marmara Üniversitesi EBE. <http://tez2.yok.gov.tr/>

Yüksel, S. (2007). Bilişsel alanın sınıflamasında (taksonomi) yeni gelişmeler ve sınıflamalar. *Türk Eğitim Bilimleri Dergisi*, 5 (3), 479-509

EKLER DİZİNİ

EK 1. ETİK KOMİSYONU ONAY BİLDİRİMİ



T.C.
HACETTEPE ÜNİVERSİTESİ
Rektörlük

Sayı : 35853172/ 433 - 3739

12 Aralık 2016

EĞİTİM BİLİMLERİ ENSTİTÜ MÜDÜRLÜĞÜNE

İlgi: 14.11.2016 tarih ve 2619 sayılı yazınız.

Enstitünüz Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı tezli yüksek lisans programı öğrencilerinden **Kenan Burak YÜKSEL**'in **Doç. Dr. Nuri DOĞAN** danışmanlığında yürüttüğü "**Yaratıcılık ve Madde Yazarlığı Eğitiminin Ölçme Aracının Psikometrik Özelliklerine Etkisi**" başlıklı tez çalışması, Üniversitemiz Senatosu Etik Komisyonunun **02 Aralık 2016** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini rica ederim.

Prof. Dr. Rahime M. NOHUTCU
Rektör a.
Rektör Yardımcısı

Hacettepe Üniversitesi Rektörlük 06100 Sıhhiye-Ankara
Telefon: 0 (312) 305 3001 - 3002 • Faks: 0 (312) 311 9992
E-posta: yazimd@hacettepe.edu.tr • www.hacettepe.edu.tr

Ayrıntılı Bilgi için:
Yazı İşleri Müdürlüğü
0 (312) 305 1008

EK 2. ARAŞTIRMA İZİN BELGESİ



T.C.
ANKARA VALİLİĞİ
Milli Eğitim Müdürlüğü

Sayı : 14588481-605.99-E.12957741
Konu : Araştırma İzni

16.11.2016

HACETTEPE ÜNİVERSİTESİNE
(Rektörlük)

İlgi: a) MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 2012/13 nolu Genelgesi.
b) 03/11/2016 tarihli ve 3304 sayılı yazınız.

Enstitünüz Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı yüksek lisans öğrencisi Kenan Burak YÜKSEL'in "**Yaratıcılık ve Madde Yazarlığı Eğitiminin Ölçme Aracının Psikometrik Özelliklerine Etkisi**" konulu tez kapsamında uygulama talebi Müdürlüğümüzce uygun görülmüş ve uygulamanın yapılacağı İlçe Milli Eğitim Müdürlüğüne bilgi verilmiştir.

Görüşme formunun (18 sayfa) araştırmacı tarafından uygulama yapılacak sayıda çoğaltılması ve çalışmanın bitiminde bir örneğinin (cd ortamında) Müdürlüğümüz Strateji Geliştirme (1) Şubesine gönderilmesini rica ederim.

Vefa BARDAKCI
Vali a.
Milli Eğitim Müdürü

EK 3. ORJİNALLİK RAPORU



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS/DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU

HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLER ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANA BİLİM DALI BAŞKANLIĞI'NA

Tarih: 19/06/2017

Tez Başlığı: Yaratıcılığın ve Madde Yazarlığı Eğitiminin Ölçme Aracının Psikometrik Özelliklerine Etkisi

Yukarıda başlığı verilen tez çalışmamın tamamı (kapak sayfası, özetler, ana bölümler, kaynakça) aşağıdaki filtreler kullanılarak **Turnitin** adlı intihal programı aracılığı ile kontrol edilmiştir. Kontrol sonucunda aşağıdaki veriler elde edilmiştir.


Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Endeksi	Gönderim Numarası
19/06/2017	157	256495	19/06/2017	%4	826098571

Uygulanan filtreler:

- 1- Kaynakça hariç
- 2- Alıntılar dâhil
- 3- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.


19/06/2017

Adı Soyadı: Kenan Burak YÜKSEL

Öğrenci No: N14219758

Anabilim Dalı: Eğitim Bilimleri Anabilim Dalı

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü: Y.Lisans Doktora Bütünleşik Dr.

DANIŞMAN ONAYI

UYGUNDUR.

Prof.Dr.Nuri DOĞAN





HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
THESIS ORIGINALITY REPORT

HACETTEPE UNIVERSITY
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES
TO THE DEPARTMENT OF EDUCATIONAL SCIENCES

Date:19/06/2017

Thesis Title: The Effect of Creativity and Education in Item Writing on Measurement Tool's Psychometric Properties

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

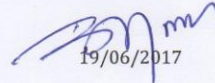
Time Submitted	Page Count	Character Count	Date of Thesis Defence	Similarity Index	Submission ID
19/06/2017	157	256495	19/06/2017	%4	826098571

Filtering options applied:

1. Bibliography excluded
2. Quotes excluded
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.


19/06/2017

Name Surname: Kenan Burak YÜKSEL

Student No: N14219758

Department: Department of Educational Sciences

Program: Division of Educational Measurement and Evaluation

Status: Masters Ph.D. Integrated Ph.D.

ADVISOR APPROVAL

APPROVED

Prof.Dr.Nuri DOĞAN



EK 4. BİLGİLENDİRİLMİŞ GÖNÜLLÜ ONAM FORMU

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Bilgilendirilmiş Gönüllü Onam Formu

Sizi Kenan Burak YÜKSEL tarafından yürütülen “Yaratıcılık ve Madde Yazarlığı Eğitiminin Ölçme Aracının Psikometrik Özelliklerine Etkisi” başlıklı araştırmaya davet ediyoruz. Bu araştırmanın amacı madde yazarının yaratıcılık düzeyi ve madde yazarlığı eğitim durumunun ölçme aracına etkisini tespit etmektir. Bu araştırma için kurumunuz Ankara İl Millî Eğitim Müdürlüğü’nden izin alınmıştır. Araştırmada sizden tahminen 40-70 dakika ayırmanız istenmektedir. Araştırmaya sizin dışınızda tahminen 79 kişi katılacaktır. Bu çalışmaya katılmak tamamen **gönüllülük** esasına dayanmaktadır. Çalışmanın amacına ulaşması için sizden beklenen, bütün soruları eksiksiz, kimsenin baskısı veya telkini altında olmadan, size en uygun gelen cevapları içtenlikle verecek şekilde cevaplamanızdır. Bu formu okuyup onaylamanız, araştırmaya katılmayı kabul ettiğiniz anlamına gelecektir. Ancak, çalışmaya katılmama veya katıldıktan sonra herhangi bir anda çalışmayı bırakma hakkına da sahiptir. Bu çalışmadan elde edilecek bilgiler tamamen araştırma amacı ile kullanılacak olup kişisel bilgileriniz **gizli tutulacaktır**; ancak verileriniz yayın amacı ile kullanılabilir. İletişim bilgileriniz ise sadece izninize bağlı olarak ve farklı araştırmacıların sizinle iletişime geçebilmesi için “ortak katılımcı havuzuna” aktarılabilir. Eğer araştırmanın amacı ile ilgili verilen bu bilgiler dışında şimdi veya sonra daha fazla bilgiye ihtiyaç duyarsanız araştırmacıya şimdi sorabilir veya levlalemeecnunveben@gmail.com e-posta adresi ve 0 505 661 11 23 numaralı telefonda ulaşabilirsiniz. Araştırma tamamlandığında genel/size özel sonuçların sizinle paylaşılmasını istiyorsanız lütfen araştırmacıya iletiniz.

Danışman: Prof.Dr.Nuri DOĞAN / Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü/
nurid@hacettepe.edu.tr

Yukarıda yer alan ve araştırmadan önce katılımcıya verilmesi gereken bilgileri okudum ve katılmam istenen çalışmanın kapsamını ve amacını, gönüllü olarak üzerime düşen sorumlulukları anladım. Çalışma hakkında yazılı ve sözlü açıklama aşağıda adı belirtilen araştırmacı tarafından yapıldı. Bana, çalışmanın muhtemel riskleri ve faydaları sözlü olarak da anlatıldı. Kişisel bilgilerimin özenle korunacağı konusunda yeterli güven verildi.

Bu koşullarda söz konusu araştırmaya kendi isteğimle, hiçbir baskı ve telkin olmaksızın katılmayı kabul ediyorum.

Katılımcının :

Adı-Soyadı:.....

İmzası: İletişim Bilgileri: e-posta:

Telefon:

İletişim bilgilerimin diğer araştırmacıların benimle iletişime geçebilmesi için “ortak araştırma havuzuna” aktarılmasını;

kabul ediyorum

kabul etmiyorum

(lütfen uygun seçeneği işaretleyiniz)

Velayet veya Vesayet Altında Bulunanlar İçin:

Veli veya Vasisinin

Adı-Soyadı:.....

İmzası:

Araştırmacının

Adı-Soyadı:.....

İmzası:

Sahidin:

Adı-Soyadı:.....

İmzası:

EK 5. TYDT UYGULAMA İZİN BELGESİ

15/10/2016

T.C.

HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ'NE

Enstitünüzün Eğitimde Ölçme ve Değerlendirme A.B.D. Yüksek Lisans programı öğrencilerinden KENAN BURAK YÜKSEL yüksek lisans tezinde Torrance Yaratıcı Düşünme Testleri A ve B formunu kullanmak istemektedir. Testlerin Türkçe formu yasal kullanım hakkı sahibi olarak testleri kullanmasına tarafımdan izin verilmiştir.

Bilgilerinize arz olunur.



Prof. Dr. A. Esra Aslan
İstanbul Üniversitesi
Hasan Ali Yücel Eğitim Fakültesi
Eğitim Bilimleri Bölümü
Rehberlik ve Psikolojik Danışma A:B.D. Öğretim Üyesi
Tel: 0212 4400000-13052
E-mail: aeaslan@hotmail.com

**EK 6. 2016-2017 EĞİTİM VE ÖĞRETİM YILI 1.DÖNEM 8.SINIF MERKEZİ
ORTAK ASIL SINAV MATEMATİK ALT TESTİ A KİTAPÇIĞI ALAN
UZMANLARININ MADDE-KAZANIM EŞLEŞTİRMESİ FREKANS VE
YÜZDELERİ**

Madde No	Öğrenme Alanı	Alt Öğrenme Alanı	Ölçtüğü Kazanım-Bilişsel Süreç	Seçilme		Seçilen Diğer Program Kazanım Kodu (n)
				f	%	
1		Üslü İfadeler	8.1.2.1. Tam sayıların, tam sayı kuvvetlerini hesaplar, üslü ifade şeklinde yazar. (A)	14	100	
2		Üslü İfadeler	8.1.2.2. Sayıların ondalık gösterimlerini 10'un tam sayı kuvvetlerini kullanarak çözümler. (U)	12	86	8.1.2.5 (2)
3		Çarpanlar ve Katlar	8.1.1.1. Verilen pozitif tam sayıların çarpanlarını bulur; pozitif tam sayıları üslü ifade ya da üslü ifadelerin çarpımı şeklinde yazar. (U)	14	100	
4		Çarpanlar ve Katlar	8.1.1.1. Verilen pozitif tam sayıların çarpanlarını bulur; pozitif tam sayıları üslü ifade ya da üslü ifadelerin çarpımı şeklinde yazar. (U)	14	100	
5		Çarpanlar ve Katlar	8.1.1.2. İki doğal sayının en büyük ortak bölenini (EBOB) ve en küçük ortak katını (EKOK) hesaplar; ilgili problemleri çözer. (U)	14	100	
6		Çarpanlar ve Katlar	8.1.1.2. İki doğal sayının en büyük ortak bölenini (EBOB) ve en küçük ortak katını (EKOK) hesaplar; ilgili problemleri çözer. (U)	14	100	
7		Üslü İfadeler	8.1.2.1. Tam sayıların, tam sayı kuvvetlerini hesaplar, üslü ifade şeklinde yazar. (U)	12	86	8.1.2.3 (2)
8		Çarpanlar ve Katlar	8.1.1.3. Verilen iki doğal sayının aralarında asal olup olmadığını belirler. (U)	14	100	
9	Sayılar ve İşlemler	Üslü İfadeler	8.1.2.3. Üslü ifadelerle ilgili temel kuralları anlar, birbirine denk ifadeler oluşturur. (U)	11	79	8.1.2.1 (3)
10		Üslü İfadeler	8.1.2.4. Sayıları 10'un farklı tam sayı kuvvetlerini kullanarak ifade eder. (A)	12	86	8.1.2.1 (2)
11		Kareköklü İfadeler	8.1.3.1. Tam kare doğal sayıları tanır. (A)	14	100	
12		Üslü İfadeler	8.1.2.3. Üslü ifadelerle ilgili temel kuralları anlar, birbirine denk ifadeler oluşturur. (U)	13	93	8.1.2.1 (1)
13		Kareköklü İfadeler	8.1.3.7. Kareköklü bir ifade ile çarpıldığında, sonucu bir doğal sayı yapan çarpanlara örnek verir. (A)	14	100	
14		Kareköklü İfadeler	8.1.3.2. Tam kare doğal sayılarla bu sayıların karekökleri arasındaki ilişkiyi belirler. (A)	14	100	
15		Kareköklü İfadeler	8.1.3.5. Kareköklü ifadelerle çarpma ve bölme işlemlerini yapar. (U)	14	100	
16		Kareköklü İfadeler	8.1.3.3. Tam kare olmayan sayıların karekök değerlerinin hangi iki doğal sayı arasında olduğunu belirler. (U)	14	100	
17		Kareköklü İfadeler	8.1.3.4. Gerçek sayıları tanır, rasyonel ve irrasyonel sayılarla ilişkilendirir. (A)	14	100	
18		Üslü İfadeler	8.1.2.5. Çok büyük ve çok küçük sayıları bilimsel gösterimle ifade eder ve karşılaştırır. (U)	14	100	
19		Kareköklü İfadeler	8.1.3.5. Kareköklü ifadelerle çarpma ve bölme işlemlerini yapar. (U)	14	100	
20		Kareköklü İfadeler	8.1.3.5. Kareköklü ifadelerle çarpma ve bölme işlemlerini yapar. (A)	14	100	

(A): Anlama, (U): Uygulama

**EK 7. 2016-2017 EĞİTİM VE ÖĞRETİM YILI 1.DÖNEM 8.SINIF MERKEZİ
ORTAK ASIL SINAV TÜRKÇE ALT TESTİ A KİTAPÇIĞI ALAN UZMANLARININ
MADDE-KAZANIM EŞLEŞTİRMESİ FREKANS VE YÜZDELERİ**

Madde Nu	Öğrenme Alanı	Alt Öğrenme Alanı	Ölçtüğü Kazanım-Bilişsel Süreç	Seçilme		Seçilen Diğer Program Kazanım Kodu (n)
				f	%	
1	T8.2 Okuma	Söz Varlığı	T.8.2.30. Okuduğu metindeki sözcük ve sözcük gruplarının cümle içinde kazandığı anlamı fark eder. (U)	12	100	
2	T8.2 Okuma	Anlama	T.8.2.22. Okuduklarındaki söz sanatlarını bulur. (H)	12	100	
3	T8.2 Okuma	Söz Varlığı	T.8.2.30. Okuduğu metindeki sözcük ve sözcük gruplarının cümle içinde kazandığı anlamı fark eder. (U)	12	100	
4	T8.2 Okuma	Anlama	T.8.2.1. Okuduklarının konusunu ve ana fikrini/ana duygusunu belirler. (A)	11	92	8.2.3 (1)
5	T8.2 Okuma	Anlama	T.8.2.1. Okuduklarının konusunu ve ana fikrini/ana duygusunu belirler. (A)	8	67	8.2.3 (4)
6	T8.3 Yazma	-	T.8.3.4.10. Sunduğu görüşleri, gerekçeleri, ve kanıtları yazma amacına uygun bir şekilde sonuca bağlar. (A)	9	75	8.2.26 (2) 8.2.22 (1)
7	T8.2 Okuma	Anlama	T.8.2.3. Okuduklarında yardımcı fikirleri belirler ve yardımcı fikirlerin ana fikri nasıl desteklediğini açıklar. (A)	10	83	8.2.30 (2)
8	T8.2 Okuma	Anlama	T.8.2.1. Okuduklarının konusunu ve ana fikrini/ana duygusunu belirler. (A)	12	100	
9	T8.2 Okuma	Anlama	T.8.2.6. Okuduklarındaki yönlendirici ifadelerin anlamda yaptığı değişikliği fark eder. (A)	12	100	
10	T8.2 Okuma	Anlama	T.8.2.3. Okuduklarında yardımcı fikirleri belirler ve yardımcı fikirlerin ana fikri nasıl desteklediğini açıklar. (A)	12	100	
11	T8.3 Yazma	-	T.8.3.4.9.Paragraflar veya metnin bölümleri arasında ana fikri, görüş veya tezi destekleyecek şekilde bağlantı kurarak metnin kendi içinde tutarlılığını sağlar. (U)	12	100	
12	T8.2 Okuma	Anlama	T.8.2.13. Hikâye edici metinlerde olay örgüsünü çözümler. (A)	7	58	8.2.5 (5)
13	T8.2 Okuma	Anlama	T.8.2.1. Okuduklarının konusunu ve ana fikrini/ana duygusunu belirler. (U)	8	67	8.2.5 (4)
14	T8.2 Okuma	Anlama	T.8.2.22. Okuduklarındaki söz sanatlarını bulur. (A)	12	100	
15	T8.2 Okuma	Anlama	T.8.2.12. Okuduğu metnin türünü belirler. (A)	12	100	
16	T8.2 Okuma	Akıcı Okuma	T.8.2.27. Yazım ve noktalama işaretlerine dikkat ederek, uygun hızda sesli veya sessiz okur. (A)	12	100	
17	T8.2 Okuma	Akıcı Okuma	T.8.2.27. Yazım ve noktalama işaretlerine dikkat ederek, uygun hızda sesli veya sessiz okur. (A)	12	100	
18	T8.2 Okuma	Söz Varlığı	T.8.2.31. Fiilimsilerin cümledeki işlevini fark eder. (A)	9	75	8.3.9 (3)
19	T8.3 Yazma	-	T.8.3.8. Cümlelerin öğelerini bilir. (A)	12	100	
20	T8.3 Yazma	-	T.8.3.8. Cümlelerin öğelerini bilir. (U)	12	100	

(H): Hatırlama, (A): Anlama, (U): Uygulama

EK 8. ARAŞTIRMADA MATEMATİK ÖĞRETMENLERİNİN GELİŞTİRDİĞİ MADELERDEN BİR KISMI

Yaratıcılığı Yüksek Gruptaki Öğretmenler Tarafından Geliştirilen Maddeler (Form I)	Yaratıcılığı Düşük Gruptaki Öğretmenler Tarafından Geliştirilen Maddeler (Form II)	Madde Yazarlığı Eğitimi Alan Deney Grubundaki Öğretmenler Tarafından Geliştirilen Maddeler (Form III)
--	--	---

Kazanım: 8.1.2.2. Sayıların ondalık gösterimlerini 10'un tam sayı kuvvetlerini kullanarak çözümler. (U)
Madde No: 2

<p>Aşağıdakilerden hangisinin sonucu diğerlerinden farklıdır?</p> <p>a) $0,0716 \cdot 10^{-1}$ b) $0,716 \cdot 10^{-2}$ c) $7,16 \cdot 10^{-3}$ d) $716 \cdot 10^{-4}$</p> <p>pi=0,43 rjx=0,30 Ç=0,70</p>	<p>10,724 sayısının ondalık gösteriminin çözümlenmiş hali aşağıdakilerden hangisidir?</p> <p>a) $1 \cdot 10^2 + 7 \cdot 10^3 + 2 \cdot 10^2 + 4 \cdot 10^2$ b) $1 \cdot 10^5 + 7 \cdot 10^3 + 2 \cdot 10^2 + 4 \cdot 10^1$ c) $1 \cdot 10^2 + 7 \cdot 10^1 + 2 \cdot 10^2 + 4 \cdot 10^1$ d) $1 \cdot 10^2 + 7 \cdot 10^{-1} + 1 \cdot 10^2 + 4 \cdot 10^{-1}$</p> <p>pi=0,50 rjx=0,36 Ç=0,66</p>	<p>$0,000256 = 25,6 \cdot 10^3$ $60000 = 6,0 \cdot 10^b$ ise $a+b=?$</p> <p>a) 1 b) 0 c) -1 d) -2</p> <p>pi=0,43 rjx=0,40 Ç=0,98</p>
---	---	---

Kazanım: 8.1.2.1. Tam sayıların, tam sayı kuvvetlerini hesaplar, üslü ifade şeklinde yazar. (U)
Madde No: 7

<p>$3^2[(3)^{-2} + (-2)^{-3}] \cdot 2^3$ işleminin sonucu kaçtır?</p> <p>a) -2 b) -1 c) 1 d) 2</p> <p>pi=0,39 rjx=0,29 Ç=0,85</p>	<p>4 tane 2'nin çarpımının, 4 tane 2'nin toplamına oranı kaçtır?</p> <p>a) 2^3 b) 2^2 c) 2^1 d) $2^{4/8}$</p> <p>pi=0,79 rjx=0,28 Ç=0,62</p>	<p>Bir büyükşehir belediyesi sular idaresi başkanlığı su tüketiminin ilk 32 m^3'ü için m^3 başına ₺2 sonraki her m^3 tüketim için ₺4 olarak birim fiyat belirliyor. Buna göre bir çiftçi tarlasını sulamak için 96 m^3 su kullandığında kaç ₺ su faturası öder?</p> <p>a) $5,2^6$ b) $3,2^6$ c) $7,2^5$ d) 2^7</p> <p>pi=0,49 rjx=0,33 Ç=0,94</p>
--	--	---

Kazanım: 8.1.3.1. Tam kare doğal sayıları tanır. (A)
Madde No: 11

<p>Şubat ayının tam kare sayı olan günleri kaç tanedir?</p> <p>a) 2 b) 3 c) 4 d) 5</p> <p>pi=0,29 rjx=0,24 Ç=0,72</p>	<p>Alanı 289 m^2 olan kare şeklindeki bir tarlanın sınırlarını belirlemek için çevresine $1\text{m}'de$ bir çubuk dikilerek tel çekilecektir. Bu iş için kaç adet çubuk gerekmektedir?</p> <p>a) 68 b) 52 c) 64 d) 76</p> <p>pi=0,54 rjx=0,36 Ç=0,63</p>	<p>Üç basamaklı 1a1 tam kare doğal sayıdır. Buna göre, üç basamaklı aa5 sayısının karekökü kaçtır?</p> <p>a) 15 b) 25 c) 34 d) 45</p> <p>pi=0,51 rjx=0,37 Ç=0,89</p>
---	--	--

Kazanım: 8.1.3.5. Kareköklü ifadelerle çarpma ve bölme işlemlerini yapar. (A)
Madde No: 20

<p>$\frac{\sqrt{32}}{\sqrt{8}} : \frac{\sqrt{20}}{\sqrt{125}} = ?$</p> <p>a) $4/5$ b) 2 c) $2\sqrt{2}$ d) 5</p> <p>pi=0,76 rjx=0,21 Ç=0,65</p>	<p>$\frac{\sqrt{0,81} + \sqrt{0,64} - \sqrt{0,36}}{\sqrt{1,21}} = ?$</p> <p>a) $0, \bar{1}$ b) 1 c) 1,1 d) $\sqrt{11}$</p> <p>pi=0,46 rjx=0,12 Ç=0,69</p>	<p>$\frac{\sqrt{80} - \sqrt{20}}{\sqrt{45}} = ?$</p> <p>a) $3/4$ b) $2/3$ c) 1 d) $4/3$</p> <p>pi=0,48 rjx=0,48 Ç=0,84</p>
---	--	--

EK 9 ARAŞTIRMADA TÜRKÇE ÖĞRETMENLERİNİN GELİŞTİRDİĞİ MADELERDEN BİR KISMI

Yaratıcılığı Yüksek Gruptaki Öğretmenler Tarafından Geliştirilen Maddeler (Form I)	Yaratıcılığı Düşük Gruptaki Öğretmenler Tarafından Geliştirilen Maddeler (Form II)	Madde Yazarlığı Eğitimi Alan Deneysel Grubundaki Öğretmenler Tarafından Geliştirilen Maddeler (Form III)
--	--	--

Kazanım: T.8.2.30. Okuduğu metindeki sözcük ve sözcük gruplarının cümle içinde kazandığı anlamı fark eder. (U)

Madde No: 3

<p>Aşağıdakilerden hangisinde "dolanmak" sözcüğünün kullanımı, parantez içindeki anlamıyla uyumamaktadır?</p> <p>a) Zaman dediğin nedir ki; ay dolanır, yıllar geçer. (Bir şeyin çevresine sarılmak, etrafını kuşatıp sarmak)</p> <p>b) Dillerinde senin ismin dolanır. (Yayılmak ,elden ele, ağızdan ağza gezmek anlatılmak)</p> <p>c) Çantamdaki yün çileleri birbirine dolanmış. (Karışmak, güç çözülür hale gelmek)</p> <p>d) Mecnun gibi dolanıyorum çöllerde. (Gezinmek, orada burada gezip durmak)</p> <p style="text-align: right;">pi=0,64 rjx=0,29 Ç=0,94</p>	<p>Kapının yanında duran iki adam, ağız ağıza bir şeyler konuşuyorlar, yer yer ----- bize bakıyorlardı. Bu cümlede boş bırakılan yere aşağıdaki deyimlerden hangisi getirilirse, cümle "anlaşılmadan" anlamı kazanır?</p> <p>a) yan gözle</p> <p>b) gözünü kırpmadan</p> <p>c) gözünü açmış</p> <p>d) gözünü ağartmış</p> <p style="text-align: right;">pi=0,78 rjx=0,23 Ç=0,60</p>	<p>Yazar, öykü yazmayı <u>toplumun üzerinde ayna gezdirmeye</u> benzetir.</p> <p>Bu cümlede altı çizili sözle anlatılmak istenen aşağıdakilerden hangisidir?</p> <p>a) Roman toplumsal gerçekliği yansıtmaya görevi üstlenir.</p> <p>b) Her sanat yapıtı yazarından izler taşır.</p> <p>c) Roman gerçeğin düşünle birleştiği yerde doğar.</p> <p>d) İyi bir roman hayal unsurlarını barındırır.</p> <p style="text-align: right;">pi=0,41 rjx=0,39 Ç=0,90</p>
---	---	--

Kazanım: T.8.3.4.9.Paragraflar veya metnin bölümleri arasında ana fikri, görüş veya tezi destekleyecek şekilde bağlantı kurarak metnin kendi içinde tutarlılığını sağlar. (U)

Madde No: 11

<p>Geç kalktım bugün çünkü okullar tatil olduğu için erkenden uyanmama gerek yoktu. Aslında uykuyu pek sevmem.Yataktan çıkasım gelmez.</p> <p>Boşluğa uygun ifade nedir?</p> <p>a) Yorulursam çok uyurum o da ayrı.</p> <p>b) Okula gitmek zor geliyor bazen.</p> <p>c) Ama okula gitmem gerekince nedense uyanmak istemem.</p> <p>d) Çok rahatsız olduğum zamanlarda uyurum ya da yorgunken.</p> <p style="text-align: right;">pi=0,81 rjx=0,09 Ç=0,70</p>	<p>En sevdiğin elbiseleri giy bugün, Şiir oku, şarkı söyle, coş,, Kirlara, parklara koş.</p> <p>Yukarıdaki dizelere en uygun ifade hangisidir?</p> <p>a) Tadını çıkar, durma günün,</p> <p>b) Sağın solun arkan önün,</p> <p>c) Belki de en güzel günü ömrünün,</p> <p>d) Çalışmaya verme ödün,</p> <p style="text-align: right;">pi=0,58 rjx=0,21 Ç=0,69</p>	<p>(I)Yaz yağmuru ıslatır ama üşütmez. (II)Aynı yaz yağmuru gibi dostları vardır insanın.(III).....(IV) Yine de dosttur, küsülmez.</p> <p>(III) ile verilen boşluğa aşağıdaki cümlelerden hangisi yerleştirilirse anlam bütünlüğü sağlanır?</p> <p>a) İyi günü olduğu gibi kötü günü de vardır.</p> <p>b) Zaman zaman istemediğin gibi davranabilir.</p> <p>c) Zarar verir, kızdırır yeri gelir bağırır, çağırır.</p> <p>d) Seni üzecek şeyler yapmaz da denemez.</p> <p style="text-align: right;">pi=0,44 rjx=0,43 Ç=0,90</p>
--	---	--

ÖZGEÇMİŞ

Kişisel Bilgiler

<i>Adı Soyadı</i>	Kenan Burak YÜKSEL
<i>Doğum Yeri</i>	Ankara
<i>Doğum Tarihi</i>	12.10.1980

Eğitim Durumu

<i>Lise</i>	Maltepe Askeri Lisesi-Güzelbahçe/İZMİR	1998
<i>Lisans</i>	Kara Harp Okulu-Sistem Mühendisliği	2002
<i>Yüksek Lisans</i>	Hacettepe Üniversitesi-Eğitimde Ölçme ve Değerlendirme	2017
<i>Yabancı Dil</i>	İngilizce: Okuma (Çok iyi), Yazma (İyi), Konuşma (İyi)	

İş Deneyimi

<i>Stajlar</i>	-	
<i>Projeler</i>	-	
<i>Çalıştığı Kurumlar</i>	Kara Kuvvetleri Komutanlığı-Takım/Bölük Komutanlığı	2002-

Akademik Çalışmalar

Yayınlar

-

Seminer ve Çalıştaylar

-

Sertifikalar

-

İletişim

<i>e-Posta Adresi</i>	leylalemecnunveben@gmail.com

<i>Jüri Tarihi</i>	13/06/2017
--------------------	------------