



T.C.

EGE ÜNİVERSİTESİ

Eğitim Bilimleri Enstitüsü

**SAYISAL YETENEK FARKLILAŞMASININ CİNSİYETE GÖRE
İNCELENMESİ: BİR DEĞİŞEN MADDE FONKSİYONU(DMF)
ÇALIŞMASI**

YASEMİN YARDIM PALAMUTLU

YÜKSEK LİSANS TEZİ

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI

**İzmir
2019**

T.C.

EGE ÜNİVERSİTESİ

Eğitim Bilimleri Enstitüsü

**SAYISAL YETENEK FARKLILAŞMASININ CİNSİYETE GÖRE
İNCELENMESİ: BİR DEĞİŞEN MADDE FONKSİYONU(DMF)
ÇALIŞMASI**

**INVESTIGATION OF QUANTITATIVE ABILITY DIFFERENCE
BASED ON GENDER: A DIFFERENTIAL ITEM FUNCTIONING(DIF)
STUDY**

YASEMİN YARDIM PALAMUTLU

YÜKSEK LİSANS TEZİ

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI

**Tez Danışmanı
Prof. Dr. Tuncay ÖĞRETMEN**

**2019
İZMİR**



T.C.EGE ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ



YÜKSEK LİSANS

TEZ SAVUNMA TUTANAĞI

ÖĞRENCİNİN

Adı Soyadı : Yasemin Yardım PALAMUTLU
Numarası : 92160000059
Anabilim Dalı : Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme B.D.
Tez Başlığı (Türkçe) : Sayısal Yetenek Farklaşmasının Cinsiyete Göre İncelenmesi: Bir Değişen Madde Fonksiyonu (DMF) Çalışması
Tez Başlığı (İngilizce) : Investigation of Quantitative Ability Difference Based on Gender: A Differential Item Functioning (DIF) Study
Tez Savunma Tarihi : 09.09.2019
Tez Başlığı Değişikliği Varsa Yeni Başlık: -

JÜRİ ÜYELERİ

Jüri Başkanı

Unvan, Adı, Soyadı : Prof. Dr. Tuncay ÖĞRETMEN (Danışman)
Karar : O Başarılı O Başarısız O Düzeltme
İmza :

Jüri Üyesi

Unvan, Adı, Soyadı : Doç. Dr. Oğuz BAŞOKÇU
Karar : O Başarılı O Başarısız O Düzeltme
İmza :

Jüri Üyesi

Unvan, Adı, Soyadı : Doç. Dr. Duygu GÜNGÖR
Karar : Başarılı O Başarısız O Düzeltme
İmza :

TEZ HAKKINDA JÜRİNİN GENEL GÖRÜŞÜ

(Jüri Başkanı Tarafından Doldurulacaktır)

Tez savunması sonucunda öğrenci tarafından hazırlanan çalışma;

Oybirliğiyle

Oy çokluğuyla

Başarılıdır


Düzeltilmelidir

Başarısızdır

- Bu tutanak üç (3) işgünü içerisinde jüri üyelerinin raporlarıyla beraber Anabilim Dalı Başkanlığı üst yazısıyla Enstitü Müdürlüğüne gönderilmelidir.
- Tezli yüksek lisans programlarında düzeltme alan öğrencinin 3 (üç) ay içerisinde yeniden savunmaya girmesi zorunludur.

EGE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ
ETİKKURALLARA UYGUNLUK BEYANI

EÜ Lisansüstü Eğitim ve Öğretim Yönetmeliğinin ilgili hükümleri uyarınca Yüksek Lisans Tezi olarak sunduğum “Sayısal Yetenek Farklılaşmasının Cinsiyete Göre İncelenmesi: Bir Değişen Madde Fonksiyonu(DMF) Çalışması” başlıklı bu tezin kendi çalışmam olduğunu, sunduğum tüm sonuç, doküman, bilgi ve belgeleri bizzat ve bu tez çalışması kapsamında elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara atıf yaptığımı ve bunları kaynaklar listesinde usulüne uygun olarak verdiğimi, tez çalışması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını, bu tezin herhangi bir bölümünü bu üniversite veya diğer bir üniversite başka bir tez çalışması içinde sunmadığımı, bu tezin planlanmasında yazımına kadar bütün safhalarda bilimsel etik kurallarına uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul edeceğimi beyan ederim.


Yasemin Yardım Palamutlu

ÖNSÖZ

Dünyada ve ülkemizde, bireylerin okul yaşantılarındaki başarılarının ölçülmesi ve kariyer basamaklarında ilerlemelerinin sağlanması için çeşitli ölçümler yapılmaktadır. Bu ölçümlerin yapılabilmesi için çeşitli testler kullanılır. Bir Matematik Öğretmeni olarak öğrencilerin girdikleri sınavlarda cinsiyet ya da kültürel farklılaşmadan kaynaklanan yanlış maddelerin olup olmadığı her zaman merak ettiğim bir konu olmuştur. Bu konu ile ilgili merakımı ÖLÇME ve DEĞERLENDİRME Anabilim dalında yüksek lisansımı yaparken, özellikle sayısal yetenek testlerinde cinsiyet temelli bir farklılaşmanın olup olmadığını araştırmaya itmiştir. Bu araştırma kapsamında uygulanmış ulusal seçme ve yerleştirme sınavlarından elde edilen verileri kullanmak yerine, kendi hazırladığım bir sayısal yetenek testi geliştirdim. Test geliştirme basamaklarını uygularken, bu çalışma mesleki gelişimime de fazlasıyla katkıda bulundu.

Meslek seçimi ve kurumlara yerleşme süreçlerinde kız ve erkek bireylerin girdikleri sınavlardaki testlerin kızların veya erkeklerin lehine ya da aleyhine olmaması gerekir.

Matematikten başarılı olma yeteneğini eğitim deneyimlerinden elde edilen başarıdan ayırt edemeyiz. Yetenek ile başarı arasındaki bulanık ayrım, cinsiyet farklılıkları gösteren testlerden aldığımız sonuçlara çok dikkat etmemiz gerektiği anlamına gelir.

Yeteneği ölçmeye çalıştığımızda bir dereceye kadar her zaman başarıyı ölçmüş oluyoruz. Hayatlarımız boyunca, ama özellikle de okul yıllarında herkes çok sayıda teste tabi tutulur. Bu testlerden bazıları yüksek öneme sahiptir, yani yaşamları etkileyen kararlar için temel oluştururlar. Yüksek öneme sahip sınavların çoğunda, cinsiyet, ırk veya etnik grup, sosyo-ekonomik durum ve coğrafi bölgenin bir fonksiyonu olarak farklılık gösteren grup farklılıkları veya ortalama farklılıklar vardır. Bu farkların ortadan kaldırılması veya en aza indirilmesi adil bir ölçme ve değerlendirme için önemlidir.

Bu araştırma, hem mesleki gelişimime katkı sağlaması, hem ilgi alanımı içermesi dolayısıyla pek çok literatür taraması yaparak bilgi birikimimi arttırması, hem de bilimsel bir çalışma yapmanın bana verdiği heyecandan dolayı keyifli fakat bir o kadar da zorlu bir süreç olmuştur.

Bu çalışmanın sonunda elde ettiğim sonuçların, ulusal test uygulayıcılarına ve bilimsel literatüre ışık tutması dileğiyle.

İzmir, 03.09.2009

Yasemin YARDIM

TEŞEKKÜR

Yüksek Lisans eğitimim süresince ve bu tezin hazırlanmasında desteğini esirgemeyen, çalışmamı görüş ve önerileriyle yönlendiren, her zaman örnek aldığım ve birlikte çalışma şansına eriştiğim, bana destek olup bilgi ve önerilerini paylaşan, her aşamada yol gösterici olan değerli danışmanım Sayın Prof. Dr. Tuncay ÖĞRETMEN'e,

Yüksek lisans eğitimim boyunca kıymetli bilgilerinden yararlandığım Ege Üniversitesi Ölçme ve Değerlendirme Anabilim Dalı değerli hocaları Sayın Doç. Dr. Hakan ATILGAN ve Sayın Doç. Dr. Tahsin Oğuz BAŞOKÇU'ya, bu süreçte birlikte çalıştığım, her aşamada manevi desteklerini esirgemeyen sevgili dönem arkadaşlarım Dilay AK, Pelin BAĞDU SÖYLER ve Simge CEYLAN'a

Hem çalışıp hem yüksek lisansımı tamamlama gayretleri içindeyken bana destek olan Karşıyaka Lisesi idarecilerine, verileri elde etme aşamasında üstün gayretleriyle her zaman yanımda olan Karşıyaka Lisesi ailesindeki tüm mesai arkadaşlarıma,

Gerek yüksek lisans ders dönemi, gerekse tez yazma dönemi boyunca sürekli onun zamanından çaldığım biricik oğlum Arda PALAMUTLU'ya, desteğini ve anlayışını sürekli üzerimde hissettiğim Cüneyt PALAMUTLU'ya, ayrıca beni her zaman destekleyen aileme ve dostlarıma teşekkürlerimi ve saygılarımı sunarım.

ÖZGEÇMİŞ

KİMLİK BİLGİLERİ

Adı Soyadı : Yasemin YARDIM PALAMUTLU

Doğum Yeri : Almanya-Münih

Doğum Tarihi : 06.03.1979

E-posta : yaseminyyp@gmail.com

EĞİTİM BİLGİLERİ

Lise : Yüce Fen Lisesi, Ankara

Lisans : Hacettepe Üniversitesi Matematik Öğretmenliği

Yabancı Dil ve Düzeyi: D (YDS 2016)

Katıldığı Eğitimler:

Practical Rasch Measurement-Further Topics (Statistic.com)

İÇİNDEKİLER

| | |
|---|------|
| ÖNSÖZ..... | iii |
| TEŞEKKÜR..... | iv |
| ÖZGEÇMİŞ..... | v |
| İÇİNDEKİLER..... | vi |
| TABLolar LİSTESİ..... | x |
| ŞEKİLLER LİSTESİ..... | xi |
| EKLER LİSTESİ..... | xii |
| KISALTMALAR..... | xiii |
| ÖZET..... | xv |
| EXTENDED ABSTRACT..... | xvi |
| BÖLÜM I..... | 1 |
| GİRİŞ..... | 1 |
| 1.1. Problem Durumu..... | 1 |
| 1.1.1.Zekanın Tanımı..... | 3 |
| 1.1.2. Çağımızın Zeka Kuramlarından Bazıları..... | 4 |
| 1.1.3. Zeka Testlerinin Tarihsel Temelleri..... | 8 |
| 1.1.4. Sayısal Akıl Yürütme Yapılarının Gelişimi..... | 10 |
| 1.1.5. Sayısal Yetenek..... | 12 |
| 1.1.6. Cinsiyete Dayalı Farklılıkların İşaretlerini Değişen Madde Fonksiyonunun Psikometrik-Çok İçerik Teorisini Kullanarak İncelenmesi..... | 16 |

| | |
|--|-----------|
| 1.1.7. Sayısal Yeteneğin Cinsiyete Göre Farklılaşmasının İncelenmesinin Önemi..... | 17 |
| 1.1.8. Yanlılık Tanımı..... | 22 |
| 1.2. Problem Cümlesi..... | 23 |
| 1.3 Alt Problemler..... | 23 |
| 1.4 Sınırlılıklar..... | 23 |
| 1.5 Araştırmanın Amacı ve Önemi..... | 23 |
| BÖLÜM II..... | 25 |
| İLGİLİ ARAŞTIRMALAR..... | 25 |
| 2.1. Genel Zihinsel Yetenek Testleri ve Zeka Testleri Üzerine Yapılan Tartışmalar..... | 25 |
| 2.2. Sayısal Yetenek Testlerinde Cinsiyet Farklılıklarına İlişkin Açıklamalar..... | 26 |
| 2.3 Konu ile İlgili Yurtdışında Yapılan Bazı Çalışmalar..... | 27 |
| 2.4. Konu ile İlgili Türkiye’de Yapılan Bazı Çalışmalar..... | 30 |
| BÖLÜM III..... | 34 |
| YÖNTEM..... | 34 |
| 3.1 Araştırmanın Türü..... | 34 |
| 3.2. Evren ve Örneklem..... | 34 |
| 3.3. Verilerin Toplanması..... | 35 |
| 3.4. Veri Toplama Aracı..... | 35 |
| 3.5. Verilerin Analizi..... | 35 |
| 3.5.1. Seçilen Bilgisayar Programı Hakkında Genel Bilgi..... | 36 |
| 3.5.2. Madde Tepki Kuramı Hakkında Genel Bilgiler..... | 36 |
| 3.5.3. Madde Yanıt Teorisinde En Çok Kullanılan Modeller..... | 39 |

| | |
|--|-----------|
| 3.5.4. Madde Karakteristiğinin Bir Fonksiyonu Olarak DMF..... | 41 |
| 3.5.5. DMF Belirleme..... | 42 |
| 3.5.6. DMF Belirleme Teknikleri..... | 44 |
| 3.5.7. Rasch Analizi Hakkında Genel Bilgiler..... | 45 |
| 3.5.7.1. Rasch Model ve Madde Yanıt Modeli..... | 46 |
| 3.5.7.2. Rasch Modelinin Varsayımları..... | 49 |
| 3.5.7.3. Rasch Modelinin Özellikleri..... | 49 |
| BÖLÜM IV..... | 55 |
| BULGULAR..... | 55 |
| 4.1. Örneklem Büyüklüğü..... | 55 |
| 4.2. Maddelerin Analizi..... | 55 |
| 4.3. Rasch Analizi Sonuçları..... | 62 |
| 4.3.1. Data-Model Uyumu..... | 63 |
| 4.3.2. Kişi-Model Uyumu ve Test-Model Uyumu Testleri..... | 66 |
| 4.3.3. DMF Araştırması..... | 67 |
| 4.3.4. Tek Boyutluluk İncelemesi..... | 71 |
| 4.3.5. Yerel Bağımsızlık Varsayımının Değerlendirilmesi..... | 73 |
| BÖLÜM V..... | 75 |
| SONUÇ, TARTIŞMA VE ÖNERİLER..... | 75 |
| 5.1. Sonuç..... | 75 |
| 5.2. SYT'nin Psikometrik Özelliklerinin Tartışılması..... | 76 |
| 5.3. SYT'nin Rasch Modelinin Varsayımlarını Karşılayıp Karşılamadığının Tartışılması..... | 77 |

| | |
|---|-----------|
| 5.4. SYT Maddelerinin Cinsiyet Gruplarına Göre Değişen Madde Fonksiyonu(DMF) İçerip İçermediğinin Tartışılması..... | 79 |
| 5.5. Sayısal Yetenekte Cinsiyete Göre Farklılaşmanın Olup Olmadığının Tartışılması..... | 82 |
| 5.6. Öneriler..... | 85 |
| KAYNAKLAR..... | 87 |
| EKLER..... | 92 |
| Ek 1:Sayısal Yetenek Testi..... | 92 |
| Ek 2:SYT'nin Uygulama İzni..... | 98 |

TABLolar LİSTESİ

| | |
|---|----|
| Tablo 1: Zekâ Tanımları..... | 3 |
| Tablo 2: Sayısal Akıl Yürütmenin Tanımları..... | 14 |
| Tablo 3: 2017 Lisans Yerleştirme Sınavı Alanlara Göre Kız ve Erkek Öğrencilerin Ortalama Puanları..... | 20 |
| Tablo 4: 2016-2017 Eğitim Öğretim Yılı LYS Sınavı Sonuçlarına Göre Sayısal Alanlara Yerleşen Kız ve Erkek Öğrenci Sayıları..... | 20 |
| Tablo 5: Test1 ve Test2 nin davranış dağılımı..... | 58 |
| Tablo 6: Test1 ve Test2 nin Madde İstatistikleri..... | 59 |
| Tablo 7: Test1 ve Test2 nin Genel Test İstatistikleri..... | 61 |
| Tablo 8: Nihai Test Genel Test İstatistikleri..... | 62 |
| Tablo 9: Maddelerin Uyum İstatistikleri..... | 63 |
| Tablo 10: Öğrenci-Model Uyum İstatistikleri..... | 66 |
| Tablo 11: Test-Model Uyum İstatistikleri..... | 67 |
| Tablo 12: DMF Anlamlılık Testi Sonuçları..... | 69 |
| Tablo 13: DMF İçeren Maddeler..... | 71 |
| Tablo 14: Özdeğer Birimi İçindeki Standart Artık Varyans Tablosu..... | 72 |
| Tablo 15: En Yüksek Standart Artık Korelasyonları..... | 73 |

ŞEKİLLER LİSTESİ

| | |
|--|----|
| Şekil 1: 2005 Ulusal Eğitimsel İlerlemeyi Değerlendirme Raporunda Erkek ve Kızların Puanı..... | 18 |
| Şekil 2: Sınava Giren İki Grup İçin Yetenek Dağılımları ve Madde Karakteristik Eğrisi..... | 37 |
| Şekil 3: DMF Görüntülemeyen Bir Madde Örneği..... | 43 |
| Şekil 4 : Uniform DMF Görüntüleyen Bir Madde Örneği..... | 43 |
| Şekil 5: Nonuniform DMF Görüntüleyen Bir Madde Örneği..... | 44 |
| Şekil 6: Bir Madde için Madde Karakteristik Eğrisi..... | 48 |
| Şekil 7: Madde 9'un Beklenen Madde Karakteristik Eğrisi..... | 64 |
| Şekil 8: Madde 10'un Beklenen Madde Karakteristik Eğrisi..... | 65 |
| Şekil 9: Madde 17'nin Beklenen Madde Kararkteristik Eğrisi..... | 65 |
| Şekil 10: DMF Ölçüsü Grafiği..... | 68 |
| Şekil 11: DMF Büyüklüğü Grafiği..... | 68 |
| Şekil 12: DMF t-değerleri Grafiği..... | 69 |

EKLER LİSTESİ

| | |
|--|-----|
| EK 1: Sayısal Yetenek Testi..... | 94 |
| EK 2: Sayısal Yetenek Testinin Uygulama İzni..... | 100 |



KISALTMALAR

| | |
|---------------|---------------------------------------|
| DMF | Değişen Madde Fonksiyonu |
| DIF | Differential Item Functioning |
| MTK | Madde Tepki Kuramı |
| MKE | Madde Karakteristik Eğrisi |
| IRT | Item Response Theory |
| SYT | Sayısal Yetenek Testi |
| QAT | Quantitative Ability Test |
| SAT | Scholastic Aptitude Test |
| SAT-M | Scholastic Aptitude Test -Mathematics |
| ACT | American College Test |
| ACTM | American College Test Mathematics |
| ÖSYM | Öğrenci Seçme ve Yerleştirme Merkezi |
| LYS | Lisans Yerleştirme Sınavı |
| YGS | Yükseköğretime Geçiş Sınavı |
| ABD | Amerika Birleşik Devletleri |
| APA | American Psychological Association |
| P.M.A. | Primary Mental Abilities Tests |
| IQ | Intelligence Quotient |
| MYM | Madde Yanıt Modeli |
| MGD | Madde Güçlük Dönüşümü |
| MH | Mantel-Haenszel |

| | |
|-----------------|--|
| LR | Lojistik Regresyon |
| IRT-RDIF | Likelihood Ratio Test |
| SDMF | Standardize Deęişen Madde Fonksiyonu |
| QSAT | Sayısal Skolastik Beceri Testi |
| SMPY | Matematiksel Hassas Gençlik Çalışması |
| TIMSS | Trends in International Mathematics and Science Study |
| OKS | Orta Öğretim Kurumlar Sınavı |
| PISA | Programme for International Student Assessment |
| SBS | Seviye Belirleme Sınavı |

ÖZET

SAYISAL YETENEK FARKLILAŞMASININ CİNSİYETE GÖRE İNCELENMESİ: BİR DEĞİŞEN MADDE FONKSİYONU(DMF) ÇALIŞMASI

YARDIM PALAMUTLU, Yasemin

Yüksek Lisans Tezi, Eğitimde Ölçme ve Değerlendirme

Supervisor: Prof. Dr. Tuncay ÖĞRETMEN

Eylül, 2019

Bu tez çalışmasında, Sayısal Yetenekteki farklılaşma cinsiyete göre incelenmiştir. Bunun için önce her bir alt testi 10'ar maddeden oluşan toplam 40 maddelik bir test oluşturuldu. Bu test yaklaşık olarak 400 kişilik 2017-2018 eğitim öğretim yılında 11. sınıfta okuyan öğrencilerden oluşan örnekleme uygulandı. Bu uygulama sonrasında madde güçlükleri ve ayırdedicilikleri hesaplanarak madde seçimi yapıldı. Güvenirlik ve geçerlilik çalışmaları yapılarak, ölçmek istediğimiz yapıları (Aritmetik Akıl Yürütme, Eşitlik Kurma, Geometri ve Görsel-Uzamsal Yetenek) en iyi ölçen 5'er madde seçilip, teste son şekli verildi. Teste son şekli verildikten sonra 20 maddeden oluşan sayısal yetenek testi 2018-2019 eğitim öğretim yılında Bornova Anadolu Lisesi, Cihat Kora Anadolu Lisesi, 15 Temmuz Şehitler Anadolu Lisesi, Atatürk Lisesi, Karşıyaka Lisesi, Atakent Anadolu Lisesi okullarının 11. Sınıfında okuyan 1382 öğrenciye uygulandı. Analizler gerekli elemeler yapıldıktan sonra 1036 kişilik örneklem büyüklüğü için gerçekleştirildi.

Sayısal yetenekte cinsiyete göre farklılık gösteren maddeler tespit edildi. Değişen Madde Fonksiyonu(DMF) içeren maddelerde, bu sonuçlara göre erkek öğrencilerin lehine DMF içeren, eşitlik kurma-sayısal akıl yürütme alanında iki soru olduğu, kız öğrencilerin lehine DMF içeren geometri ve görsel uzamsal yetenek alanında üç madde olduğu bulunmuştur.

Ayrıca oluşturulan SYT'nin Rasch modeli ile uyumu test edilerek raporlandı. SYT maddelerinden bir tanesi hariç diğer maddelerin modele iyi uyum sağladığı, data-model, kişi-model ve test-model uyumunun iyi olduğu, yerel bağımsızlık ve tek boyutluluk varsayımlarını karşıladığı görüldü.

Anahtar kelimeler: Sayısal Yetenek Testi, Rasch Model, Değişen Madde Fonksiyonu(DMF)

EXTENDED ABSTRACT

INVESTIGATION OF QUANTITATIVE ABILITY DIFFERENCE BASED ON GENDER: A DIFFERENTIAL ITEM FUNCTIONING(DIF) STUDY

YARDIM PALAMUTLU, Yasemin

MS, Department of Measurement and Evaluation

Supervisor: Prof. Dr. Tuncay ÖĞRETMEN

September, 2019

1. Introduction

It is common for women to have stronger verbal skills and better memory for events, words, objects, faces and activities. Men are also often thought to be successful in manipulating objects mentally and in numerical tasks based on visual presentations. Because grades and general test scores depend on many factors, psychologists tend to assess better defined cognitive skills to understand these gender differences. Preschool children were born with neither more nor less numerical ability, because girls and boys perform equally well in early cognitive skills related to numerical thinking and knowledge of surrounding objects. However, when school starts, these skills begin to differentiate between genders (Perry, 2015).

Numerical reasoning; arithmetic, number meaning, deductive reasoning, mathematical literacy, numerical literacy, problem solving, contextualized mathematics, mathematical modeling and quantitative reasoning is a complex structure with many names and definitions (Mayes, Peterson, & Bonilla, 2013, p. 2nd).

Reasoning skills and content form the basis for defining the structure of numerical reasoning for evaluation purposes in the world of mathematics. In measurement theory and practice, the conceptual parallel study by Messick et al. (Dwyer et al., 2003).

The ability to think arithmetically is crucial to being a democratic citizen because it enables them to make informed decisions on complex national and international issues that affect their home, workplace and local communities.

It is important at this stage to investigate whether there is a gender-based differentiation in numerical ability tests. In the process of choosing a profession and

settling in institutions, the tests in the examinations of girls and boys should not be in favor of or against girls or boys.

Item bias is that different groups respond to the same test item differently. These differences require discovery as they can shed light on both the test item and the experiences and backgrounds of the different groups taking the exam (Holland & Thayer, 1986a). Regardless of whether a test makes a student's cognitive ability a biased or unbiased criterion, the test is likely to make biased predictions. For example, an academic ability test may routinely overestimate men's future academic performance and underestimate women's. In this case, if this test is used to make academic choice or placement decisions, women will be disadvantaged.

Differences in tests are examined on the basis of item and test. These studies are mostly carried out by Differential Item Functioning (DIF) analysis. It is common to make DIF analyzes based on Item Response Theory (IRT) (Murk & Davidshofer, 2005).

The aim of this study is to investigate whether the items in the Arithmetic Ability Test for High School Grades 11 contain DIF by using Rasch model within the frame of item response theory (IRT).

At this thesis study, the differentiation in quantitative ability between sexes is studied. For this, first of all, a test is formed that is totally 40 items each of which has sub-tests composed of 10 items. The test was applied to 400 sample students that were 11th class in 2017-2018 educational year. After the application item selection was made by accounting item difficulties and item discrimination parameters. Studies for reliability and validity were made, 5 items from each that best measure the subject structures (arithmetical reasoning, forming equality, geometry, visual-spatial ability) were chosen and the final form of the test was decided. After that the test was applied to 1382 students who attend 11th class in 2018-2019 at Bornova Anadolu, Cihat Kora Anadolu, 15 Temmuz Şehitler Anadolu, Atatürk Lisesi, Karşıyaka Lisesi and Atakent Anadolu high schools. After making necessary eliminations, series of analysis were made for a sample of 1036 students.

When the researches on this subject are examined, it is concluded that boys are more successful than the girls in the Numerical Ability tests. In this study, contrary to many previous studies, instead of using ready data such as PISA, TIMMS, OKS, SBS data, a quantitative ability test was developed. After the validity and reliability studies were carried out and item parameters were calculated, the test was applied to 11th grade high school students.

The problem sentence presented in this study is as follows:

Is there any differentiation in quantitative ability across the gender groups?

The sub-problems examined within the framework of the problem sentence are as follows:

1. What are the psychometric properties of the Quantitative Ability Test (QAT) developed within the scope of this research?
2. Does QAT meet the assumptions of the Rasch Model?
3. Do QAT items contain differential item function (DIF) that varies across the gender groups?

2.Method

The population of the study consists of students who are studying in İzmir and are in the 11th grade as of the 2018-2019 academic year. The sample selected for this study consisted of 11th grade students attending 1036 State Anatolian High Schools. While selecting the sample, students in Anatolian High Schools with a minimum percentage of 8% settling in Anatolian High Schools were preferred. The students who participated in the study answered 20 items in the quantitative ability test consisting of Arithmetic Reasoning, Equation, Geometry and Visual-Spatial Ability subtests. As a result, the data of the students' responses to this test were evaluated.

The Quantitative Ability Test consists of four subtests and each subtest contains 5 items. The test consisted of 20 items and the students were given 40 minutes as the response time. The answers were evaluated as true false (1-0).

The data were obtained from the answers given by the 11th grade students of some state Anatolian High Schools in Izmir to the Quantitative Ability Test.

Before the actual application of the Quantitative Ability Test, a total of 40 items consisting of 10 items were created for each subtest. After this application, item difficulties and discrimination were calculated and item selection was made. The reliability and validity studies were conducted and the five items that best measure the quantitative ability (Arithmetic Reasoning, Equation, Geometry and Visual-Spatial Ability) that we want to measure were selected and the test was finalized.

In the first sub-sentence of the study, the psychometric properties of the Quantitative Ability Test were applied to the sample consisting of 400 students of 11th grade students. In order to finalize the test, after the data obtained as a result of the application, item difficulties and discrimination were calculated and item selection was made. Reliability and validity studies were conducted and 5 items that best measure the structures we want to measure were selected and the test was finalized. The final form of the test was applied to 1382 11th grade students. The data obtained were analyzed by using WINSTEPS program according to Rasch model within the framework of item response theory. In order to examine the second sub-problem of the research, it was investigated whether the data set meets the assumptions of Rasch model. In the investigation of the third sub-problem, DIF-containing substances were identified.

3. Results

The Test1 application includes the Equality (top 10 questions) and Numerical Reasoning (11-20 questions) subtests. In Test 2 application, Visual-Spatial Ability (top 10 questions) and Geometry (Questions 11-20) tests are available. Item analyzes of Test 1 and Test 2 were performed in the exel program in order to determine the items that could be put directly into the test and to eliminate the substances that could not be tested. According to this; of the 20 items in Test1; item difficulties were selected from 0.35 to 0.77, item discrimination between 0.36 and 0.65 and item standard deviations between 0.42 and 0.50 were selected (2, 3, 4, 6, 7, 11, 17, 18, 19, and 20. items).

Of the 20 items in Test2, which consisted of Geometry and Visuospatial Ability Questions, 10 items were selected, ranging from 0.27 to 0.84, difficulty in distinguishing between 0.31 and 0.51, and standard deviations between 0.36 and 0.50 (1, 2, 5, 9), 10, 11, 13, 16, 18, and 20). When reliability and validity studies were performed, the Kr-20 reliability coefficient of Test1 was 0.653 and the reliability coefficient of Test2 was 0.597. As a result, 5 items in each subtest that best measures the structures we want to measure (Arithmetic Reasoning, Equation, Geometry and Visual-Spatial Ability) were selected and the test was finalized. The reliability coefficient of the final test was found to be 0.61, item difficulty average was 0.411 and item discrimination average was 0.286. The observed correlation $>$ Expected correlation has high discrimination power for high and low gifted students. According to the analysis, 15, 5, 16, 6, 7, 20.8., 19., 9. And 10th items have high discrimination power.

The observed correlation $<$ expected correlation has low discrimination power for high and low gifted students. At the end of the analyzes, 17th, 4th, 14th, 12th, 13th, 11th, 1st, 2nd. The items were found to have low discrimination power.

When we look at the INFIT and OUTFIT meansquare statistics, the most incompatible item is item 17 (infit mnsq = 1.37 and outfit mnsq = 1.65). When the infit and outfit mnsq values of all other items in the scale were examined, it was seen that all of them were close to 1. In other words, the items were in good fit with the model.

Looking at the expected item characteristic curves of item 9 and item 10, it is seen that although item fit statistics (outfit mnsq values) are close to 1, these items are easier for students with low talent than expected and are more difficult for high-skilled students than expected.

When we look at the INFIT and OUTFIT fit statistics for the students, it is seen that it is 1.00 and .99 respectively. The fact that these statistics are close to 1 indicates perfect fit. Standard deviation model expectation is reported as .95. Standard deviation of infit mnsq and outfit mnsq values are .17 and .50, respectively.

When we look at the INFIT and OUTFIT statistics for the whole test, it is seen that 1.00 and .99 respectively. Standard deviation model expectation is reported as .86

(close to 1). Standard deviation of infit mnsq and outfit mnsq values are .11 and .20, respectively.

If we want to find out if the item is biased in favor of or against a group, we look at the value of the DIF score. If the DIF score value is positive, item bias is in favor of that group, if negative, item bias is against that group ($p < .05$).

When we look at the DIF score values, the items with positive value for men are 1,2,7,9,10,12 and 16th items. These items appear to be biased in favor of men.

For males, items with a negative DIF score are 3,4,5,6,11,15,17,18,19,20. these items appear to be biased against men. For more information, check the probability (p) values. According to these values, the results that test the hypothesis H_0 "This item does not contain DIF statistically" are evaluated.

The hypothesis H_0 was rejected for items 7,9,15,17 and 20 ($p < .05$). Items 7 and 9 show bias in favor of boys, while items 15, 17 and 20 show bias in favor of girls. However, this H_0 hypothesis was accepted for other items ($p > .05$).

When the unidimensionality of the quantitative ability test is examined, the ratio of the raw variance eigenvalue explained by the model to the unexplained raw variance eigenvalue in 1. Contrast is found as $7.0636 / 2.0705 = 3.4115$. Since this ratio is greater than 3, it shows that the test measures a single implicit feature, namely quantitative ability. It was concluded that the Quantitative Ability Test was unidimensional.

In order to evaluate the local independence assumption, the highest inter-item standard residual correlations are examined. It is said that if the correlation between the items is less than $|.30|$, the assumption of local independence is provided (Bruin, 2012). Only the residual correlation between items 9-10 was found to be 0.55. Since this value is greater than $|.30|$, this value violates the assumption of local independence. Since residual correlations between other items are lower than $|.30|$, local independence assumption is provided for these items. Since the correlation is greater than the value $|.30|$ between the 9th and 10th items, the assumption of local independence of the test is met (Bruin, 2012).

4. Discussion and Conclusion

Accordingly, the Arithmetic reasoning and equality of the sub-test consisting of 20 items in Test1' de 20 items; item difficulties were selected from 0.35 to 0.77, item discrimination between 0.36 and 0.65 and item standard deviations between 0.42 and 0.50. Of the 20 items in Test2, which consisted of Geometry and Visuospatial Ability Questions, 10 items were selected with item difficulty ranging from 0.27 to 0.84, discrimination between 0.31 and 0.51, and item standard deviations between 0.36 and 0.50. When reliability and validity studies were performed, the Kr-20 reliability coefficient of Test1 was 0.653 and the reliability coefficient of Test2 was 0.597. As a result, 5 items in each subtest that best measures the structures we want to measure (Arithmetic Reasoning, Equation, Geometry and Visual-Spatial Ability) were selected and the test was finalized. After finalizing the test, the 20-item quantitative ability test was applied to 1382 students in the 11th grade of Bornova Anatolian High School, Cihat Kora Anatolian High School, 15 Temmuz Şehitler Anatolian High School, Atatürk High School, Karşıyaka High School and Atakent Anatolian High School. After screening, 1036 students were analyzed and reported. In the Rasch analysis, the outfit mnsq value was 0.99 and the infit mnsq value was 1.00, which shows that the person-model fit was excellent.

The reliability coefficient of the final test was found to be 0.61, item difficulty average was 0.411 and item discrimination average was 0.286. Although the reliability coefficient seems to be low, in the Rasch analysis, test reliability was found to be 0.99 in 1036 students. When the sample size increased, the overall reliability of the test increased.

As a result of Rasch analysis, data-model fit, person-model fit, test model fit, DIF research, local independence and unidimensionality assumptions were checked. In Özalp Ateş (2015) 's study of the construct validity of the scales, Rasch Analysis suggested that the internal construct validity was also evaluated in terms of different characteristics such as item fit, person fit and item bias compared to factor analysis methods examining only the dimension structure of the scale (Özalp Ateş , 2015). The findings in this study were based solely on Rasch analysis and other methods such as factor analysis were not used during the test development stage.

According to the findings 15th, 5th, 16th, 6th, 7th, 20th, 8th, 19th, 9th. and 10) have high discrimination power. 17, 4, 14, 12, 13, 11, 1, 2. items have low discrimination power. When the INFIT and OUTFIT meansquare statistics are examined, it is seen that the most incompatible item is item 17 (infit mnsq = 1.37 and outfit mnsq = 1.65). It is recommended to remove this item from the scale. When the infit and outfit mnsq values of all other items in the scale were examined, it was seen that all of them were close to 1. In other words, the items were in good fit with the model. In addition, the expected and observed matching percentages were examined to examine the data model fit. When these percentages are equal, the data is said to be fit with the model. Except item 17., the expected and observed matching percentages of other items are very close to each other or equal. Looking at the expected item characteristic curves of item 9 and item 10, it is seen that although item fit statistics (outfit mnsq values) are close to 1, these items are easier for students with low talent than expected and are more difficult for high-skilled students than expected. In other words, only the expected and observed percentages of matching are insufficient to comment on the data-model fit. For healthier decisions, item characteristic curves should also be examined.

Looking at the INFIT and OUTFIT statistics for the overall test, it was found to be 1.00 and .99, respectively, and the standard deviation model expectation was reported as .86 (close to 1). The standard deviation infit mnsq and outfit mnsq values were .11 and .20, respectively, suggesting good test-data fit.

When the unidimensional feature of the quantitative ability test was examined, the ratio of the raw variance eigenvalue explained by the model to the unexplained raw variance eigenvalue in the 1st Contrast was found to be $7.0636 / 2.0705 = 3.4115$. Since this ratio is greater than 3, it shows that the test measures a single implicit feature, namely quantitative ability. It was concluded that the Quantitative Ability Test was unidimensional.

In order to evaluate the assumption of local independence, the correlation of residuals between items was examined and only the residual correlation between items 9-10 was found to be 0.55. This can be avoided by eliminating pairs (9-10) of the items

that violate the local independence assumption or by combining these items and testing them as a single item.

Finally, it should be ensured that DIF studies are routinely performed for all national exams within the scope of validity studies of tests.

Key words: Quantitative Ability Test(QAT), Rasch Model, Differential Item Function(DIF)



BÖLÜM I

GİRİŞ

1.1 Problem Durumu

Kadınların, daha güçlü sözel becerileri ve olaylar, kelimeler, nesnelere, yüzler ve etkinlikler için daha iyi hafızaları olduğu düşüncesi yaygındır. Erkeklerin de genellikle nesnelere zihinsel olarak manipüle etmede ve görsel sunumlara dayanan sayısal görevlerde başarılı oldukları düşünülür. Notlar ve genel test puanları pek çok faktöre bağlı olduğu için, psikologlar bu cinsiyet farklılıklarını anlamak için daha iyi tanımlanmış bilişsel becerileri değerlendirmeye yönelmişlerdir. Okul öncesi çocuklar ne daha fazla ne de daha az sayısal yetenekle doğmuşlardır, çünkü kızlar ve erkekler, sayısal düşünme ve çevredeki nesnelere bilgisi ile ilgili erken bilişsel becerilerde eşit derecede iyi performans gösterirler. Bununla birlikte, okul başladığı zaman bu beceriler cinsiyetlerde farklılaşmaya başlar. Eğitim-Öğretim dönemi boyunca ve okulun sonuna gelindiğinde kızlar, sözel yeteneklerin değerlendirilmesinde daha iyi performans gösterirler. Bununla birlikte, erkeklerin üst seviyede olduğu spesifik bir başka yetenek türü vardır: visuospatial olarak adlandırılan beceri seti: Zihinsel olarak gezinme ve nesnelere üç boyutlu hareketi modelleme becerisi. Dört ve beş yaş arasındaki erkek çocuklar, standartlaştırılmış testlerde labirentleri çözmeye ölçülebilir ölçüde daha iyidir. Erkeklerin başarı gösterdiği visuospatial becerinin bir başka belirtisi de "zihinsel rotasyon" a aittir ve üç boyutlu bir nesneyi belleğe alırken eşzamanlı olarak onu dönüştürür. Beklediği gibi, bu yetenekler aynı zamanda erkeklere zihinsel bir görüntü yaratmaya dayalı matematik problemlerini çözmeye bir avantaj sağlar (Perry, 2015).

Erkekler bilim, mühendislik ve matematik alanlarına neden hakim oluyor? (Perry et al., 2015) (By Diane F. Halpern, Camilla P. Benbow, David C. Geary, Ruben C. Gur, Janet Shibley Hyde and Morton Ann Gernsbacher from Scientific American Mind, November 2007). Amerika Birleşik Devletleri (ABD) ve diğer ülkelerde, erkeklerin kızlara göre matematikte daha iyi performans sergilediğine dair yaygın bir inanç vardır. Bununla birlikte, Amerika Birleşik Devletleri'nin verileri bu inancı desteklememektedir (Chipman & Thomas, 1985; Hyde, Fennema, & Lamon, 1990).

ABD öğrenci popülasyonlarıyla ilgili geniş temsilli çalışmalar, orta öğretim yıllarından önce matematiğin genellikle ABD'de isteğe bağlı hale geldiğinden genel matematik performansında cinsiyet farkı az ya da hiç bulunmamaktadır. İlkokuldaki erkeklerin veya kızların bir ya da daha fazla sayıdaki matematik testinde daha iyi performans gösterdiklerini gösteren ara raporlara rağmen, bir meta-analiz, ortaöğretimden önce böyle bir farklılığın olmadığı sonucuna varmıştır (Hyde ve ark. 1990). Bununla birlikte, ortaokul sonuna kadar, erkeklerin lehine olan matematik testi performansındaki cinsiyet farklılıkları genellikle Amerika Birleşik Devletleri'nde rapor edilmiştir ve performans farklılıkları problem çözme testlerinden veya maddelerden kaynaklanmaktadır (Hyde ve diğerleri, 1990). Bu durum, matematiksel performansın farklı ölçümlerinin cinsiyet farklılıklarına ilişkin farklı mesajlar üretmesinden dolayı daha da karmaşıktır (Gallagher & Kaufman, 2000).

Toplumumuz cinsiyetçi, ırkçı veya diğer herhangi bir şekilde önyargılı olduğu sürece, grup farklılıkları bulma iddiası, er ya da geç, daha güçlü bir grubun üstünlüğünün bir kanıtı olarak kabul edilir (Eccles & Jacobs, 1986). Toplumsal / motivasyonel faktörler hakkındaki araştırmalar incelenirken, matematik performansını etkileyen çok şey olduğu (biyolojik farklılıklar gibi), hormonların erkeklerin ve dişilerin gelişen bilişsel yeteneklerini belirlemede rol oynayabileceğini göstermektedir (Perry et al., 2015) ve kızların yaşadıkları çok sayıdaki engelin matematiksel başarılarını olumsuz yönde etkilediği görülmüştür. Bütün engeller atıldığında kızların matematik performansının erkeklerinkinden nadiren daha düşük olduğunu akılda tutmak önemlidir.

Byrnes ve Takahira (1993), önceki bilgi ve stratejilerin bir sonucu olarak Scholastic Aptitude Test-Mathematics (SAT-M) performansındaki cinsiyet farklılıklarını açıkladılar. Pajares ve Miller (1994), bireylerin cinsiyete bağlı olarak matematik problemi çözme konusunda farklı tecrübeleri olduğunu buldular (Perry et al., 2015).

Erkekler en az otuz yıldır SAT-M'de kadınlardan daha iyi performans göstermektedirler. Aradaki fark genel olarak yaklaşık 40 puan ($d = 0.39$) aralığındadır ve en son 2000'de 33 puana düşmüştür. American College Test Mathematics (ACTM)'deki performans benzer bir eğilim göstermiştir; 1970'deki 3.1 puanlık fark,

2001'de 1.2 puana kadar gerilemiştir ($d = 0.24$) (Langenfield, 1997; ABD Eğitim Bakanlığı,2001). Bu durum özellikle önemlidir, çünkü genellikle üniversiteye gitmek isteyen öğrenciler üst düzey öğrencilerdir ve SAT veya ACT bir üniversite giriş şartı olarak kabul edilir (Perry, 2015).

Sayısal yetenekte cinsiyet farklılaşmasını incelemeye önce zekâ tanımları ve zekâ kuramları hakkında aşağıda genel bilgiler verilmiştir.

1.1.1. Zekânın Tanımı

Zekâ, diğer pek çok psikolojik değişken gibi, doğrudan gözlenemeyen çok karmaşık yapılardan (construct), hatta en önemlilerinden biridir. Zekâ geçmişten günümüze çok farklı şekillerde tanımlanmıştır. Bu tanımlardan bazıları aşağıda Tablo 1'de verilmiştir (Toker, 1968).

Tablo 1: Zekâ Tanımları

| | |
|----------------|--|
| A.Binet | Zekâ, iyi akıl yürütme, iyi hüküm verme ve kendi kendini eleştirme kapasitesidir. |
| E.L.Thorndike | Zekâ, gerçek ya da olgular açısından iyi tepkilerde bulunabilme yeteneğidir. |
| W.Stern | Zekâ, bireyin düşüncesini yeni durumlara bilinçli olarak uydurma yeteneğidir. |
| L.M. Terman | Birey soyut düşünebildiği ölçüde zekidir. |
| D.Wechsler | İşevruk bir tanımla zeka, bireyin gayeli davranma, mantıklı düşünme ve çevresiyle ilişkilerinde etkili olma kapasitesinin tümüdür. |
| J.M.Baldwin | Zekâ, bilme yeti ya da kapasitesidir. |

| | |
|--------------|---|
| J .Piaget | Zekâ organizma ile çevresi arasındaki kendine uydurma(assimilatory) ve kendini uydurma(accomodatory) ile ilgili etkileşimlerin olduğu kadar tüm duyuşal-hareketsel ve bilişsel nitelikteki ardışık uyumların yöneldiği denge durumunu kurar. |
| G.D.Stoodord | Zekâ, zor karmaşık, soyut, ekonomik, amaca uygun, sosyal değer taşıyan, orijinal etkinliklerde bulunma ve bu çeşitten etkinlikleri, enerjiyi bir noktada toplamayı ve heyecansal güçlere dayatmayı gerekli kılan durumlarda sürdürme yeteneğidir. |
| F.S.Freeman | Zekâ, yaşantıları bütünleştirme ve yeni durumlara uyumu sağlayan tepkilerde bulunma kapasitesidir. |
| H.E.Garrett | Zekâ, anlamayı ve sembolleri kullanmayı gerektiren problemlerin çözümünde ihtiyaç duyulan yetenekleri kapsar. |
| N.L.Munn | Zekâ, uyumun esnekliği ya da çok taraflılığı olarak tanımlayabileceğimiz bir fonksiyondur. |

Zekâ kavramı üzerine çeşitli kuramlar da üretilmiştir.

1.1.2. Çağımızın Zekâ Kuramlarından Bazıları

1.İki Faktör Kuramı:

Bu kuramı ortaya atan ve geliştiren Charles E.Spearman dır (1863—1945). Henüz zekâ testleri mevcut olmadığından Spearman, öğrencilerin genel zekâ seviyelerini, öğretmen kanaatlerine ve öğrencilerin birbirlerini bu yönden değerlendirmelerine dayanarak saptamıştır. Yukarıda adı geçen değişkenler arasında pozitif ve yüksek korelasyonlar elde eden Spearman, bu sonucu, bütün zihin etkinliklerinde rol oynayan genel bir zekânın varlığına bağlamıştır (Toker, 1968).

Spearman, değişik zihni yetenekleri ölçtüğü kabul edilen testlerin birbirleriyle olan korelasyonlarına faktör analizi tekniğini uygulamış ve sonuç olarak bu testlerin

ölçtüğü birbirinden farklı zihin güçlerinin tek bir yanı olduğu kanısına varmıştır (Toker, 1968).

Bunun üzerine, her türlü zihin etkinliğinde rol oynayan genel bir zihnî enerjinin var olduğunu ileri sürmüş ve buna “g” adını vermiştir. Spearman, farklı zihin yeteneklerini ölçen testler arasındaki korelasyonların mükemmel olmayışını zihinde özel faktörlerin varlığına bağlamış, bu özel faktörlere de ‘s’ demiştir. Özel faktör (s), belirli bir zihnî etkinliğin gösterilebilmesi için genel zihni yeteneğin (g) dışında ihtiyaç duyulan zihin gücüdür. Ayrıca Spearman, her zihnî etkinlik için gerekli olan genel (g) ve özel (s) yetenek miktarının farklı olduğunu ileri sürmüştür (Toker, 1968).

Spearman 'a göre, bireyler sahip oldukları genel zihnî yetenek (g) yönünden birbirlerinden farklıdır. Zekâyı ölçmek demek, “g” yi ölçmek demektir. Genel zihni yetenekle daha çok yüklü bulunan etkinlikler:

- a) Kendi yaşantılarını anlama
- b) İlişkileri bulma
- c) Benzer ilişkileri bulma gibi etkinliklerdir (Toker, 1968).

2. Çok Faktör Kuramı:

A. Thorndike'in Çok Faktör Kuramı:

Thorndike'in zekâ görüşü atomistik bir görüştür. Ona göre zekâ birbirinden ayrı faktörlerden meydana gelir. Faktörler birbirlerinden bağımsızdır. Bu duruma göre genel bir zekânın sözü edilemez. Zekâ değil, zekâlar vardır. Bir zihni problemin çözümünde birden fazla faktör rol alır. Ölçme gibi pratik amaçlarla, birbirine benzeyen zihni işlerde çoğu birlikte çalışan faktörleri gruplamak olabilir. Böylece, zihnin ayrı güçleri (faktörleri) olduğu, faktörlerin ortak özelliklerine göre gruplanabileceği görüşünden hareket eden Thorndike, çalışmaları sonucu aşağıdaki grupları bulmuştur:

- a) Kelime anlamı
- b) Aritmetik akıl yürütme
- c) Kavrama
- d) İlişkileri görsel algılama

Thorndike, zekâyı, soyut zekâ, sosyal zekâ ve mekanik zekâ olmak üzere üçe ayırmıştır. Ona göre soyut zekâ, sayı ve kelime cinsinden sembolleri, mekanik zekâ, çeşitli araç gereç ve makinaları anlama ve kullanma yeteneğidir. Sosyal zekâ ise

insanları anlama ve onlarla başarılı ilişkiler kurabilme yeteneği olarak tanımlanabilir (Toker, 1968).

B. J.P.Guilford 'un Çok Faktör Kuramı:

J. P.Guilford, faktör analizi yolu ile birbirinden kesin şeklide bağımsız zihinsel faktörler saptamış, “Zihnî Yapısı” adını verdiği zekâ kuramını 1959 yılında yayınlamıştır. Ona göre, zekânın tabiatı ancak öğelerinin bilinmesi ve bunların bir sistem içinde düşünülmesi ile anlaşılabilir. "Zihnî Yapısı" adlı kuramının dayandığı başlıca sayılılar şunlardır:

- Zihin birbirinden bağımsız faktörlerden meydana gelir. Faktörler belli yönlerden birbirlerine benzedikleri için sınıflandırılabilir. Sınıflandırma, faktörlerin ayrılığı ilkesine aykırı düşmez.
- Birey, her zihinsel etkinlik alanında aynı ölçüde yeteneğe sahip olmayabilir. Belli bir işte üstün başarı sağlayan bir kimse diğer bir işte aynı ölçüde başarı gösteremeyebilir.
- Her zihinsel etkinliğin "muhteva", "işlem" ve "ürün" olmak üzere üç yönü vardır. Bu boyutlardan sadece birinin olmaması halinde dahi zihinsel etkinliğin varlığı düşünülemez. Muhteva olmadan zihinsel işlem yapmak, zihinsel işlem yapmadan zihinsel ürün elde etmek mümkün değildir.

Zihnî faktörlerden her biri belli bir muhtevayı belli bir işlemden sonra belli bir ürün haline getiren zihin yeteneğidir. Guilford'a göre insan zihninde birinden ayrı 120 faktör vardır (Toker, 1968).

Akışkan ve Kristalize Zekâ: Cattell (1963), genel zekâ ile ilişkili fakat kavramsal bakımdan birbirinden ayrı iki yönü öne sürmüştür; Akışkan ve kristalize zekâ. Akışkan zekâ , analogilerde ve harflerdeki ve sayı serilerindeki ilişkileri görme kabiliyeti olarak tanımlanır; akıcı zekâ öncelikle akıl yürütme yeteneğini ele alır. Kristalize zekâ, diğer taraftan, bir bireyin edindiği bilgi ve becerileri ifade eder. Bir anlamda, kristalize zekâ, kişinin gerçek bilgi deposunun büyüklüğü anlamına gelir. Akıcı zeka alanındaki kişilerin hızlı bir şekilde olgusal bilgi deposu kurmalarını kolaylaştırdığı gerçeğini yansıtan, akışkan ve kristalize zeka ölçümleri arasında yüksek korelasyon eğilimi vardır (Murphy & Davidshofer, 2005).

C. Grup Faktör Kuramı:

Louis L. Thurstone (1887-1955) çok sayıda testin sonuçlarına, gelişmesine büyük ölçüde katkıda bulunduğu faktör analizi tekliğini uygulayarak grup faktör kuramı diye adlandırılan zekâ kuramını ortaya koymuştur. Bu kurama göre zihni etkinliği gerekli kılan işler gruplandırılabilir. Belli bir grupta yer alan işler belli ve diğerlerinden ayrı bir zihin gücünü gerektirir. Gruplardan her biri için gerekli olan zihnin gücüne temel faktör veya yetenek adı verilmiştir.

Thurstone çalışmalarında 12 faktör bulmuşsa da bunlardan ancak 7 tanesini adlandırabilmiştir. Bu 7 temel faktör şunlardır:

- Sayısal: Sayısal işlemleri doğru ve çabuk yapabilme yeteneği.
- Sözel: Kelimeler arasındaki ilişkileri görebilme yeteneği.
- Yersel: Bir cismin uzaydaki çeşitli durumlarını göz önünde canlandırabilme yeteneği.
- Kelime akıcılığı: Belli bir süre içinde çok kelime söyleyebilme yeteneği.
- Akıl yürütme: Kural veya ilkeyi bulabilme yeteneği. Burada tümevarımsal ve tümdengelimsel düşünme kastedilmekle beraber tümevarımsal düşünme daha önemli sayılmıştır.
- Anlamsız belleme: Mümkün olduğu kadar çabuk belleme yeteneği.
- Algısal: Çeşitli şekiller arasındaki ince farkları algılayabilme yeteneği.

Thurstone, bulmuş olduğu yetenekleri ölçmek amacı ile Temel Yetenekler Testi (Primary Mental Abilities Tests-P.M.A) adını verdiği bir test geliştirmiştir. Bu test her temel yetenek için ayrı bir puan verir (Toker, 1968).

D. Piaget'in Zekâ Kuramı:

Jean Piaget (1896), zekâyı uyum süreci olarak görür. Uyum, organizma-çevre etkileşiminde denge demektir. Her etkinliğin amacı dengeye ulaşmaktır. Organizma ile çevresi arasında hiçbir zaman tam ve sürekli bir denge kurulamazsa da tüm çabalar daima ona dönüktür.

Her dengesizlik halinde birey, dengeyi yeniden kurmak için etkinlikte bulunur. Uyumu sağlayan etkinlikler, en basit duyuusal-hareketsel tepkilerden en üst seviyedeki düşünme sürecine kadar uzanan bir devamlılık gösterir. Zekâ, akıl yürütmeyi gerekli kılan karmaşık durumlarda, algı ve alışkanlık gibi esnek olmayan tek yönlü etkinliklerle duruma uymanın olduğu hallere kıyasla kendini daha çok gösterir (Toker, 1968).

E. Vernon'un Hiyerarşik Modeli:

Vernon'un modeli, zekânın tüm testlerine etki eden ve büyük grup faktörleri olarak adlandırılan iki geniş kategoriye ayrılabilen genel bir faktörü göstermektedir. Sözel ve mekansal faktörler gibi daha küçük grup faktörlerine bölünebilen, sözel-egitimsel ve mekansal-motor olmak üzere iki ana grup faktörü vardır. Bazı açılardan, bu küçük grup faktörleri birincil zihinsel yeteneklere benzemektedir. Küçük grup faktörleri, bireysel testler arasındaki ilişkinin kalıplarını özetler. Zekâ hiyerarşik modelleri, g'yi ölçmek üzere tasarlanan testlerin ve zekanın daha spesifik yönlerini ölçmek için tasarlanmış testlerin yerini belirtir. Hiyerarşideki belirli bir seviyenin seçimi, kişisel tercih yerine test amacı ile belirlenmelidir (Murphy & Davidshofer, 2005, s.25).

Tüm bu zekâ tanımlarına ve zekâ kuramlarına baktığımızda çoğunun ortak özelliği olarak zekânın ölçülmesinin tek bir yolu olmadığı ve zekâ ölçmenin aslında yetenekleri ölçmek anlamına geldiğidir.

1.1.3 Zekâ Testlerinin Tarihsel Temelleri

"Zihinsel Test" terimi 1890 yılına kadar (modern anlamda) görünmemekteydi. Genel zekânın standartlaştırılmış ve oldukça geçerli bir ölçümünü sağlayan ilk test 1908 yılına kadar görülmemiştir. Genel zekâyı ölçmek için tasarlanan test teorilerinin erken gelişimi, esasen 1900'lerin başında İngiltere ve Fransa'da toplanmıştır. İngiltere'de, Galton'un öncü çalışması bireysel farklılıkların sistematik olarak incelenmesi için gerekli ölçümleri ve araçları sağlamıştır. Fransa'da deli ve zihinsel engellilere yönelik tutumların değişmesi bireysel farklılıkların ölçümünün yapılmasına yardımcı olmuştur. Galton ve öğrencileri, özellikle Pearson, bugün bireysel farklılıkların analizinde kullanılan temel prosedürlerden birkaçını geliştirdi. James

McKeen Cattell, bireysel farklılıkların incelenmesini Birleşik Devletlere getirmeye yardımcı oldu. Cattell ve Wundt, Galton'dan etkilenmiş ve üniversite öğrencilerinin entelektüel seviyesinin ölçülmesinde kullanılan bir dizi sensorimotor test geliştirmiştir. Cattell' in testleri Galton'a benzemekle birlikte kuvvet, tepki süresi, ağrı duyarlılığı ve kilo ayrımcılığı ölçülerini içeriyordu. Bu prosedürleri açıklayan bir makalede Cattell (1890) "zihinsel test" terimini ilk kullanan kişi oldu (Murphy & Davidshofer, 2005, s.29-30-31).

1890'lar, sensorimotor tipinde zihinsel testlerin geçiciliğini gördüler. Galton ve Cattell tarafından kullanılan sensorimotor testlerin son derece güvenilir olmadığını ve daha kötüsü, diğer ölçümler ile çok az korelasyon gösterdiğini gösteren bazı erken çalışmalar, en önemlisi Wissler (1901), bu zihinsel testlere ölümcül bir darbe indirmiştir. Bununla birlikte, Fransa'da zekâ ölçümünün pratik yöntemleri başarıyla geliştirilmiştir. Başka bir Fransız hekim Seguin'in çalışması, zekâ testlerinin geliştirilmesi için bir ivme kazandırmada özellikle önemlidir. Binet, bellek, anlama ve yargılama gibi işlevleri içeren daha karmaşık konularda örnekleme yaparak zekânın daha iyi ölçüleceğini ısrarla vurgulayan kişiydi (Binet & Henri, 1895). 1905'te Binet, Simon ile işbirliği içinde, 1905 Binet-Simon ölçeği (Binet & Simon, 1905) ile genel zekânın pratik ve geçerli bir ölçümünü sağlamak için ilk ölçeği geliştirmiştir. 1905 Binet-Simon ölçeği zorluk sırasına göre düzenlenmiş 30 madde veya test içermektedir. Bu teste dahil edilen birkaç psikomotor görev olsa da, çoğu madde dil, akıl yürütme veya okuduğunu anlama kullanımını gerektirmektedir. Binet-Simon ölçeğinin ilk revizyonu 1908'de ortaya çıkmıştır. 1908 ölçeğinde, daha önceki ölçekten memnun olunmayan birkaç madde çıkarılarak 1908 Binet-Simon ölçeğindeki toplam madde sayısı 58'e çıkarılmıştır. Binet'in testleri çocuklarda kullanılmak üzere geliştirilmişti. Yetişkinler için karşılaştırılabilir bir teste ihtiyaç duyulduğunu belirten Bellevue Hastanesi'nden bir psikiyatrist David Wechsler, Wechsler-Bellevue Zeka Testi'ni geliştirmiştir. Stanford Üniversitesi'nde Terman (1916) tarafından gerçekleştirilen etkili uygulamalardan biri, Stanford-Binet'in Intelligence Quotient(IQ) zekâ testinin, ilk 1908'deki bazı zekâ güçlükleri (IQ) ile ilgili olarak bir test geliştirmesiydi (Murphy & Davidshofer, 2005, s.30-31).

Bugün yaygın olarak kullanılan standartlaştırılmış, makine puanlaması kullanan zekâ testleri Binet'in testlerinden oldukça farklıdır. Stanford-Binet ve benzeri

testler, sınav uygulanan kişilerin zorlu görevlerle karşı karşıya kaldıkları karmaşık sosyal etkileşimleri temsil eder. Bu testler, özel eğitim görmüş psikologlar tarafından bireysel olarak uygulanmalıdır ve yönetim ve yorumlamaları için önemli özen ve profesyonel yargılama gerektirir. Birçok psikolog, bunlardan en önemlisi olan Otis, 1900'lerin başında standart koşullar altında öğrenci gruplarına uygulanabilecek daha basit testler geliştirmeye başladı. Bununla birlikte, grup testlerinin geliştirilmesi için gerçek itici güç, Amerika'nın I. Dünya Savaşı'na girişiyle geldi. Savaşın önce, Amerikan ordusu küçüktü. Savaşa girildiğinde aniden asker sayısının ve ordu büyüklüğünün karmaşıklığının artmasıyla birlikte emir komuta zincirinin zorlaştı. Yeni askerlerin sınıflandırılmasına ve yerleştirilmesine yönelik geleneksel yöntemler, özellikle iyi subay adaylarının belirlenmesi için yeterli olmadığından, askere alınan kişileri sınıflandırmanın basit ve ekonomik bir yöntemi gerekiyordu. Amerikan Psikolog Birliği, hizmet çalışmalarına gönüllü oldu ve Yerkes'in başında olan ve Otis ve Terman gibi psikologlar da dahil olmak üzere bir American Psychological Association (APA) komitesine, çok sayıda askerin entelektüel seviyesini ölçmek için pratik bir yöntem geliştirme görevi verildi. Bu çabalar, Ordu Alfa ve Ordu Beta olmak üzere iki test geliştirilmesine yol açtı. Ordu Alfa, büyük gruplardaki askere alınanlara ve genel zekânın kaba bir ölçümü olarak uygulanabilen yazılı bir testti. Ordu Beta ise, okuryazar olmayan, az da olsa İngilizce konuşan askerler için grup halinde uygulanabilen ve basit resimsel ve sözsüz talimatlar içeren bir testti. Ordu Alfa ve Ordu Beta'nın kesin değeri ve geçerliliği hakkında bugün bazı tartışmalar olmasına rağmen (Gould, 1981), testler o dönemde faydalı olarak algılanıyordu. Ordunun grup testlerinin görünürdeki başarısı, iş çevrelerinde ve eğitim ortamlarında fark edilmedi (Murphy & Davidshofer, 2005, s.31-32).

1.1.4. Sayısal Akıl Yürütme Yapılarının ve Değerlendirmelerinin Gelişimi

Sayısal akıl yürütmenin değerlendirilmesi, 20. yüzyılın başlarından itibaren bir ölçme hedefi olmasına rağmen (örneğin, Üniversite Giriş Sınav Kurulu, 1926; McCall, 1923; Monroe, DeVoss ve Kelly, 1917; Thorndike ve diğerleri, 1924; Yerkes, 1921), sayısal akıl yürütmenin içerik ya da müfredat olarak matematikten farklı bilişsel bir

süreç olarak sistematik olarak ele alınması çok da geç şekillenmeye başlamamıştır (Dwyer, Gallagher, Levin, & Morley, 2003).

1920'ler ve 1930'lar boyunca, Dewey'in görüşleri (1933), sayısal akıl yürütme analizine hakimdi. Dewey, akıl yürütme yeteneğini geliştirmeyi, Amerika'daki ilk ve orta öğretimin temel hedeflerinden biri haline getirme gereğini vurguladı. Dewey, sayısal akıl yürütmeyi, eğitimdeki mantık sürecinin bir parçası olarak değerlendirdi ve öğretim ve öğrenme süreci için akıl yürütme tavsiyelerine odaklandı.

E. L. Thorndike, cebir üzerine yaptığı araştırmalar yoluyla sayısal akıl yürütme psikolojisinde ilk öne çıkan kişilerden biriydi (Thorndike et al., 1924). Sayısal akıl yürütme biçiminde soru türlerinin ilk tipolojisini sundu. Sayısal akıl yürütmenin hesaplama gibi alt düzeydeki matematiksel becerilerden esasen farklı olduğunu açıkça ortaya koymak onun döneminde alışılmadık bir durum olduğundan. Thorndike, sayısal akıl yürütmenin ölçülmesinin geçerliliğine yönelik bazı tehditleri belirtmek için de öncülük etmiştir. Test yapanlara, bugün yapısal olmayan varyans dediğimiz şeylerden kaçınmak için çok sayıda öneride bulundu (Dwyer et al., 2003).

1940'lı yıllarda, çoğu yazar, eleştirel düşüncüyü, çoğunlukla matematiksel düşünme perspektifinden ele aldı, ancak sayısal akıl yürütme üzerine yoğunlaşmadı. Bu trendin önemli bir istisnası, sayısal akıl yürütmeyi bir problem çözme süreci olarak nitelendiren ve bu süreci açıkça öğretmeyi savunan Polya'dır (1957, 1962, 1965). Polya'nın problem çözme konusundaki ana teması ve nasıl öğretileceği, 1960'lı ve 1970'li yıllarda tekrar ele alındı (Dwyer et al., 2003).

Resnick ve Resnick (1977), sosyal sınıf boyutuna sahip üst düzey bilişsel süreçlere yönelik tutumda bir değişiklik olduğunu savunuyorlardı. Resnick ve Resnick, çağdaş düşüncede sayısal akıl yürütmede yeni olan şeyin müfredatı dahil olmadığını söylediler. Uzun yıllar elitlerin okullarında okutulan içerikler, daha sonra "kitlelere" yönelik müfredat içerisine dahil edildi. Aynı dönemde Greeno (1978; Greeno & Simon, 1988) gibi eğitim odaklı bilişsel psikologlar çeşitli matematiksel görev içeriğinin bilgi ve işlemlerinin birçok sınıftaki etkinlik ve görevlerin ön koşulları olduğunu göstermekle ilgilendiler. Sayısal akıl yürütmenin bilişsel temelini anlama konusunda teorik gelişmeler ortaya çıktı (örn., Newell, 1980; Newell & Simon, 1972; Schoenfeld, 1983, 1985, 2002). Sayısal akıl yürütmenin etki alanına ve özgüllüğüne vurgu yaparak (ör. Ball, 1991; Glaser, 1984; Glaser & Baxter, 2002), bu içerik ve

ampirik verileri sınıfta ve değerlendirme uygulamasına çevirmek için yoğun çaba sarf edildi (Dwyer et al., 2003).

John Carroll (1993) ise, erken çocukluktan erişkinliğe kadar olan bireylerdeki sayısal akıl yürütme hakkındaki kanıtları sunmuştur. Onun ana sonucuna göre yalnızca üç temel akıl yürütme yeteneği vardı: sıralı (tümdengelimli), endüktif ve sayısal. Bu sonuç, ilginç bir şekilde, çocukluk çağının 5 yaşı ile yetişkinliğe kadar olan süreç için geçerlidir. Çok küçük çocuklar için akıl yürütücü faktörler anlamlı bir şekilde ayrılamaz (Dwyer et al., 2003).

Böylece 1970'lerin başına kadar, bilişsel psikologlar, matematik eğitimcileri ve matematikçiler sayısal akıl yürütme sürecinin ve mantıksal işlemin gerçekleştiği matematiksel içerikle olan karmaşık etkileşimlerinin anlaşılmasının önemi üzerinde birleşmeye başlamışlardır.

Markman ve Gentner (2001), sayısal akıl yürütmenin doğası hakkında birçok tartışmanın merkezinde yer alan etki alanı özgülüğü ve etki alanı genelliği sorunlarının bir analizi de dahil olmak üzere, akıl yürütme sürecinin ayrıntılı bir incelemesini yaptılar. Sayısal akıl yürütme üzerine teori ve araştırmaların tarihi, özellikle sayısal akıl yürütmenin kesin kavramlaştırmalarında, bireysel farklılıkların değerlendirilmesi bağlamında az sayıda girişimi göstermektedir. NCTM (2000); Amerika Matematik Birliği (MAA) (Amerika Matematik Birliği [MAA], 2003; Sons, 1996); Amerikan Matematik Derneği (AMS) (Howe, 1998); ve American Mathematical Association of Two-Year Colleges (AMATYC-Amerikan Matematik Derneği İki Yıllık Kolejler, 1995), matematik eğitiminin amaçları hakkında yaptıkları açıklamalarda, sayısal mantık yeteneğini tüm lise ve üniversite öğrencilerinin geliştirebileceğini ve geliştirmelerini söylemişlerdir.

1.1.5. Sayısal Yetenek

Sayısal yetenek sayıları hızlı ve doğru bir şekilde ustaca kullanabilme becerisidir. Sayısal Yetenek matematik yeteneğinden çok tümevarımsal akıl yürütme, mekânsal-uzamsal yetenek ve iki faktörlü zekâ ile ilişkilidir. Elbette ki sayılar, matematik kuralları dışında kullanılamaz; bu nedenle, sayısal yetenek testindeki maddeler belirli bir minimum aritmetik beceri gerektirir (Kline, 2000, s.219).

Sayısal akıl yürütme; aritmetik, sayı anlamı, tündengelimli akıl yürütme, matematiksel okur yazarlık, sayısal okur yazarlık, problem çözme, bağlamsallaştırılmış matematik, matematiksel modelleme ve niceliksel akıl yürütme gibi birçok isim ve tanıma sahip karmaşık bir yapıdır (Mayes, Peterson, & Bonilla, 2013, s.2).

Akıl yürütme becerisi ve içeriği, matematik dünyasında değerlendirme amaçlı sayısal akıl yürütme yapısını tanımlamak için temel oluşturur. Ölçme teorisi ve uygulamasında, Messick ve diğerleri tarafından geçerlik teorisinde kavramsal olarak yapılan paralel çalışma, akıl yürütme testlerinin, akıl yürütme yapısının unsurları arasında ayırım yapmak için yorumlanmasının kritik olduğunu anlamamızı sağlar. Bu, değerlendirmenin hedefi ve testi uygulayanların sınava getirdiği varsayılmış olan içerik bilgisinin ortak özüdür (Dwyer et al., 2003).

Sayısal mantık da, bir çözüm bulabilmek için belirli bir probleme hangi becerilerin ve prosedürlerin uygulanabileceğinin belirlenmesi de dahil olmak üzere nicel bilgiyi analiz etme kabiliyeti olarak tanımlanır (Dwyer et al., 2003).

Sayısal akıl yürütme, sayısal bilgileri analiz etme yeteneği olarak tanımlanır. Bu nedenle, matematik derslerinde edinilen becerilerle sınırlı değil, bütçeleme ve alışveriş gibi gündelik faaliyetlerin yanı sıra, neredeyse tüm lise veya üniversite derslerinde uygulama yoluyla zamanla geliştirilen akıl yürütme becerilerini de içerir.

Sayısal akıl yürütme, aşağıdaki altı yeteneği içerir:

- Tablolarda, geometrik şekillerde, matematik formüllerinde veya metin halinde (örneğin, gerçek yaşam problemlerinde) çeşitli biçimlerde verilen bilgileri okur ve anlar;
- Nicel bilgiyi yorumlamak ve ondan uygun çıkarımlar yapmak;
- Problem çözme, aritmetik, cebirsel, geometrik veya istatistiksel yöntemler kullanma;
- Cevapları tahmin etmek ve makul olup olmadığını anlamak için cevapları kontrol etmek;
- Nicel bilgileri sözlü, sayısal, cebirsel veya grafiksel olarak ilişkilendirmek;
- Matematiksel veya istatistiksel yöntemlerin sınırlamalarını tanımak.

Bu yetenekler, tüm üniversite mezunlarının sahip olması gereken yetkinliklerdir (Dwyer et al., 2003).

Sayısal düşünme yeteneği, demokratik bir vatandaş olmak için çok önemlidir, çünkü evde, işyerinde ve yerel topluluklarını etkileyen karmaşık ulusal ve uluslararası konularda bilinçli kararlar almalarını sağlar. Sayısal akıl yürütmenin tanımları ve sayısal okuryazarlık ile ilgili kavramlar Tablo 2'de verilmektedir (Dwyer et al., 2003).

Tablo 2: *Sayısal Akıl Yürütmenin Tanımları*

| Sayısal Akıl Yürütmenin Tanımları | |
|--|---|
| Steen (2004) | Sayısal okur yazarlık, sofistike akıl yürütmeye dayanan temel matematikten çok, temel matematikle sofistike akıl yürütmeyi içerir. |
| Langkamp and Hull (2007) | Çevresel sorunlar, sayıların anlamını, temel cebiri, basit modelleri ve istatistiksel tanımları kullanılarak daha iyi anlaşılabilir. sayısal akıl yürütme, temel matematiksel kavramları ve teknikleri sofistike şekilde kullanmayı gerektirir. |
| International Life Skills Survey (ILSS 2000) | Sayısal okur yazarlık, insanların yaşam ve iş hayatında ortaya çıkan sayısal durumlarda etkili bir şekilde yer alabilmek için ihtiyaç duyduğu beceri, bilgi, inanç, eğilim, alışkanlık, iletişim yetenekleri ve problem çözme becerilerinin toplamıdır. |

Programme for International Student Assessment (PISA 2000).

Matematik okur yazarlığı, matematiğin dünyada oynadığı rolü tanımlamak ve anlamak, iyi kurulmuş matematiksel yargılamalar yapmak ve bireyin şimdiki ve gelecek hayatındaki ihtiyaçlarını, yapıcı, düşündürücü ve yansıtıcı bir şekilde matematikle ilişkilendirmektir.

Hollins University (2011)

Sayısal mantık, gerçek dünyadaki problemleri çözmek için matematiksel kavram ve becerilerin uygulanmasıdır. Profesyoneller ve vatandaşlar olarak etkin bir şekilde performans göstermek için, öğrenciler sayısal verilerin okunması ve kullanılması, niceliksel kanıtların anlaşılması ve gerçek yaşam problemlerinin çözümüne temel sayısal becerilerin uygulanması konusunda yetkin olmalıdırlar.

National Numeracy Network (NNN 2011)

Sayısal akıl yürütme, sayısal bilgileri araştırmak, tartışmak, üzerinde düşünmek ve kamu, kişisel ve mesleki yaşamda uygulamak için zihin gücü ve alışkanlığıdır.

BYU (2011)

Öğrencilerin niceliksel olarak dünyayı anlama ve açıklama kapasitesine sahip olmalarına yarayan sayısal yetenekler; sayısal verileri yorumlamak; sayısal

bilgiye ve yaklaşımlara dayanan argümanları değerlendirmektir.

Kolata (1997).

Aritmetik ve geometri ötesinde, sayısal okur yazarlık mantık, veri analizi ve olasılık gerektirir. Bireylerin kanıtları analiz etmelerini, grafikleri okumalarını, mantıklı argümanları anlamalarını, mantıksal yanılgıları tespit etmelerini, kanıtları anlamalarını ve riskleri değerlendirmelerini sağlar. Sayısal okur yazarlık, akıl ve düşüncenin bilinmesi anlamına gelir

Bennett and Briggs (2008)

Sayısal okur yazarlık nicel bilgilerle, matematiksel fikirleri veya sayıları içeren bilgileri yorumlama ve akıl yürütme yeteneğidir. Sayısal muhakeme, niceliksel bilgi ile yorumlama ve mantık yürütme sürecidir.

Mekansal-Uzamsal yetenek ise, yönleri değiştirildiğinde iki veya üç boyutlu figürleri görselleştirme yeteneğidir. Bu, geometri ve topoloji gibi veya mühendislik ve kimyanın bazı alanları gibi, yapıların önemli olduğu bazı matematik formlarında önemli bir yetenektir (Kline, 2000, s.219).

1.1.6. Cinsiyete Dayalı Farklılıkların İşaretlerini Değişen Madde Fonksiyonunu ve Psikometrik-Çok İçerik Teorisini Kullanarak İncelenmesi

Neden erkekler ve kadınlar, standart testlerin analitik bölümleri üzerinde farklı performans gösterme eğilimindedirler? Psiko / sosyal araştırma, sıklıkla kadınların performansının "rol beklentileri veya haksız yetersizlik korkuları gibi değişkenlerden

daha fazla etkilenebileceğini" iddia etmektedir (Basinger, 1997, s.2; bkz. Sternberg ve Williams, 1997). Cinsiyete dayalı DMF'yi (Değişen Madde Fonksiyonu) matematik maddeleri üzerinde incelemek için yapılan psikometrik çabalar büyük oranda cinsiyet veya etnik köken gibi açık bir gruba üye olmanın bir fonksiyonu olarak farklılıkların incelenmesine odaklanmıştır (örn., Carlton & Harris, 1989; Scheuneman & Gerritz, 1990; Wild & McPeck). Bu yaklaşım, potansiyel olarak yanlı maddelerin testten çıkarılmasına veya yanlılığa neden olabilecek maddenin/maddelerin düzeltilmesine yönelik yaygın bir şekilde kullanılmıştır. DMF analizinin bazı biçimleri, çoğu test şirketi tarafından, özellikle madde analiz aşamasında maddelerin taranması sürecinin bir parçası olarak kullanılır (Casey, 2005, s.143-144).

1.1.7. Sayısal Yeteneğin Cinsiyete Göre Farklaşmasının İncelenmesinin Önemi

“Gerçekten araştırma gibi bir şey yoktur. Sadece arama, daha fazla arama, aramaya devam etme vardır.” Elizabeth Bowering (1988, p. 95).

Eğitimde, seçme ve yönlendirme için eğitim sektöründe yetenekleri ölçmeye yönelik testler yapılmıştır ve endüstriyel psikoloji alanında bireyleri doğru işlere seçebilmek için de yetenekleri ölçmeye yönelik testler geliştirilmiştir. Örneğin, sivil helikopterleri uçurma eğitimi için bireyleri seçmek zorunda kalınırsa, bireylerin mekânsal-uzamsal yeteneğinin yüksek olmasını istenir ve buna konsantre olunur (Kline, 2000, s.217).

Cinsiyet farklılıklarını araştırmamız gerektiği gerekçesiyle ilgili soruya yanıt olarak, büyük bir uluslararası araştırmamanın yazarları şu noktalarla cevap verdi:

Cinsiyet farklılıklarını incelemek için en az üç neden vardır:

- i) Eşitsizliklerin kaynağını anlamak;
- ii) Ortalama performansı artırmak; ve
- iii) Öğrencilerin nasıl öğrendiğine ilişkin anlayışımızı geliştirmek (Halpern, 2013, s 22).

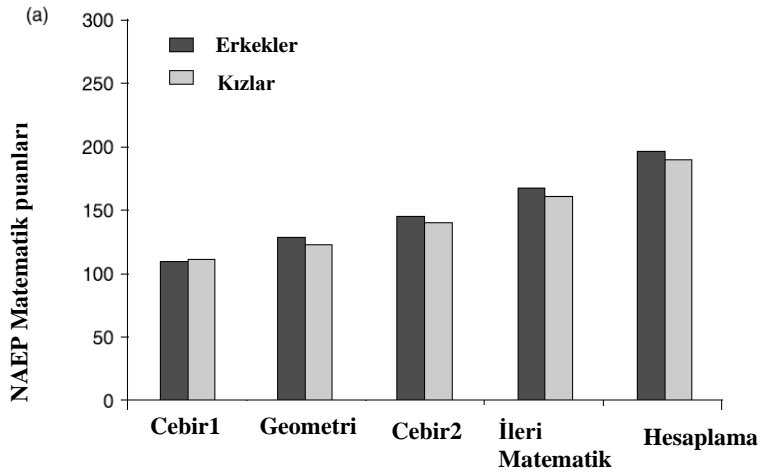
Cinsiyet farklılıkları, öğrencilerin geçmişinin ve öğrenci özelliklerinin öğrenci performansını önemli derecede etkilediği alanlara işaret etmektedir. Farklı öğrenci performansını yönlendiren şeyleri anlamak, kalite ve hakkaniyete ilişkin endişeleri gidermek için etkili eğitim politikalarının tasarımını teşvik edebilir.

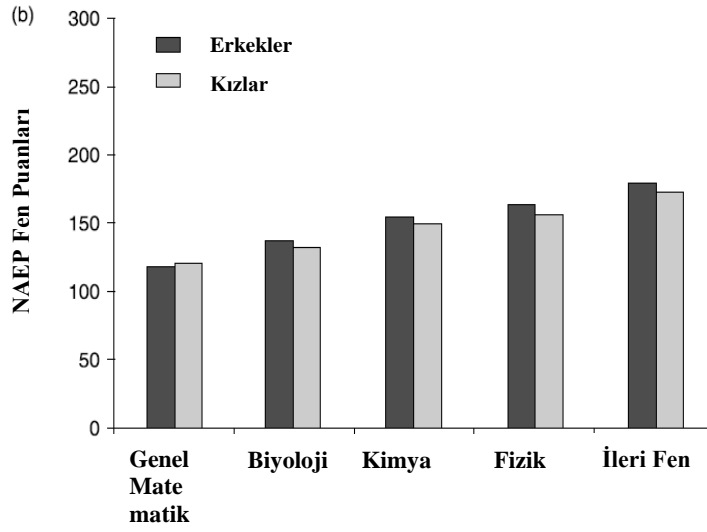
Ortalama grup farklılıklarını öğrenmek, bize farklılıkların kökeni hakkında hiçbir şey söylemez. Son birkaç on yılda, kariyer seçiminde, farklı eğitim alanlarına katılım ve aile yapılarında büyük değişiklikler olmuştur. Araştırmalar, cinsiyet farklılıklarının bulunduğu alanları tanımladığında, çıkan sonuç şu anki farklılıklarla ilgilidir.

Matematik ve fende başarılı olma yeteneğini eğitim deneyimlerinden elde edilen başarıdan ayırt edemeyiz. Yetenek ile başarı arasındaki bulanık ayrım, cinsiyet farklılıkları gösteren testlerden aldığımız sonuçlara çok dikkat etmemiz gerektiği anlamına gelir.

Yeteneği ölçmeye çalıştığımızda bir dereceye kadar her zaman başarıyı ölçüyoruz. Bu, olası cinsiyet farklılıklarını anlamak isteyen psikologlar için zahmetli bir sorundur. Amerikan ve diğer batı toplumlarında kız öğrenciler erkeklerden daha az ileri matematik dersleri almışlardır, ancak son yıllarda lise ve üniversite ileri matematik derslerinde kız ve erkeklerin oranı eşittir (D. F. Halpern, 2013). Şaşırtıcı olmayan bir şekilde, yakın tarihli araştırmalar, erkeklerin ve kızların, üniversitesindeki matematik ve fen başarı testlerinde benzer puanlar aldığını göstermektedir.

Şekil 1: 2005 Ulusal Eğitimsel İlerlemeyi Değerlendirme Raporunda Erkek ve Kızların Puanı (NAEP).





NAEP, Ulusal Rapor Kartı olarak bilinir. Bu veriler, matematikte (a) ve fende (b) kızlar ve erkekler arasındaki farkları göstermektedir. ABD Eğitim Bakanlığı, Eğitim Bilimleri Enstitüsü, Ulusal Eğitim İstatistikleri Merkezi (2005) (Halpern, 2013, sf 23).

Bilişsel kabiliyetin saf ölçüsü, her bir cinsiyetin aslında ne yaptığını (başarı) ve her birinin ne yapabileceğini (yeteneği) ayırmalıdır ama bu mümkün değildir. Örneğin, ne kadar matematik bildiğini ölçmek istediğimiz birine, bir matematik başarı testi veririz. Lisede çok az matematik dersi alan biri için, almadığı lise derslerinde öğretilen matematik türü hakkında pek fazla şey bilemeyeceğini umarız. Örneğin trigonometri veya hesap problemlerini çözemez. Bu, bu sorunları uygun öğretim ile çözemeyeceğimiz anlamına gelmez ve bu matematiksel kavramları öğretmekte güçlük çekeceğimiz anlamına gelmez. Bu testten alınan düşük puan, yalnızca test anındaki matematik problemlerini çözemediği anlamına gelir. Yetenek testleri, doğru talimatlar alındıysa ve görevi gerçekleştirmek için gereken becerileri öğrenmek ve göstermek için motive edildiysen, ileride belirli görevlerde başarılı olma ihtimalini değerlendirmeye çalışır.

Sadece 1980'lerde veya 1990'larda veya 21. yüzyılın tek haneli yıllarında uygulanan testlerin cinsiyet farklılıklarını göstermesinin, 2015 yılı ya da sonraki yıllarda yapılan testlerin de göstereceği anlamına gelmiyor (Diane F. Halpern, 2012).

Dünyada, bireylerin okul yaşantılarındaki başarılarının ölçülmesi ve kariyer basamaklarında ilerlemelerinin sağlanması için çeşitli ölçümler yapılmaktadır. Bu ölçümlerin yapılabilmesi için çeşitli testler kullanılır. Ölçme sonuçlarına göre de eğitim ve iş yaşamıyla ilgili tercihler yapılır. Ülkemizde, ÖSYM (Öğrenci Seçme ve

Yerleştirme Merkezi) tarafından üniversiteye yerleşmek isteyen adaylar için hazırlanan Lisans Yerleştirme Sınavı (LYS) yapılmaktadır. LYS, sınav sisteminin ikinci aşamasıdır. LYS yani Lisans Yerleştirme Sınavı üniversiteye yerleşmek isteyen adayların YGS(Yükseköğretime Geçiş Sınavı)'de 180 barajını geçmeleri halinde girebilecekleri ve 4 yıllık lisans programlarına yerleşebilecekleri genel sınavın adıdır.

Tablo 3'deki, 2017 LYS (Lisans Yerleştirme Sınavı) sınavı ortalama puanlarının cinsiyete göre sayısal değerlendirmesine baktığımızda Matematik-Fen (MF), Türkçe-Matematik (TM) ve Türkçe-Sosyal (TS) alanlarında erkeklerin ortalama puanlarının kızlarınkinden daha düşük olduğu görülmektedir ("Öğrenci Seçme ve Yerleştirme Merkezi [ÖSYM]," 2017).

Tablo 3: 2017 Lisans Yerleştirme Sınavı Alanlara Göre Kız ve Erkek Öğrencilerin Ortalama Puanları

| Cinsiyet | Aday Sayısı | MF | | TM | | TS | |
|----------|-------------|----------|-------------|----------|-------------|----------|-------------|
| | | Ortalama | Aday Sayısı | Ortalama | Aday Sayısı | Ortalama | Aday Sayısı |
| Erkek | 512.975 | 233,760 | 228.933 | 209,160 | 355.020 | 213,248 | 365.794 |
| Kız | 559.423 | 237,408 | 216.496 | 215,624 | 433.168 | 220,216 | 440.425 |

Dünya genelinde ve Türkiye'de üniversitelerdeki fen, matematik ve teknoloji alanlarına ya da bu alanlarda hizmet veren birimlere bakıldığında kızlardan daha çok erkeklerin olduğu görülmektedir. 2016-2017 Eğitim alanlarına göre öğrenci sayılarına bakıldığında erkeklerin fen, teknoloji, bilgisayar, matematik ve mühendislik fakültelerine yerleşme oranının kızlarınkinden azımsanmayacak derecede fazla olduğu görülmektedir. Tablo 4'de 2016-2017 eğitim öğretim yılı LYS sınavı sonuçlarına göre sayısal alanlara yerleşen kız ve erkek öğrenci sayıları gösterilmiştir (ÖSYM, 2017).

Tablo 4: 2016-2017 Eğitim Öğretim Yılı LYS Sınavı Sonuçlarına Göre Sayısal Alanlara Yerleşen Kız ve Erkek Öğrenci Sayıları

| Fakülteler | Erkek | Kız |
|---------------------------------------|--------|--------|
| Havacılık ve Uzay Bilimleri Fakültesi | 926 | 281 |
| Mühendislik Fakültesi | 35.915 | 14.413 |
| Mühendislik ve Mimarlık Fakültesi | 7377 | 3137 |

| | | |
|------------------------------------|------|------|
| Tıp Fakültesi | 6256 | 6803 |
| Bilgisayar ve Bilişim Fakültesi | 129 | 33 |

Kadınların bu alanları daha az tercih etmesinin nedeni nedir? Yeteneklerin cinsiyete göre farklılaşması mı, kadınların ve erkeklerin kariyer ve ilgi alanlarının farklı olması mı, yoksa cinsiyet rollerinin sosyal çevre tarafından farklı yönlendirilmesi mi bu duruma sebep olmaktadır? Bilhassa farklı alanlarda erkekler ve kızların yetenek testlerindeki puanlarının farklılaşması; bunun yanında kız öğrencilerin meslek seçimlerinde matematik ve fen bilimlerindeki alanları çok fazla tercih etmemesi ve sosyal çevrenin ve ailelerin kız ve erkelere olan etkisi, sosyo-ekonomik ve biyolojik etkenler göz ardı edilemez.

Sayısal yetenek testlerinde cinsiyet temelli bir farklılaşmanın olup olmadığını araştırmak bu aşamada önemlidir. Meslek seçimi ve kurumlara yerleşme süreçlerinde kız ve erkek bireylerin girdikleri sınavlardaki testlerin kızların veya erkeklerin lehine ya da aleyhine olmaması gerekir.

Hayatlarımız boyunca, ama özellikle de okul yıllarında herkes çok sayıda teste tabi tutulur. Bu testlerden bazıları yüksek öneme sahiptir, yani yaşamları etkileyen kararlar için temel oluştururlar. Yüksek öneme sahip sınavların çoğunda, cinsiyet, ırk veya etnik grup, sosyo-ekonomik durum ve coğrafi bölgenin bir fonksiyonu olarak farklılık gösteren grup farklılıkları veya ortalama farklılıklar vardır (Willingham & Cole, 1997). Grup farklılıklarının olması testlerin yanlı olduğu anlamına mı geliyor? Genel olarak, testler üç amaçla kullanılır: Geçmiş öğrenmeyi değerlendirmek, derslerdeki gibi bir müdahale sonrasında öğrenme kazanımlarını ölçmek ve geleceği tahmin etmek. SAT ve American College Test (ACT) gibi standart kolej giriş sınavları, sınavın yapıldığı zamana kadar geliştirilen yetenek seviyesini ölçer ve önceki öğrenmelerin bu endeksine dayanarak gelecekteki öğrenme görevlerinde muhtemel başarı ile ilgili öngörülerde bulunur. Bu, yordama geçerliliği adı verilen bir geçerlilik türüdür. İstatistikçiler, psikometri uzmanları (ölçme bilimlerinde uzmanlaşmış, genellikle zekâ veya kişilik gibi psikolojik yapıları ölçen kişiler) ve diğer test

uzmanları için bir madde, bir deęişkendeki başarıyı farklı gruplardaki bireyler için eşit derecede iyi yordarsa, yansızdır (. Halpern, Diane F., 2012, s. 54-55).

1.1.8. Yanlılık Tanımı

Madde yanlılığı, sınava giren farklı grupların, aynı test maddesine farklı yanıt vermesidir. Bu farklılıklar, hem test maddesinde hem de sınava giren farklı grupların deneyimlerine ve geçmişlerine ışık tutabileceği için keşfedilmeyi gerektirir (Holland & Thayer, 1986a).

Test yanlılığı, bir testin belirli bir grubun üyeleri için nasıl ölçüm yaptığı konusunda geçersizlik veya sistematik hata olarak tanımlanır. Yanlılık, belirli bir grubun üyesi için test sonuçlarında bozulma yaratması bakımından sistematiktir (Camilli & Shepard, 1994).

Yanlılık, test puanlarının veya bu puanlara dayanan tahminlerin istatistiksel bir karakteristiğidir (özelliğidir). Yanlılık istatistiksel bir sorundur, ampirik olarak tanımlanabilir ve varlığı bilimsel olarak belirlenebilir. Yani, yanlılığın, bir testin ölçüm veya tahminlemede sistematik hatalar yapması durumunda var olduğu söylenir. Üniversite başvurularında kullanılan bir test, sistematik olarak erkeklerin performansını yüksek tahmin ederse ve kadınların performansını düşük tahminlerse, bu test yanlı tahminleyici olarak değerlendirilir. Test verileri incelenerek, bir testin yanlı ölçümler veya yanlı tahminler sağlama derecesi belirlenebilir (Murphy & Davidshofer, 2005).

Bir testin öğrencinin bilişsel yeteneğini yanlı ya da yansız bir ölçüt haline getirip getirmediğine bakılmaksızın, testin yanlı tahminlerde bulunması olasılığı yüksektir. Örneğin, akademik bir yetenek testi, erkeklerin gelecekteki akademik performansını rutin olarak fazla tahmin edebilir ve kadınlarınkini düşük tahminleyebilir. Bu durumda akademik seçim veya yerleştirme kararlarını almak için bu test kullanılıyorsa, kadınlar dezavantajlı olacaktır.

Testlerdeki farklılaşmalar madde ve test bazında incelenmektedir. Bu çalışmalar çoğunlukla DMF analizi ile yapılmaktadır. DMF analizlerinin Madde Tepki Kuramı (MTK) yani IRT (Item Response Theory) temelli yöntemlerle yapılması yaygındır (Murphy & Davidshofer, 2005).

1.2. Problem Cümlesi

Sayısal yetenekte cinsiyet gruplarına göre farklılaşma var mıdır?

1.3. Alt Problemler

1. Bu araştırma kapsamında geliştirilen Sayısal Yetenek Testi (SYT)'nin psikometrik özellikleri nelerdir?
2. SYT Rasch Modelinin varsayımlarını karşılamakta mıdır?
3. SYT maddeleri cinsiyet gruplarına göre değişen madde fonksiyonu (DMF) içermekte midir?

1.4. Sınırlılıklar

Bu çalışmada uygulanan SYT'nin sadece İzmir ilindeki %8'lik dilim ve üstündeki başarı sıralamasıyla liseye yerleşen 11. Sınıf öğrencilerine uygulanmış olması araştırmanın sınırlılıklarından biridir.

Diğer bir sınırlılık ise DMF araştırmasının sadece Rasch analizi ile yapılmış olmasıdır. Diğer DMF belirleme teknikleri kullanılmamıştır.

1.5. Araştırmanın Amacı ve Önemi

Bu çalışmanın amacı, değişen madde fonksiyonunu (DMF) belirlemek için Rasch modelini kullanarak Madde Tepki Kuramı çerçevesinde, Lise 11. Sınıflara yapılacak olan Sayısal Yetenek testindeki maddelerin cinsiyete göre DMF içerip içermediğini araştırmaktır.

Bu konu ile ilgili yapılan araştırmalara bakıldığında Sayısal Yetenek testlerinde erkeklerin kızlardan daha başarılı olduğu sonucuna varılmaktadır. Bu çalışmada, daha önce yapılan birçok araştırmanın tersine, PISA, TIMMS, OKS, SBS dataları gibi hazır data kullanmak yerine bir sayısal yetenek testi oluşturulmuştur. Bu teste, geçerlilik,

güvenirlilik çalışmaları yapıp madde parametreleri hesaplanarak asıl şekli verildikten sonra, test lise 11. Sınıf öğrencilerine uygulanmıştır. Çalışma, uygulanacak yaş gurubunun daha önce çalışılmamış olması ve hazır data kullanılmaması açısından orjinaldir. Sayısal yetenek testindeki maddelerin cinsiyete göre DMF içermesi durumu olası bir yanlılık kaynağına işaret edebilir ve buna bağlı olarak bu sonucun eğitimde fırsat eşitsizliğinin sınırlı bir kanıtı olarak literatüre katkısı olacaktır.



BÖLÜM II

İLGİLİ ARAŞTIRMALAR

Bu çalışmada,11. sınıfa devam eden kız ve erkek öğrenciler arasında Sayısal Yetenek testinden aldıkları puanlara göre farklılık olup olmadığı araştırılacaktır.

Geçmişten günümüze kadar konu ile ilgili birçok çalışma yapılmış ve Sayısal Yetenek konusunda belli yaş gruplarında cinsiyete dayalı farklar bulunmamış ama belli yaş gruplarında da bu fark açıkça gözlenmiştir (Brulle, 2004). Aşağıda Sayısal Yetenekte Cinsiyet farklılığı konusu ayrıntılı ele alınarak, bu konuda yapılmış olan çalışmalardan bahsedilmiştir.

2.1. Genel Zihinsel Yetenek Testleri ve Zekâ Testleri Üzerine Yapılan Tartışmalar

Genel zihinsel yetenek testleri ya da zekâ testleri üzerine yapılan tartışmalar, bu yüzyılın ilk yıllarında başlar (Cronbach, 1975; Hartigan & Wigdor, 1989; Hilgard, 1989; Sokal, 1987). Son 20-30 yıl içinde, bu tartışmada itici güç, ırk temelli, etnik kökenli ve cinsiyete dayalı farklılıkların test puanlarındaki varlığı ve anlamı olmuştur (Gottfredson, 1988; Lerner, 1989; Scarr, 1989; Sheppard, 1982). Erkeklerin ve kadınların, bazı akademik yetenek testlerinde aldığı puanlar farklılık göstermektedir. Scholastic Aptitude Test(SAT) sınavındaki puanlar, özellikle sınavın matematik bölümünden alınan puanlar, burs vermek için tek ölçüt olarak alındığında kadınlar için daha az burs ile sonuçlanmıştır (Burton, Lewis & Robertson, 1988) (Murphy & Davidshofer, 2005, s.55-56).

Bilişsel yetenek testlerinin azınlıklara, kadınlara ve diğer belirli alt gruplara yardım edip etmeyeceği veya zarar verip vermeyeceği konusunda sorular vardır. Test tartışmaları, bu testlerin eşit istihdam, okullara eşit erişim ve diğer fırsatlar için yapay engeller yarattığını öne sürmektedir (Goldstein & Patterson, 1988) (Murphy & Davidshofer, 2005,sayfa 57).

Bilişsel yetenek testlerinin yanlı olduğuna dair yaygın bir inanç vardır ve eğitim istihdamı, akademik kabuller ve personel seçimlerinde bu testlerin kullanılması temelde adil olmayan kararlarla sonuçlanmaktadır. Testlerin yanlı olduğunu

düşündüren bir kanıt da, test puanlarındaki sosyo-ekonomik durum, cinsiyet ve ırkın bir fonksiyonu olarak sistematik farklılıkların varlığıdır (Jensen, 1980; Linn, 1982). Test yanlılığı üzerine çalışan araştırmacılar, genel olarak şu şekilde sonuçlara varmışlardır:

- Orta ve üst sosyoekonomik sınıftaki çocuklardan ve erişkinlerden elde edilen test puanları, alt sosyoekonomik sınıflardaki çocuk ve erişkinlerden elde edilen test puanlarından daha yüksektir.
- Beyazlar, azınlıklardan daha yüksek puanlar almaya meyillidir.
- Erkekler bazı testlerde sistematik olarak kadınlardan daha yüksek puanlar alırlar ve bazı diğer testlerde puanlar sistematik olarak kadınlardan daha düşüktür (Murphy & Davidshofer, 2005, s.316).

2.2. Sayısal Yetenek Testlerinde Cinsiyet Farklılıklarına İlişkin Açıklamalar

Sayısal yetenek testi performansındaki cinsiyet farklılıklarının gelişimi için yapılan açıklamalar genel olarak dört biçimdedir: sosyal, bilişsel, biyolojik ve bunların kombinasyonları. Wilder (1997), Willingham ve Cole (1997) ,cinsiyet farklılıklarının nedenlerinin analizinde şu sonuca varmıştır: Denklem - biyolojik ve sosyal - her iki tarafının da test performansına yansıyan farklılıklara yol açtığı tartışmasız olduğu görülmektedir. Bununla birlikte, bu büyük etki kategorilerinin var olan farklılıkları üretmek için etkileşim kurma şekilleri hakkında büyük sorular bulunmaktadır. Farklılıkların zaman içinde değiştiğini ve ölçülme biçimlerinin bir fonksiyonu olarak değişiklik gösterdiği eklendiğinde, erkek-kadın farklılıklarının öncülerinin olması, kadınlar arasındaki farklılıkların verimli bir alanı olacağına inanmak için iyi bir nedendir. Halpern ve LaMay (2000) ,bilişsel yeteneklerde cinsiyet farklılıklarına "psikobiyososyal" bir modeli savunan bir makalede de benzer bir sonuca varmışlardır. Onların psikobiyolojik modeli, bilişsel yetenek üzerindeki bazı etkilerin hem biyolojik hem de toplumsal olduğu ve bir veya ikisi olarak kolaylıkla sınıflandırılmayacağı fikrine dayanmaktadır (Casey, Nuttal, Byrnes, & Pajares, n.d., 2005, s.102).

2.3. Konu ile İlgili Yurtdışında Yapılan Bazı Çalışmalar

Hyde, J.S. Fennema, Elizabeth ve Susan J. matematik performansındaki cinsiyet farklılıklarına ilişkin 1963-1988 yılları arasında yayınlanmış 100 adet çalışmanın bir meta-analizini gerçekleştirdiler. Araştırmacılar, sürekli olarak erkeklerin matematik testlerinde kızlardan daha iyi performans gösterdiğine karar vermiştir. Yaş eğilimlerinin incelenmesi sonucunda, kızların ilkokul ve ortaokulda hesaplamada testinde hafif bir üstünlük gösterdiğini, ilkokul ve ortaokulda ise problem çözme alanında cinsiyet üstünlüğü olmadığını; Lise ve kolejde ise erkekler lehine olan farklı sonuçlar ortaya çıkmıştır. Matematik performansındaki cinsiyete dayalı başarı farklılıklarının küçük olduğu ve bununla birlikte, problem çözmede lisedeki kızların performansının erkeklerinkinden düşük olduğu sonucuna varmışlardır (Hyde, J. S., Fennema, E., & Lamon, S. J., 1990).

S. A. Burnett, D. M. Lane ve L. M. Dratt (1979)'ın yaptıkları başka bir çalışmada mekansal görselleştirme yeteneğindeki cinsiyet farklılıklarının, matematiksel beceride cinsiyet farklılıklarına neden olabileceği hipotezi, 183 erkek ve 81 kız üniversite öğrencisinden oluşan bir grup için desteklenmiştir. İstatistiksel olarak kontrol edilen mekansal görselleştirme ile Sayısal Skolastik Beceri Testi (QSAT) puanlarında anlamlı bir cinsiyet farklılığı bulunmamıştır; tahminleyici değişken olarak cinsiyet, artan varyansın % 1'den azını açıklayabilmiştir.. Mekansal görselleştirme üzerine matematik regresyonunun eğimi cinsiyete göre farklılık göstermezken, erkeklerin kadınlara göre biraz daha iyi tahminlenebilir olduğu görülmüştür. Yüksek mekansal görselleştirme becerisine sahip, hem erkek hem de kızların QSAT puanları , düşük beceriye sahip olanların QSAT puanlarından daha iyi tahminlenebilir olduğu görülmüştür. Tahminlemedeki cinsiyet farkının, kadınlardan daha fazla mekansal görselleştirme kabiliyetine sahip olan erkeklerden kaynaklandığı sonucuna varılmıştır (Burnett, Lane, & Dratt, 1979).

CP Benbow, ve JC Stanley (1980) tarafından yürütülen çalışmada ise 9927 entelektüel yetenekli ortaokul öğrencisi üzerinde yürütülen çalışmada ise, matematiksel akıl yürütme yeteneğinde (Scholastic Aptitude Test-Matematik Testi) erkek öğrenciler lehine önemli bir cinsiyet farkı bulunmuştur. Veriler, matematiksel

beceride gözlemlenen cinsiyet farklılıklarının farklı dersler alma ile açıklanabileceği hipotezi ile çelişmekte fakat bu farklılıkların çevresel etkilerle biraz arttığı hipotezini desteklemektedir (Benbow & Stanley, 1980).

Julia Sherman (1980), Sekizinci Sınıfta okuyan, matematik arka planı kontrol edilmiş bilişsel becerileri ve matematiğe karşı tutumları benzer 135 kız ve 75 erkek üzerinde yaptığı çalışmada, erkeklerin kızlara göre matematikte 11. sınıfta anlamlı olarak daha iyi performans gösterdiğini tespit etmiştir. Mekansal görselleştirme konusunda cinsiyete bağlı bir fark gelişmiştir. Ölçümlere, Bilişsel Yetenek Testi-Sözel bölüm, Matematiksel Kavramlar Testi ve Fennema-Sherman Matematik Tutum Ölçekleri II de dahil edilmiştir. Bu dönemde kadınların matematiğe yönelik tutumları erkeklerin tutumlarından daha olumsuz bir hale geldiğinden, sonuçlar cinsel rolün sosyokültürel etkilerinden kaynaklanıyor gibi görünmektedir (Gallagher & Kaufman, 2000, Sherman, J., 1980).

Plomin ve Foch'un araştırması, bilişsel cinsiyet farklılıklarının, toplam değişkenliğin çok az bir kısmını oluşturduğu sonucuna varmıştır; çünkü mekansal görselleştirme maddeleri içeren karmaşık bir testte, test puanlarındaki değişkenliğinin önemli bir bölümünü erkekler ve kızlar arasındaki farkların oluşturduğunu bulmuştur. İki ölçekli uzaysal görselleştirme-ETS Kart Rotasyonları testi ve Shepard / Metzler Zihinsel Rotasyon testi üniversite öğrencilerinin büyük bir örneğine uygulanmıştır. Kart Rotasyonları testi, düzlem içerisinde döndükten sonra oldukça basit özet formların tanımlanmasını gerektiren bir test; Zihinsel Dönüşler testi, 3 boyutlu uzayda rotasyondan sonra daha karmaşık 3 boyutlu şekillerin tasvirlerini tanımlamasını gerektiren bir testtir. Her iki testte de erkekler kızlardan önemli derecede yüksek puan almışlardır. Cinsiyet, Kart Rotasyonları testindeki varyansın yalnızca % 2'sini açıklarken, Mental Rotasyon testindeki varyansın %16'sını açıklamıştır. Shepard/Metzler Zihinsel Rotasyon testindeki benzer cinsiyet farklılıkları diğer araştırmacılar tarafından da bulunmuştur (Sanders, Soares, & D'Aquila, 1982).

Erin Leahey ve Guang Guo, makalelerinde erkek öğrencilerin matematikteki kız öğrencilere göre daha iyi performans gösterdikleri hipotezini test etmişlerdir. Büyük ulusal veri kümelerini ve eğrisel büyüme modellerini kullanarak, ilkokuldan liseye kadar matematiksel yörüngelerdeki cinsiyet farklılıklarını incelemişlerdir. Yüksek puan alan öğrencilerin alt grupları ile mantık ve geometri gibi farklı matematik

alanlarını analiz etmişlerdir. İlkokulda eşit başlangıç noktalarına ve eşit eğimlere sahip olmalarına rağmen, erkeklerin daha hızlı bir hızlanma oranına sahip olduklarını görmüşlerdir. Bu durum, 12. sınıfa kadar, en çok geometri alanında görülen az bir cinsiyet farklılığına neden olur. Bu hafif ve gecikmiş cinsiyet farklılıklarının farkına vararak, Benbow ve Stanley (1980, 1983) gibi daha önceki araştırmaların güçlü sonuçları değerlendirilmiş ve cinsiyet farklılıklarının ortaokul yıllarında ortaya çıktığı tespit edilmiştir (Leahey & Guo, 2001).

Akademik yetenek testinin (SAT) matematiksel ve sözel bölümleri ile yüzlerce entelektüel açıdan yetenekli 12 ila 13 yaş arası çocuklar son 16 yılda ülke genelinde test edilmiştir. Sözel yetenekte hiçbir cinsiyet farklılığı bulunmamakla birlikte, matematiksel akıl yürütme yeteneğinde, SAT matematik bölümü (SAT-M) ile ölçülen tutarlı cinsiyet farklılıkları erkeklerin lehine olmuştur. Bu farklılıklar, matematiksel akıl yürütmenin en üst düzeylerinde daha çok belirginleşir, istikrarlıdır ve diğer ülkelerde de gözlemlenir. Matematiksel akıl yürütme yeteneğindeki cinsiyet farkı, matematik ve fen başarısındaki cinsiyet farklılıklarını tahminleyebilir ve bu nedenle pratik bir öneme sahiptir. Şu ana kadar, yetenek farklılığı için yapılan çevresel açıklamalar, Matematiksel Hassas Gençlik Çalışması (SMPY) ve diğerleri tarafından uzun yıllar süren sayısız çalışmadan destek almadı. Klasik çevresel hipotezlerden bazıları: matematiğe yönelik tutumlar, matematiğin yararlılığının algılanması, güven, ebeveynlerden ve diğerlerinden beklentiler / teşvik, cinsiyet tipi ve farklı ders alımı şeklinde sıralanabilir. Ek olarak, son derece yüksek matematiksel akıl yürütme yeteneğinin çeşitli fizyolojik korelasyonları tespit edilmiştir (solaklık, alerjiler, miyopi ve bilişsel işlevlerin ve prenatal hormonal maruziyetin ikili gösterimi). Bu nedenle, entelektüel yetenekli öğrenciler arasındaki SAT-M puanlarındaki cinsiyet farkının erkek değişkenliği ile ilişkili olabileceği hem çevresel hem de biyolojik faktörlerden kaynaklandığı düşünülmektedir (Thomas, 1993).

Başka bir araştırmada TIMSS matematik testlerindeki büyük cinsiyet farklılıklarının orta öğretimin bitimine kadar var olduğu görülmüştür. Ortaöğretim öğrencileri için cinsiyete göre sonuçlar ortaöğretim öncesi sınıflarda olanlar ile farklılık göstermiştir. Ülkelerin çoğunda, matematik okuryazarlığı ve ileri matematikte erkekler kızlardan daha fazla ortalama (tahmin edilen ortalama $d=5.38$) başarı oranına sahiptir

(Mullis, Martin, Fierros, Goldberg, & Stemler, 2000, s.13)(Halpern Diane F. et al., 2007).

2.4. Konu ile İlgili Türkiye’de Yapılan Bazı Araştırmalar

Öğretmen (1995) yüksek lisans tez araştırması kapsamında Türkiye’deki üniversite giriş sınavının birinci basamağının sözel yetenek bölümündeki maddelerin madde yanlılığı analizini Madde Tepki Kuramı çerçevesinde dört farklı alan indeksini kullanarak yapmıştır. Türkçe ve Sosyal bilimler testlerindeki maddeleri konu alanları ve taksonomik seviyelerine göre karşılaştırmış ve bu karşılaştırmaları cinsiyet guruplarına göre yapmıştır. Ağırlıksız alan indekslerinin Türkçe ve Sosyal Bilimler testlerindeki çoğu maddenin farklı yetenek seviyelerinde cinsiyet guruplarına göre değişen madde fonksiyonu gösterdiği sonucuna varmıştır. Türkçe testinde ağırlıklı ve ağırlıksız alan indekslerine göre madde yanlılığının gözlendiğini, Sosyal Bilimler testinde ise sadece ağırlıksız alan indeksine göre madde yanlılığı olduğunu ortaya koymuştur. Cinsiyet gurupları maddelerin konu alanlarına göre değerlendirildiğinde madde yanlılığı gözlenmesine rağmen, her iki alt testte de genel farklılığın azalma eğiliminde olduğunu belirtmiştir (Öğretmen, 1995).

Yenal (1995), Yüksek Lisans Tezi kapsamındaki çalışmasında Madde Tepki Kuramına göre elde edilen dört alan indeksiyle 1993 Öğrenci Seçme Sınavı Sayısal Testi'nin cinsiyete göre madde yanlılığı incelenmiştir. Kullandığı Alan indeksleri şunlardır: İşaretli alan, işaretli alan ağırlıklandırılmış işaretli alan ağırlıklandırılmış işaretli alan indeksleri. Elde ettiği alan indekslerini farklı soru guruplarına göre karşılaştırmıştır. Ayrıca, alan indekslerini kendi aralarında da karşılaştırmıştır. Bu çalışmanın denekleri Ankara'daki beş özel ve devlet okulundaki 1993 ÖSS'ye giren öğrencilerdir. Madde Yanlılığı Analizi, işaretli ve işaretli alan indekslerinin Madde Karakteristik Eğrilerinde düzgün olmayan bir yanlılık gösterdiğini ortaya çıkartmıştır. Konu alanlarının karşılaştırılması sonucunda geometri maddelerinin Matematik testinin

diğer konu alanlarındaki maddelere göre daha fazla yanlılık gösterdiği ve bu yanlılığın erkekler aleyhine olduğu gözlenmiştir. Fen Bilgisi testinde ise yanlılık en fazla biyoloji maddelerinde ve erkeklerin aleyhinde olmuştur. Ayrıca, yapılan analizler Fen Bilgisi

testinin Matematik testinden daha fazla yanlılık taşıdığını göstermiştir. Matematik ve Fen Bilgisi testlerini alan indeksleri kendi aralarında karşılaştırıldığında, işaretli alan indeksleriyle ağırlıklandırılmış işaretli alan indeksleri arasında anlamlı ilişki olduğu gözlenmiştir. Buna ek olarak Matematik testinde ağırlıklandırılmış işaretli alan indeksi ile ağırlıklandırılmış işaretsiz alan indeksi arasında, işaretli alan indeksi ile işaretsiz alan indeksi arasında anlamlı ilişki olduğu gözlenmiştir. Matematik ve Fen Bilgisi testlerinde konu alanlarına göre yanlılık olmasına rağmen testlerin tamamına bakıldığında bu yanlılığın kaybolduğunu ortaya koymuştur (Yenal,1995).

Büyükatacık , PISA 2012'deki matematiğe yönelik duyuşsal özelliklerin bölge, okul türü ve cinsiyete göre sınıflama doğruluğunun incelenmesi başlıklı yüksek lisans tezindeki araştırma sonunda; "Matematiğe Yönelik Duyuşsal Özelliklerin öğrenci performansı üzerindeki açıklayıcılıklarının düşük olması sebebiyle öğrencilerin bölge, okul türü ve cinsiyete göre sınıflandırmasında çok etkili olmadığı sonucuna ulaşmıştır. Bu kapsamda Matematiğe Yönelik Duyuşsal Özelliklerin öğrencileri yaşadıkları bölgelere göre ayırmada kullanılamayacağı, bu özelliklerin okul türleri açısından anlamlı bir fark yaratmadığı fakat duyuşsal özelliklerin cinsiyete göre çok anlamlı olmasa dahi şansın ötesinde doğru sınıflandırma yapabilmesi sebebiyle öğrencileri ayırmada kullanılabilir" sonucuna varmıştır (Büyükatacık, 2016).

Köse (2012),PISA 2003, 2006 ve 2009 Türkiye Uygulaması Matematik Ortak Maddelerindeki Başarıların İncelenmesi başlıklı yüksek lisans tezinde PISA 2003, PISA 2006 ve PISA 2009 uygulamaları matematik okuryazarlığı ortak maddelerine göre başarının madde özellikleri ve öğrenci cinsiyeti açısından incelenmesi konusunu araştırmıştır. Yaptığı analizler sonucunda ulaşılan bulgular göstermiştir ki; Türkiye'deki öğrenciler, 'uzay ve şekil' matematiksel içerik alanında daha az başarılı olmuşlardır. Türkiye'deki öğrenciler, 'üretici' yeterlilik kümesinde daha fazla başarılı, "yansıtıcı" yeterlilik kümesinde daha az başarılıdır. Türkiye'deki öğrenciler, 'kişisel'durum maddelerinde daha fazla başarılıdır. Türkiye'deki öğrenciler, 'çoktan seçmeli'madde tipinde daha fazla başarılı, 'açık uçlu' madde tipinde daha az başarılı oldukları görülmüştür. Türkiye'deki öğrencilerin genel olarak PISA 2003'te daha başarılı, PISA 2006'da daha az başarılı olduğu görülmüştür. İkinci alt problem için incelenen kitapçıklardan hepsinde erkek öğrencilerin ortalama puanlarının, kız

öğrencilerinkinden daha fazla olduğu ancak bu farkın sadece PISA 2003 kitapçık 1’de anlamlı olduğu görülmüştür (Köse, 2012).

Ülkemizde Rasch analizi kullanılarak yapılan araştırmalar sınırlıdır. Bunlardan bir tanesi Ankara üniversitesi Biyoistatistik anabilim dalı Ölçeklerde Yapı Geçerliliğinin Değerlendirilmesinde Faktör Analizi ve Rasch analizi yaklaşımlarının karşılaştırıldığı yüksek lisans tezidir. Özalp Ateş, bu tez çalışması kapsamında, romatoid artritli hastalar için “kendine bakım-hareket-ev işleri” alanlarında özürlülük düzeylerinin alt ve üst ekstremité açısından ayrı ayrı değerlendirilmesi için geliştirilen ölçme araçlarının yapı geçerliliğinin değerlendirilmesinde AFA, DFA ve Rasch Analizi uygulanmış ve sonuçları irdelenmiştir. Ölçeklerde yapı geçerliliğinin incelenmesinde, ölçeğin sadece boyut yapısını inceleyen faktör analizi yöntemlerine kıyasla, içsel yapı geçerliliğinin madde uyumu, birey uyumu, madde yanlılığı gibi farklı özellikler bakımından da değerlendirildiği Rasch Analizi’nin beraber kullanılmasını önermektedir (Özalp Ateş, 2015).

Yüksel, Ankara Üniversitesi Biyoistatistik anabilim dalı, Ölçeklerde Saptanan Madde İşlev Farklılığının Karma Rasch Analizi ile İncelenmesi başlıklı doktora tezinde, heterojen veri setlerinde Karma Rasch Modelleri ile elde edilen kişi parametre kestirimlerinin altın standart değerlere oldukça yakın olduğu ve Kısmi Kredi Modeline göre daha iyi kestirimler elde edildiği söylemiştir. Literatürde iki sonuçlu maddeler içeren ve küçük örnekleme sahip veri setlerinde Karma Rasch Modelleri’nin etkin sonuç vermediğini ifade etmiştir. Bu çalışması ile çok sonuçlu maddeler içeren ve küçük örnekleme sahip veri setlerinde Karma Rasch Modelleri’nin Madde İşlev Fonksiyonuna neden olan etkeni tanımlamada başarılı olamadığını kanıtlamıştır (Yüksel, 2012).

Dadaş, Ege Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik ve tıbbi Bilişim Anabilim Dalı yüksek lisans tezinde, sağlık alanında Romatoid Artritli (RA) hastaların “kendine bakım-hareket-ev işleri” alanlarında özürlülük düzeylerinin değerlendirildiği ölçme aracı kullanmış, bu araçta yer alan maddelerden sırasız eşik değerine sahip maddelerdeki kategorilerin birleştirilmesinde kullanılan dört farklı stratejinin Rasch modeline uyum üzerindeki etkisini incelemiştir. Sonuç olarak bu veri için, kategori birleştirme sonrası elde edilen ölçme araçlarında kategorilerin birleşmesine bağlı bilgi kaybı gerçekleşmiş olup bireylerin θ ’larına ait standart

hataların yükseldiğini ifade etmiştir. Fakat birleştirme stratejileri sonrasında, ölçme araçlarının değişmezlik özelliğinin sağlandığı, model uyumunun ve iç tutarlılığının arttığını tespit etmiştir. Ayrıca dört farklı strateji sonrası elde edilen ölçme araçlarında, hem her ölçekten madde olması hem de her ölçeğin alt bölümlerinin temsil edilmesi, bu ölçme araçlarının kapsam geçerliğine ilişkin önemli bir kanıt niteliğinde olduğunu belirtmiştir (Dadaş, 2018).

T.Öğretmen ve N.Doğan, 2004'te yayınladıkları bir araştırmada Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme sınavı(OKÖSYS) Matematik alt testindeki maddelerin madde yanlılığı analizlerini madde tepki kuramı çerçevesinde incelemişlerdir. Araştırmadan elde edilen sonuçlar, OKÖSYS Matematik alt testinin, cinsiyet gruplarına göre testin her maddesinin yanlılık gösterdiğini ortaya koymuştur. Özellikle geometri soru grubunun hemen hepsinin yanlılık değerlerinin yüksek ve kızların lehine olduğunu bulmuşlardır (Öğretmen & Doğan, 2004)

BÖLÜM III

YÖNTEM

Bu bölümde Araştırmanın Türü, Çalışma Gurubu/Evren ve Örneklem, Verilerin Toplanması, Veri Toplama Aracı, Verilerin Analizi ve Analizlerde kullanılacak olan istatistik paket programı ile Madde Tepki Kuramı, DMF ve Rasch Analizi açıklanmıştır.

3.1. Araştırmanın Türü

Bu araştırmada kapsamında bir Sayısal Yetenek Testi geliştirilmiştir. Geliştirilen bu test İzmir ilindeki devlet Anadolu Liselerinde okuyan ve rasgele seçilen 11. Sınıf öğrencilerine uygulanmıştır. Araştırmanın bu yönü göz önüne alındığında deneysel bir çalışma olduğu söylenebilir.

Geliştirilen Sayısal Yetenek Testi'ni oluşturan alt testler Aritmetik Akıl Yürütme Alt Testi, Eşitlik Kurma Alt Testi, Geometri Alt Testi ve Görsel-Uzamsal Yetenek Alt Testi şeklindedir. Sayısal yetenek testindeki maddelerin DMF içerip içermediğine bakılmıştır. Son olarak, Rasch modeli kullanılarak, testin model varsayımlarını karşılayıp karşılamadığına bakılmıştır. Araştırma bu yönüyle betimsel bir araştırmadır.

3.2. Evren ve Örneklem

Araştırma evreni İzmir ilinde okuyan ve 2018-2019 eğitim öğretim yılı itibarıyla 11. Sınıfa devam etmekte olan öğrencilerden oluşmaktadır. Bu çalışma için seçilen örneklem ise 1036 kişilik devlet Anadolu Liselerinde okuyan 11. Sınıf öğrencisinden oluşmaktadır. Örneklem seçilirken Anadolu Liselerine yerleşme yüzdeleri en az %8 olan Anadolu Liselerindeki öğrenciler tercih edilmiştir. Çalışmaya katılan öğrenciler Aritmetik Akıl Yürütme, Eşitlik Kurma, Geometri ve Görsel-Uzamsal Yetenek alt testlerinden oluşan Sayısal yetenek testindeki 20 maddeyi cevaplamışlardır. Sonuç olarak öğrencilerin bu teste verdikleri cevapların verileri değerlendirmeye alınmıştır.

Sayısal Yetenek Testi dört alt testten oluşacak olup, her bir alt test 5'er soru içermektedir. Test toplam 20 madde olup, öğrencilere cevaplama süresi olarak 40 dakika verilmiştir. Yanıtlar doğru yanlış(1-0) olarak değerlendirilmeye alınmıştır.

3.3 Verilerin Toplanması

Veriler, İzmir ilinde bulunan bazı devlet Anadolu Liselerinin 11.sınıf öğrencilerinin Sayısal Yetenek Testi'ne verdikleri cevaplardan elde edilmiştir.

3.4 Veri Toplama Aracı

Sayısal Yetenek testinin asıl uygulaması yapılmadan önce her bir alt testi 10'ar maddeden oluşan toplam 40 maddelik bir test oluşturulmuştur. Bu test yaklaşık olarak 400 kişilik 11. Sınıf örnekleme uygulanmıştır. Bu uygulama sonrasında madde güçlükleri ve ayırdedicilikleri hesaplanarak madde seçimi yapılmıştır. Güvenirlilik ve geçerlilik çalışmaları yapılarak, ölçmek istediğimiz Sayısal yeteneği oluşturan yapıları (Aritmetik Akıl Yürütme, Eşitlik Kurma, Geometri ve Görsel-Uzamsal Yetenek) en iyi ölçen 5'er madde seçilip, teste son şekli verilmiştir.

3.5 Verilerin Analizi

Araştırmanın birinci alt problem cümlesindeki Sayısal Yetenek Testi'nin psikometrik özelliklerinin incelenmesi aşamasında 10'ar maddelik Aritmetik Akıl Yürütme, Eşitlik Kurma, Geometri ve Görsel-Uzamsal Yetenek alt testlerinden oluşan Sayısal Yetenek Testi 400 kişilik 11.sınıf öğrencilerinden oluşan örnekleme uygulanmıştır. Teste son şeklini vermek için uygulama sonucunda elde edilen veriler sonrasında madde güçlükleri ve ayırdedicilikleri hesaplanarak madde seçimi yapılmıştır. Güvenirlilik ve geçerlilik çalışmaları yapılarak, ölçmek istediğimiz yapıları en iyi ölçen 5'er madde seçilip, teste son şekli verilmiştir. Son şekli verilen test 1382 adet 11. Sınıf öğrencisine uygulanmıştır. Elde edilen veriler madde tepki kuramı çerçevesinde Rasch modeline göre WINSTEPS programı kullanılarak analiz edilmiştir. Araştırmanın ikinci alt problem cümlesini incelemek için veri setinin Rasch

modelinin varsayımlarını karşılayıp karşılamadığı araştırılmıştır. Üçüncü alt problemin araştırılması kapsamında DMF içeren maddeler tespit edilmiştir.

3.5.1. Seçilen Bilgisayar Programı Hakkında Genel Bilgi

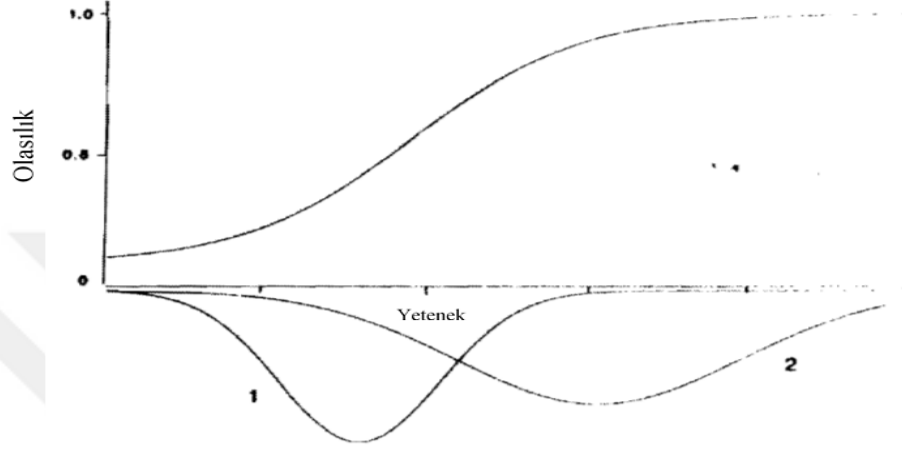
Literatürde, çoğu Rasch analizi özel bir yazılımla gerçekleştirilir. En yaygın kullanılan paketler WINSTEPS, RUMM2020 ve ConQuest'tir, ancak daha birçokları mevcuttur. Her biri, bulguları biraz farklı bir şekilde rapor etse de, temel öncül, verilerde gözlemlenen cevap modelinin model tarafından beklenen teorik kalıpla uyuşup uyuşmadığını test etmektir (yani, Guttman ölçeklendirmenin olasılıksal formu). Gözlemlenen ve beklenen arasındaki bu fark verilerin modele uygun olup olmadığını test etmek için kullanılan istatistiklerin merkezini oluşturur. Bu araştırmanın analizleri WINSTEPS programı ile yapılmıştır. Mike Linacre (2012a) tarafından geliştirilen Rasch yazılımı olan Winsteps kullanılmasının amacı son derece kullanıcı dostu, kullanıcı el kitabı çok ayrıntılı ve neredeyse anında çevrimiçi destek özelliklerine sahip olmasıdır. Winsteps bir analiz için kullanıldığında, bir ölçme aracının güvenilirliğini ve geçerliliğini birçok yönden değerlendirebilir, ek olarak, ölçmeye katılanların parametreleri kolayca hesaplanabilir. Bu ölçümler herhangi bir parametrik istatistiksel test için kullanılması gereken ölçümlerdir.

3.5.2. Madde Tepki Kuramı Hakkında Genel Bilgiler

Madde Tepki Kuramı (MTK) iki temel varsayım üzerinde durmaktadır: (a) Bir sınava tabi tutulan kişinin bir test maddesindeki performansı, özellikler, örtük özellikler veya yetenekler denilen bir dizi faktör tarafından tahmin edilebilir (veya açıklanabilir). (b) Sınava tabi tutulanların madde performansı ile madde performansını temel alan örtük özellikler arasındaki ilişki, bir madde karakteristik fonksiyonu veya madde karakteristik eğrisi (MKE) olarak adlandırılan monoton olarak artan bir fonksiyon tarafından tanımlanabilir. Bu fonksiyon, yetenek seviyesi arttıkça, bir maddeye doğru yanıt verme olasılığının arttığını belirtir (Hanlbleton, Swaminathan, & Rogers, n.d.sayfa 7). Şekil 2,iki gurup öğrencinin yetenek dağılımıyla birlikte, bir maddenin karakteristik fonksiyonunu göstermektedir. Özellik üzerinde daha yüksek

değerlere sahip olan bireylerin, grup üyeliğine bakılmaksızın, özellik üzerinde daha düşük değerlere sahip olan bireylerden daha yüksek yanıt olma olasılıklarına sahip olduklarına dikkat ediniz (Hanbleton, 1991, s.8).

Şekil 2: Sınava Giren İki Grup İçin Yetenek Dağılımları ve Madde Karakteristik Eğrisi



Madde karakteristik fonksiyonunun matematiksel formunda ve / veya modelde belirtilen parametrelerin sayısı ile ilişkili birçok madde tepki modeli vardır. MTK modelleri, maddeyi tanımlayan bir veya daha fazla parametre ve öğrenciyi tanımlayan bir veya daha fazla parametre içerir. Herhangi bir MTK uygulamasındaki ilk adım, bu parametrelerin hesaplanmasıdır. Madde yanıt modelleri, klasik gerçek puan modelinin aksine çürütülebilir modellerdir. Belirli bir madde tepki modeli, belirli bir test verisi seti için uygun olabilir veya olmayabilir; Yani model, verileri yeterince tahmin edemez veya açıklayamaz. Herhangi bir MTK uygulamasında, modelin veriye uyumu değerlendirmek gerekir. Belirli bir MTK modeli ilgilenilen test verilerine uyduğunda, arzu edilen birçok özellik elde edilir. Öğrenci yeteneği tahminleri test bağımlı değildir ve madde indisleri grup bağımlı değildir. Madde yanıt teorisinde madde ve yetenek parametrelerinin değişmez olduğu söylenir. Yukarıda bahsedilen istenen özelliklere ek olarak, klasik test teorisinde olduğu gibi, MTK tüm öğrenciler için tek bir hata tahmininden ziyade bireysel yetenek tahminleri için standart hata tahminleri sunar (Hanbleton, 1991, s.8).

MTK'da kullanılan matematiksel modeller, bir öğrencinin belirli bir maddeyi doğru şekilde cevaplama ihtimalinin, kişinin yetenek ve yeteneklerine ve maddenin özelliklerine bağlı olduğunu belirtmektedir. MTK modelleri, modelin uygulandığı verilerle ilgili bir dizi varsayım içerir. Varsayımların uygulanabilirliği doğrudan belirlenemese de, bazı dolaylı deliller toplanabilir ve değerlendirilebilir. Ve modelin test verilerine genel uyumu da değerlendirilebilir (Hanbleton et al., n.d.ssayfa 9).

En çok kullanılan MTK modellerindeki ortak olan bir varsayım, testi oluşturan maddelerle yalnızca bir yetenek ölçülmesidir. Buna, tek boyutlu olma varsayımı denir. Tek boyutlu olma ile ilgili bir kavram, yerel bağımsızlık kavramıdır. Tüm MTK modellerinde yapılan bir diğer varsayım, belirtilen madde özellik fonksiyonunun gözlemlenemeyen değişkenler (yetenekler) ile gözlenebilir değişkenler (madde yanıtları) arasındaki gerçek ilişkiyi yansıtır olmasıdır. Varsayımlar, bir araştırmacı tarafından bir maddenin performansı ile alakalı olan madde özellikleri olmaksızın yapılır. Genel kullanımdaki MTK modelleri arasındaki başlıca ayırım, sınav yapılan kişinin performansını etkilediği kabul edilen madde özelliklerinin sayısı ve türüdür (Hanbleton, 1991).

Tek Boyutluluk Varsayımı:

Yukarıda belirtildiği gibi, MTK modellerinin ortak bir varsayımı, bir teste bir takım maddeler tarafından yalnızca bir yetenek ölçülmesidir. Bu varsayım tam olarak yerine getirilemez, çünkü birçok bilişsel, kişilik ve test faktörleri daima test performansını en azından bir ölçüde etkiler. Bu etkenler, motivasyon düzeyi, test endişesi, hızlı çalışabilme becerisi, cevaplama şans başarısı eğilimini ve test maddeleri seti ile ölçülen baskın olana göre bilişsel becerileri içerebilir. Tek boyutluluk varsayımı için gereken şey test verilerinin bir kümesi tarafından yeterince karşılanan, test performansını etkileyen "baskın" bir bileşen veya faktörün varlığıdır (Hanbleton et al., n.d., 1991).

Yerel Bağımsızlık Varsayımı:

Yerel bağımsızlık, test performansını etkileyen kabiliyetlerin sabit tutulması durumunda, herhangi bir madde çifti için verilen cevapların istatistiksel açıdan bağımsız olduğunu gösterir. Başka bir deyişle, sınavdaki yetenekleri de hesaba kattıktan sonra, araştırmacıların farklı maddelere verdiği cevaplar arasında bir ilişki yoktur. Basitçe söylemek gerekirse, bu, modelde belirtilen yeteneklerin, öğrencilerin

test maddelerine verdiği yanıtı etkileyen tek faktör olduğu anlamına gelir. Bu yetenekler kümesi tam örtük alanı temsil eder. Tek boyutlu olma varsayımı karşılandığında, örtük alan tamamen tek bir yetenekten oluşur (Hanbleton et al., n.d., 1991).

3.5.3 Madde Yanıt Teorisinde En Çok Kullanılan Modeller

Bir madde karakteristik fonksiyonu veya madde karakteristik eğrisi (MKE), bir maddenin başarı olasılığını (yani, doğru bir yanıt vermek), maddenin karakteristiği ve test ile ölçülen yetenekle ilişkilendiren matematiksel bir ifadedir. Sonsuz sayıda madde yanıt modeli(MYM) tasarlamak mümkündür ancak yalnızca birkaç model geçerlidir. Üç en popüler tek boyutlu madde yanıt modeli, bir, iki ve üç parametrelili lojistik modellerdir. Her biri içerdiği madde parametrelerinin sayısı nedeniyle bu adları almışlardır. Bu modeller, dikotom madde yanıt verisi için uygundur (Hanbleton et al., n.d., 1991, s.12).

Bir Parametrelili Lojistik Model:

Tek parametrelili lojistik model en çok kullanılan MTK modellerinden biridir. Tek parametrelili lojistik modeli olan madde karakteristik eğrileri aşağıdaki Denklem 1 ile verilmektedir.

$$(1) \quad P_i(\theta) = \frac{e^{(\theta - b_i)}}{1 + e^{(\theta - b_i)}} \quad i=1, 2, \dots, n$$

Denklem (1)'de;

$P_i(\theta)$: θ yetenek düzeyinden rasgele seçilen bir öğrencinin i . maddeyi doğru bir şekilde cevaplandırılması ihtimalini,

b_i : i . Maddenin zorluk parametresidir,

n : testteki madde sayısıdır.

e : yaklaşık değeri 2.71 olan bir sonsuz sayıdır (π gibi)

Bir parametrelili model, tahmin faktörünün sıfır olduğunu ve madde ayırıcılık gücünün tüm maddeler için sabit olduğunu varsayar (Hanbleton et al., n.d., 1991, s.12-13).

İki Parametrelili Lojistik Model:

Lord (1952), yığılmalı normal dağılıma (normal ogive) dayalı iki parametrelili bir madde yanıt modeli geliştiren ilk kişiydi. Birnbaum (1968) iki parametrelili lojistik fonksiyonunu iki parametrelili normal ogive fonksiyonları için madde karakteristik fonksiyonunun formülü olarak değiştirmiştir. Lojistik fonksiyonların normal ogive fonksiyonlarına göre daha rahat çalışıldıkları için önemli bir avantajı vardır. Lojistik model, normal ogive modelinden matematiksel olarak daha kolaydır, çünkü lojistik model entegrasyon gerektirir, diğeri madde ve yetenek parametrelerinin açık bir fonksiyonudur ve ayrıca önemli istatistiksel özelliklere sahiptir.

Birnbaum tarafından geliştirilen iki parametrelili lojistik model için madde karakteristik eğrileri aşağıdaki Denklem 2 ile verilmiştir:

$$(2) \quad P_i(\theta) = \frac{e^{D a_i (\theta - b_i)}}{1 + e^{D a_i (\theta - b_i)}} \quad i = 1, 2, \dots, n$$

İki parametre modelinde tahmin faktörü oluşmaz, ancak her bir madde için madde ayırıcılık gücü farklı değerdedir (Hanbleton et al., n.d.sayfa 14-15).

Denklemde;

$P_i(\theta)$: θ yetenek düzeyinden rasgele seçilen bir öğrencinin i . maddeyi doğru bir şekilde cevaplandırılması ihtimalini,

b_i : i . Maddenin zorluk parametresidir

n : testteki madde sayısıdır.

a_i : i . maddenin ayırıcılık gücü parametresidir.

D : Değeri 1,7 olan bir sabittir (ölçekleme sabiti)

Üç Parametrelili Lojistik Model:

Üç parametrelili MKE'nin matematiksel formu,

$$(3) \quad P_i(\theta) = c_i + (1 + c_i) \cdot \frac{e^{D_{a_i}(\theta - b_i)}}{[1 + e^{D_{a_i}(\theta - b_i)}]} \quad i = 1, 2, \dots, n$$

c_i parametresi, MKE'nin y-eksenini kestiği noktayı ve düşük yetenekli bir öğrencinin i . maddeyi doğru yanıtlama olasılığını belirtir.

3.5.4. Madde Karakteristiğinin Bir Fonksiyonu Olarak DMF

DMF, birçok kaynakta benzer şekilde tanımlanmıştır. Bu tanımlara dayanarak DMF, aynı yetenek düzeyinde olan, fakat cinsiyet, sosyoekonomik düzey, etnik köken, inanç vb. gibi farklı gruplardan gelen bireylerin, test maddelerine doğru cevap verme olasılıklarının değişmesi biçiminde sentezlenebilir (Lim & Drasgow, 1990; Raju, 1990a; Mellenbergh, 1989; Tittle, 1988; Adams & Rowe, 1988; Shepard, Camilli, & Williams, 1985; Mellenberg, 1983; Osterlind, 1983; Devine & Raju, 1982). Aynı yetenek düzeyine sahip bireylerin bir maddeye doğru cevap verme olasılığının değişmesinin iki temel kaynağı vardır: a) madde yanlılığı veya b) gerçek bilgi, beceri vb. farklılığı. Maddelerin DMF içerip içermediğini belirleme, yanlılık için daha objektif bir yaklaşım olarak görüldüğünden daha sık kullanılan bir tekniktir. DMF içeren bir maddenin DMF kaynağının ne olduğunu belirlemek için içerik analizi veya uzman kanısına başvurma yollarından yararlanılabilir (Allan S. Cohen and Robert A. Ibarra).

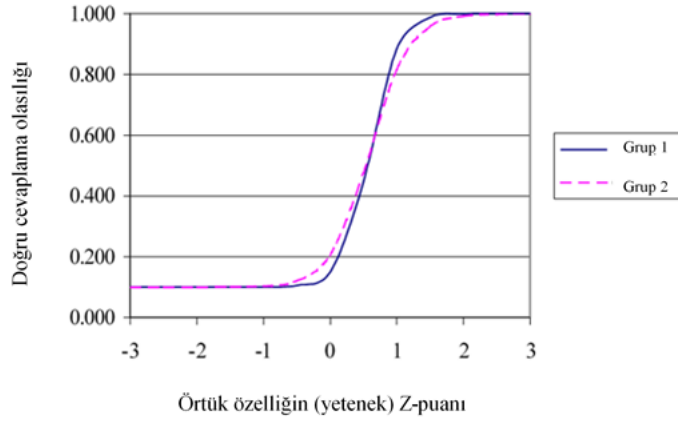
DMF de kendi içinde tek biçimli (uniform) ve tek biçimli olmayan (nonuniform) şeklinde ikiye ayrılmaktadır. Bir maddenin doğru cevaplanma olasılığı, tüm yetenek düzeyleri için bir grubun lehine işlediğinde tek biçimli, farklı yetenek düzeylerinde farklı gruplar lehine işlediğinde tek biçimli olmayan DMF' den söz edilmektedir (Zumbo, 1999). Tekniklerin hemen hepsi tek biçimli DMF'yi belirleyebilmesine rağmen, tek biçimli olmayan DMF'yi belirlemek her teknik için olanaklı değildir (Murphy & Davidshofer, 2005).

3.5.5. DMF Belirleme

DMF' nin tespit edilmesine ilişkin çalışmaların çoğu, toplumsal cinsiyet veya etnik köken gibi açık bir gruba üyelik durumunun yanıt modellerindeki farklılıkların belirlenmesine odaklanmaktadır. Açık veya örtük gruplara dayalı DMF' yi incelemek için standart yaklaşıma bir alternatif, DMF' yi maddelerin yapısal özelliklerinin bir fonksiyonu olarak incelemektir. Bu yaklaşım, yanıt eğilimlerini, testteki maddelerin yapısal özelliklerine yerleştirilen kültürel beklentilerin bir fonksiyonu olarak inceler. Başka bir deyişle DMF, farklı gruplardan gelen bireyler, maddenin ölçtüğü öngörülen yeteneğe eşleştirildikten sonra, maddenin doğru cevaplanma olasılıklarında farklılık görüldüğünde ortaya çıkar (Zumbo, 1999).

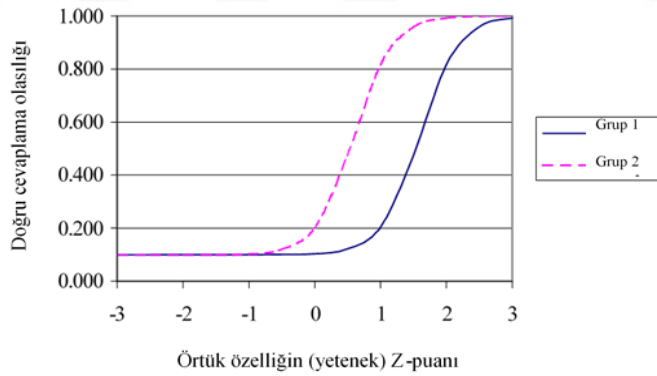
DMF, Modern psikometrinin gelgitinde ortaya çıkan doğal psikometrik bir tekniktir. DMF, varyasyonun sürekliliğini kavramsal olarak akılda tutarak bir madde üzerinde farklı grupların madde karakteristik eğrilerini (MKE) karşılaştırarak değerlendirilir. Aynı maddenin MKE'sinin, araştırmacının değerlendirmesini istediği her grup için (ör. cinsiyet) ayrı ayrı çizildiği düşünülebilir. MKE'leri her grup için özdeş veya çok yakınsa, maddenin DMF görüntüsü vermediği söylenebilir. Bununla birlikte, MKE'leri gruplar arasında birbirinden önemli ölçüde farklıysa, maddenin DMF gösterdiği söylenir. Çoğu bağlamda, DMF, iki MKE'nin yerleşim farkı (yani, zorluk veya eşik) olarak düşünülür, ancak her zaman bu durum olmayabilir. DMF'yi gösteren bazı MKE örnekleri ve DMF'yi göstermeyen bazı madde örnekleri aşağıda sunulacaktır. Şekil 3, DMF görüntülemeyen bir maddeye örnektir. Görüldüğü gibi, eğriler arasındaki alan çok küçük ve her eğri için parametreler neredeyse eşdeğerdir (Zumbo, 1999).

Şekil 3: DMF Görüntülemeyen Bir Madde Örneği



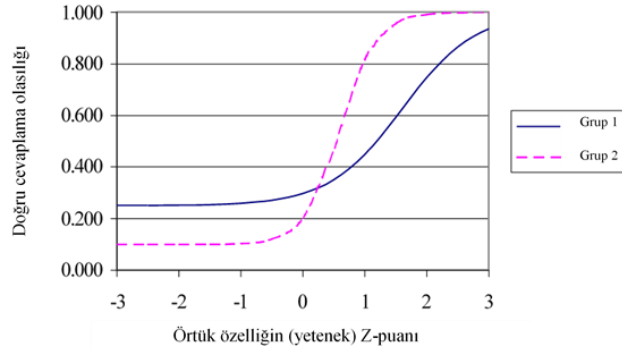
Şekil 4, iki MKE arasında çok büyük bir alana sahip önemli DMF görüntüleyen bir maddeye örnek olarak verilebilir. MKE'leri kesişmediđi için bu tür DMF uniform DMF olarak bilinir. Şekil 4'de gösterilen bu madde, her iki grup için aynı örtük deđişkenin eşdeđer bir ölçütü olmayabilir (Zumbo, 1999).

Şekil 4 : Uniform DMF Görüntüleyen Bir Madde Örneđi



Şekil 5, önemli derecede düzensiz DMF görüntüleyen bir maddeye örnektir (diđer bir deyişle, MKE'leri birbirini keser). Ortalamanın altında (yani $z < 0$) puan alan veya bu ortalamaların altında olan kişiler için Grup 1 tercih edilirken, ortalamanın üstünde puan alanlar için (yani, $z > 0$) Grup 2 tercih edilir, uniform olmayan DMF'yi gösterir (Zumbo, 1999).

Şekil 5: *Nonuniform DMF Görüntüleyen Bir Madde Örneği*



3.5.6. DMF Belirleme Teknikleri

DMF belirleme teknikleri, klasik test kuramına, örtük özellikler kuramına (madde tepki kuramına) veya bazı istatistiksel tekniklere dayanarak açıklanabilir. Klasik test kuramı kapsamında madde ayırıcılık gücü, madde güçlüğü, faktör analizi, varyans analizi, madde güçlük dönüşümü (MGD) vb; örtük özellikler kuramı kapsamında işaretli ve işaretsiz alan indeksleri, Lord'un Ki-kare'si, madde parametreleri ya da en çok olabilirlik oranları farklarının karşılaştırılması vb; bazı istatistiksel tekniklere dayananlar arasında ise Ki-kare, Mantel-Haenszel (MH) test istatistiği ve lojistik regresyon (LR), IRT-LR DIF (Likelihood ratio test), parametre karşılaştırma yöntemi, alan indeksleri tekniği gibi teknikler sıralanabilir (Dorans & Holland, 1993; Zumbo, 1999; Hambleton, Swaminathan, & Rogers, 1991; Lim & Drasgow, 1990; Adams & Rowe, 1988; Raju, 1990b; Mellenbergh, 1989; Mellenbergh, 1989; Price, 1987; Paul W Holland & Thayer, 1986; Hanbleton et al., n.d.; Shepard, Camilli, & Williams, 1985; Hills, 1984; Devine & Raju, 1982); Rudner, Getson, & Knight, 1980).

Araştırmanın üçüncü alt problem cümlesindeki SYT maddelerinin DMF içerip içermediğinin araştırması kapsamında Rasch modeli kullanılmıştır. DMF belirlemek için kullanılan Rasch Modeli aşağıda kısaca açıklanmıştır.

3.5.7. Rasch Analizi Hakkında Genel Bilgiler

Rasch analizi, 1960 yılında Danimarkalı matematikçi Georg Rasch tarafından iki sonuçlu maddelerden oluşan ölçeklerin iç tutarlılığını test etmek için geliştirilmiştir. Daha sonra David Andrich (1978) çoklu sonuçlu maddelerden oluşan ölçme araçlarını test etmek için sıralı sonuçlu modeli(Rating Scale Model) geliştirmiştir. Daha sonra da Geofferey Masters (1982) Kısmi Kredi Model'ini (Partial Credit Model) Rasch modeline dahil etmiştir.

Model, ölçmeyi destekleyen aksiyomları işler hale getirir. Bu aksiyomlar ölçme yapmanın kurallarıdır ve sıralı ölçekler mi, aralıklı ölçekler mi oluşturulduğunu belirler. Rasch modeli, aralık ölçeği ölçümü yapılacaksa maddelere verilen cevaplarda ne beklenmesi gerektiğini gösterir. Modelin iki sonuçlu maddeler için (evet-hayır, doğru yanlış, katılıyorum-katılmıyorum gibi) ve çok sonuçlu maddeler için olan versiyonları mevcuttur, ayrıca maddeleri birden fazla yanıt kategorisine sahip olan ölçekler için de kullanılır.

Bir ankette yer alan ve bir araya getirilmesi amaçlanan bir dizi maddeden elde edilen cevap çiftleri, Guttman ölçeklemesinin olası bir formu olduğu ortaya çıkan model tarafından beklenenlere karşı test edilir. Guttman ölçeklendirmesi, (dikotom durumda) bir kişinin ortalama zorluk derecesindeki bir maddeyi doğru yanıtlamasının ardından ölçek üzerindeki o görevi temsil eden o maddenin altındaki tüm maddelerin(yani, daha kolay) doğru yanıtlanmasını sağlayan kesin bir düzendir. Rasch modeli, daha zor bir görevi temsil eden bir madde doğru yanıtlanırsa, daha kolay görevleri temsil eden maddelerin de doğru yanıtlanabileceğinin yüksek bir olasılık olduğunu söylemek için bunu destekler.

Model, bir maddeyi yanıtlamak için belirli bir cevap verme olasılığının, madde güçlüğü ile cevap verenlerin yetenek düzeyi arasındaki farkın doğrusal bir ölçekte lojistik bir fonksiyon olduğunu varsaymaktadır. Başka bir deyişle, bir kişinin bir maddeyi yanıtlama olasılığı, bir kişinin yeteneği ile madde güçlüğü arasındaki farkın lojistik bir fonksiyonudur. Rasch analizi, sıralı bir puanın, Rasch modelinin beklentilerine uygun veriler verilerek doğrusal, aralık düzeyinde bir değişkene dönüşümünü sağlayarak bu işlemi yapar.

Sonuç olarak, Rasch analizi, aralıklı ölçeklendirmeye geçiş transferinin geçerliliği için gerekli olan çeşitli ölçüm konularına birleşik bir yaklaşım sağlar:

- Geçerli bir toplam ham (sıralı) puan için gerekli olan tek boyutluluk ölçeğinin iç yapı geçerliliğini test eder.
- Aralık düzeyinde ölçeklendirme için gerekli olan maddelerin değişmezliğini test eder (yani, herhangi bir madde çifti arasındaki zorluk oranı, katılımcıların yetenek düzeyleri boyunca sabit kalır),
- Uygun kategori sıralaması yapar (çok sonuçlu maddelerin kategori sıralamasının beklendiği gibi çalışıp çalışmadığı),
- Değişen madde fonksiyonu inceler (DMF; örneklemdaki alt gruplar arasında bir madde için yanlılığın olup olmadığı) (Conaghan, 2007).

Temel Rasch varsayımları şöyledir:

- Her birey bir yetenek düzeyi ile karakterize edilir.
- Her madde bir güçlük düzeyi ile karakterize edilir.
- Ve bu değerler bir satır boyunca sayılarla ifade edilebilir.
- Bu sayılar arasındaki farktan, herhangi bir belirli puanlı yanıt gözlemlene olasılığı hesaplanabilir (Tennat&Conaghan, 2007).

3.5.7.1. Rasch Model ve Madde Yanıt Modeli

MYM(Madde Yanıt Modeli) ve Rasch Modelin hesaplamalar açısından birbirlerine benzer olduğu bilinse de, felsefi temellerinin birbirinden oldukça farklı olduğunu ifade etmek önemlidir. MYM uygunluğa dayalı bir model iken Rasch Modeli sadeliğe dayalı bir modeldir. Madde yanıt modelleri üç parametrelidir kadar kullanılabilirken, Rasch modeli yalnızca bir parametreye dayalıdır (madde güçlüğü).

MYM doğası gereği betimseldir çünkü modeli dataya uydurmayı amaçlar. Tersine Rasch kurala dayalıdır, datayı modele uydurmaya önem verir.

Klasik Test Teorisinde aynı öğrencilerin test puanları bir testten diğerine test gücüne bağlı olarak değişirken, MYM'de madde güçlük saptaması örneklemden bağımsızdır ve öğrenci yetenek düzeyi de maddelerden bağımsızdır.

MYM'nin en güzel özelliklerinden biri, madde ve öğrenci özelliklerinin aynı skalada gösterilebilmesidir. Buna "logit" denir. Logit'i açıklamadan önce "olasılık oranı" (odd ratio) kavramını açıklamak gerekir. Madde boyutu için olasılık oranı, istenmeyen durumların sayısının ($q=1-p$) istenen durumların sayısına (p) oranıdır.

Olasılık oranı (odd ratio) ise istenmeyen çıktılardan olasılığının, istenen çıktılardan olasılığına oranı olarak kavramsallaştırılabilir. Olasılık oranı,

Odds = $\frac{1-p}{p}$ şeklinde ifade edilebileceği gibi olasılıklar ve odds arasındaki ilişki şöyledir:

$$P = \frac{odds}{1+odds}$$

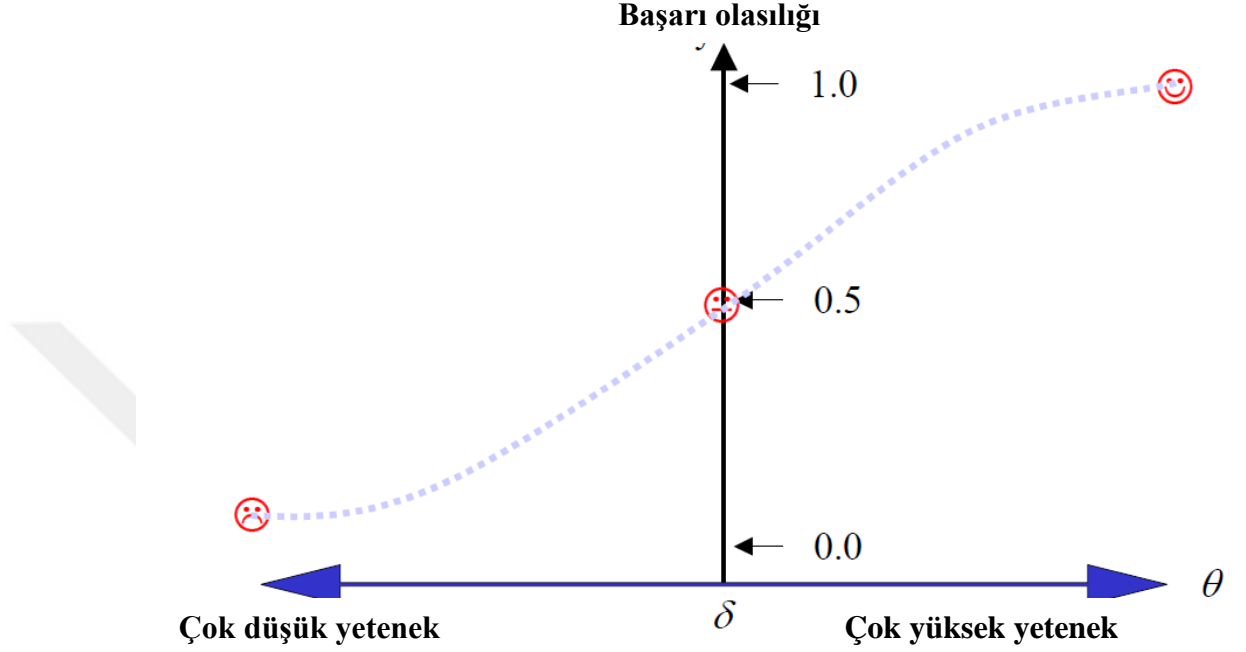
Logit ise olasılık oranının doğal logaritma fonksiyonudur.

$$\text{Logit} = \text{Log}(\text{odds})$$

MTK'de öğrenci ve madde özelliklerini aynı skalaya koyabileceğimizi söylemiştik. İşin püf noktası iki ölçekten gelen değerleri tek bir ortak ölçüğe çevirmektir. Genelde ham puanlar arasındaki farkın karşılaştırılabilir olduğu yanılığımıza düşeriz. Aynı şekilde bir ölçekteki aralıklandırma başka bir ölçekteki aralıklandırma ile karşılaştırılmaz. Logit yardımıyla yeniden ölçeklemek bu iki sorunu da çözer.

Madde yanıt modelleri, öğrencinin bir maddeye yanıt verme olasılığını öğrencinin "yetenek" düzeyinin bir fonksiyonu olarak modellemek için genellikle matematiksel bir fonksiyon uygular. Madde karakteristik eğrisi olarak bilinen bu olasılık fonksiyonu, tipik olarak aşağıdaki şekilde gösterildiği gibi "S" şekline sahiptir (Wu & Adams, 2007).

Şekil 6: Bir Madde için Madde Karakteristik Eğrisi



İncelenen özellik düzeyi θ ve madde zorluk düzeyi β olmak üzere bu olasılık aşağıdaki gibi hesaplanmaktadır:

$$P(X=1) = \frac{\exp(\theta - \beta)}{1 + \exp(\theta - \beta)}$$

X, maddedeki başarı veya başarısızlığı gösteren rastgele bir değişkendir. X = 1, maddedeki başarıyı (veya doğru bir yanıtı) ve X = 0, maddedeki başarısızlığı (veya yanlış bir yanıtı) gösterir (Wu & Adams, 2007).

Logit'in madde ve kişi özellikleri için uygulanma konsepti farklıdır. Kişiler için olasılık oranı:

$p/1-p$ iken madde için olasılık oranı $1-p/p$ dir.

Winsteps isimli IRT yazılımında madde güçlük parametresi ve öğrenci yetenek düzeyi(θ) logit ölçeğinde gösterilir ve bu ikisi arasındaki ilişki Madde-kişi haritasında gösterilir (Yu, C.H., 2017).

3.5.7.2. Rasch Modelinin Varsayımları

Rasch modelinin varsayımları şöyledir:

- 1) Tek boyutluluk: Ölçme aracındaki tüm maddeler tek bir yapıyı ölçmelidir.
- 2) Yerel Bağımsızlık: Yetenek düzeyine bağlı olarak, bir maddeye verilen yanıtlar aynı ölçme aracındaki diğer maddelere verilen yanıtlardan istatistiksel olarak bağımsızdır.
- 3) Örtük özellik monotonluğu: Bir kişinin bir maddeyi doğru yanıtlama olasılığı, yetenek düzeyi arttıkça monoton olarak artar.
- 4) Kesişmeyen madde yanıt fonksiyonları(MYF): Değişmezlik özelliğini elde etmek için MYF birbiriyle kesişmemelidir.

3.5.7.3. Rasch Modelinin Özellikleri

Rast(2001), Rasch modelini bir veya daha çok sayısal değişkeni ölçmeyi amaçlayan madde yanıt modeli olarak tanımlamıştır. Rasch modelinin özellikleri şöyle sıralanabilir:

- 1) Yeterlik/Uygunluk: Yeterlilik özelliğinin anlamı maddeler ve kişiler için yanıtların toplamının, yetenek ve güçlük kestirimi için yeterli bir istatistik olmasıdır. Diğer bir deyişle toplam puanların, bireyin ölçülen özellik üzerindeki yetenek düzeyi hakkında tüm bilgiyi içermesidir.
- 2) Ayrılabilirlik: Ayrılabilirlik özelliğinin anlamı, madde parametrelerini kestirmek, kişi parametrelerini kestirmeden ya da bilmeden de yapılabilir. Ya da tam tersi de geçerlidir.
- 3) Tarafsızlık: Bu özellik değişmez karşılaştırmalar için bir gereksinimdir. Bunun anlamı kişinin yetenek düzeyi kestirimi ölçme için hangi belirli maddelerin kullanıldığından bağımsız olması gerektiğidir. Aynı şekilde madde güçlük kestirimi, ölçmeye katılan bireylerden bağımsız olmalıdır. Bu özellik bireylerin karşılaştırmasının maddelerden bağımsız olarak ve madde karşılaştırmasının bireylerden bağımsız olarak yapılmasına izin verir.

4) Örtük katılabilirlik: Bu özellik kişi ve madde parametrelerinin toplama ya da çıkarma yoluyla birbirine bağlanması anlamına gelmektedir (Smith, 2019 n.d.).

Bu çalışmada cinsiyete göre sayısal yetenekte farklılaşma olup olmadığı araştırıldığından, kızlar ve erkekler arasındaki farkın maddeler boyunca aynı olup olmadığı veya cinsiyetin maddeler boyunca farklılaşıp farklılaşmadığına bakılmıştır. Bunun için Rasch modeli çerçevesinde Winsteps programındaki analizler sonucunda öğrenci DMF grafikleri incelenerek cinsiyete göre DMF içeren maddeler tespit edilmiş ve bu DMF'lerin büyüklükleri ile manidarlık dereceleri rapor edilmiştir.

Araştırmanın ikinci alt problem cümlesindeki, SYT'nin Rasch modelinin varsayımlarını karşılayıp karşılamadığının araştırılması doğrultusunda aşağıdaki çalışmalar yapılmıştır:

1) Model Seçimi:

Ölçekteki maddeler yalnızca 2 yanıt seçeneğine sahip olduğunda, iki boyutlu model seçilir. Maddeler 3 veya daha fazla seçeneğe sahipse, Rasch modeli, Andrich Derecelendirme Ölçeği Modeli veya Master Kısmi Kredi Modeli olarak bilinen biraz farklı bir biçim alır. Bu modeller aynı Rasch modelini kullanır, ancak matematiğinde biraz farklılık gösterir ve onları geliştirenlerin ismini alır. İkisi arasındaki temel fark, çoklu yanıt maddelerden oluşan ölçeğin eşikler arasındaki mesafenin maddeler arasında eşit olmasını beklenmesidir. Bir eşik, her iki bitişik kategori arasındaki olasılıksal orta noktadır (yani, 50/50). Bu, örneğin, kategori 1 ve 2'yi ayıran eşikler arasındaki metrik mesafe ile kategori 2 ve 3'ü ayıran metrik mesafenin tüm maddeler arasında aynı olduğu anlamına gelir (Conaghan, 2007). Bu çalışmada maddeler iki yanıt seçeneğine sahip olduğundan iki boyutlu Rasch modeli seçilmiştir.

2) Modele Uyum Testleri:

Gözlemlenen yanıt ile modelin beklediği yanıt arasındaki farka vurgu yapıldığında, uygun istatistiklerin çoğu ki-kare temellidir. WINSTEPS'de bunlara INFIT ve OUTFIT istatistikleri denir. WINSTEPS'in standartlaştırılmış uyum istatistikleri de vardır(ZSTD), gözlemlenen ve beklenen değerler arasındaki tüm farkların tüm bireyler üzerinde toplandığı standartlaştırılmış bir toplamdır. Her istatistik gözlemlenen ve beklenen cevap arasındaki fark hakkında biraz farklı bilgiler verir. Örneğin, INFIT, kişinin yetenek seviyesinin yakınında zorluk seviyesine sahip olan maddeler için gözlemlenen ve beklenen yanıtlar arasındaki farkı özellikle ifade

eder. OUTFIT, madde güçlüğünün, kişinin yeteneğinden ne kadar uzakta olduğuna bakılmaksızın, tüm maddeler için farklılıkları içerir. Bu nedenle, birincisi, kişinin yetenek seviyesine yakın olan maddelere verilen yanıtlara daha fazla ağırlık vermesi nedeniyle, ağırlıklı bir uyum istatistiğidir. Sonuç olarak, belirli bir madde için, birkaç ki-kare hesaplanır (grupların sayısı ve örneklem büyüklüğüne göre değişir) ve daha sonra bu ki-kare değerleri, madde için genel ki-kare değerini vermek üzere toplanır ve serbestlik dereceleri grup sayısı-1 olarak hesaplanır. Değer 0,05'ten küçükse (veya Bonferroni ayarlı bir değer) ise, madde modelin beklentisine uygun değildir.

INFIT ve OUTFIT Meansquare değerleri şu şekilde yorumlanır:

>2.0 ise ölçmeyi bozar veya değerini aşağıya indirir.

1.5 < <2.0 ise ölçmenin yapısı için verimsizdir, işe yaramaz.

0.5 < <1.5 ise ölçme için verimlidir.

<0.5 ise ölçme için daha az verimlidir. Ama ölçmenin değerini indirmez.

Yanıtıcı olarak daha iyi güvenilirlik ve ayırışma üretebilir.

Madde uyumuna ek olarak, kişilerin uygunluğunun incelenmesi önemlidir. Tuhaf yanıt çiftleri (yüksek pozitif artıklarla tanımlanır) olan az sayıda katılımcı madde düzeyinde uyumu ciddi şekilde etkileyebilir. Bu tür anormal yanıt kalıpları, kayıt edilmemiş eşanlılıktan veya bilişsel açıkları olan yanıtlayıcılardan kaynaklanabilir. Bu nedenle, bazı katılımcıların bu şekilde uyuşmadığı durumlarda, analizden çıkarılması ölçeğin iç yapı geçerliliği için önemli bir fark yaratabilir (Conaghan, 2007).

3)Değişen Madde Fonksiyonu Testleri:

DMF veya madde yanlılığı da modele uyumu etkileyebilir. Bu olduğunda örneklem içindeki farklı gruplar (örneğin, genç ve yaşlı insanlar, kız ve erkekler), ölçülen temel özelliklerin eşit seviyelerine rağmen, bir maddeye farklı şekilde yanıt verir. İki tip DMF tanımlanabilir. Birincisi, grubun ölçülen öznelik aralığı boyunca, bir maddeye verdikleri cevaplarda tutarlı bir sistematik farklılık göstermesidir; uniform DMF olarak adlandırılır. Gruplar arasındaki farklarda tek biçimlilik olmadığında (örneğin, farklılıklar nitelik seviyelerine göre değişir), o zaman buna uniform olmayan DMF denir (Conaghan, 2007).

4) Kişi ve Maddelerin Hedeflenmesi ve Örneklem Büyüklüğü:

Rasch analizi, ölçeğin ortalama zorluk derecesini temsil eden maddeyi daima sıfır logit üzerinde merkezlenmiştir. Maddeler için belirlenen sıfır değerine sahip kişiler için elde edilen ortalama konum puanlarının karşılaştırılması, maddelerin örneklemedeki kişiler için ne kadar iyi hedeflendiğinin bir göstergesidir. İyi hedeflenmiş bir ölçme için (çok kolay değil, çok zor değil), kişilerin ortalama konumu da sıfır değerinde olacaktır. Kişiler için pozitif bir ortalama değer, bir bütün olarak örneklemin ölçeğin ortalamasından daha yüksek bir seviyede bulunduğunu belirtirken, negatif bir değer bunun tam tersini belirtir (Conaghan, 2007).

5) Yanıt Bağımlılığı Ve Tek Boyutluluk:

Rasch modellerinin temelini oluşturan yerel bağımsızlık varsayımı, Rasch faktörünü, yani ana ölçeği çıkardıktan sonra, artıklarda artık kalıpların olmaması gerektiği anlamına gelir. Maddelerin yerel bağımsızlığı varsayımında bir ihlal, yanıt bağımlılığı ve çok boyutlulukla 2 şekilde bulunabilir. Yanıt bağımlılığı, maddelerin bir şekilde bağlandığı, bir maddedeki cevabın başka bir cevabı belirleyeceği şekildedir. Bu tip maddeler klasik güvenilirliği şişirir ve Rasch analizinde parametre tahminini etkiler.

Tek Boyutluluk Varsayımının Karşılanması

Lumsden tek boyutlu test oluşturmak için yaptığı inceleme sonucunda, test oluşturanlara deneysel kanıtlar üzerinde seçilen test modellerinin ilk havuzunu oluşturmalarını önermiştir. Bu şekildeki bir madde seçim prosedürü madde havuzu içindeki tek boyutlu bir madde kümesi bulunması olasılığını arttıracaktır. Eğer test maddelerine önseçim yapılmazsa, madde havuzu tek boyutlu bir madde kümesinin ortaya çıkması için çok heterojen olacaktır. Lumsden'in metodunun içinde madde faktör analizi uygulanır ve faktör çözümlenmesinden elde edilen baskın faktörü ölçmeyen maddeler çıkarılır. Kalan maddelere faktör analizi yapılır ve tekrar sapan maddeler çıkarılır. Bu işlem istenen çözüm elde edilene kadar devam ettirilir. Kümeleşme (yakınsama-convergence) ilk madde havuzu dikkatlice seçildiğinde ortak bir özelliği ölçüyor görünen maddeleri içerecektir. Lumsden, ilk faktör analizi varyansının ikinci faktör analizi varyansına oranı "tek boyutluluk indeksi" olarak kullanmayı önermiştir (Ronald K. Hambleton & Swaminathan, n.d., 1991, s.21-22).

Yerel Bağımsızlık Varsayımının Karşılanması

Tek boyutluluk varsayımına eşit olan ve "Yerel bağımsızlık" varsayımı olarak bilinen bir varsayım vardır. Bu varsayım bir öğrencinin bir testteki farklı maddelere verdiği cevaplar istatistiksel olarak bağımsızdır der. Eğer $U_i (i=1,2,3,\dots,n)$ bir öğrencinin n test maddesine verdiği yanıtları gösteriyorsa, bir öğrenci tarafından i . Maddeye doğru cevap verme olasılığı P_i ve $Q_i = 1 - P_i$ ise yerel bağımsızlık varsayımı aşağıdaki duruma öncülük eder:

$$\text{Prob}(U_1=u_1, U_2=u_2, \dots, U_n=u_n, | \theta) = \text{Prob}(U_1=u_1 | \theta) \cdot \text{Prob}(U_2=u_2 | \theta) \dots \text{Prob}(U_n=u_n | \theta).$$

Eğer $P_i(\theta) = \text{Prob}(U_i=1 | \theta)$ ve $Q_i(\theta) = \text{Prob}(U_i=0 | \theta)$ olarak alınırsa,

$$\text{Prob}(U_1=u_1, U_2=u_2, \dots, U_n=u_n, | \theta)$$

$$= P_1(\theta)^{u_1} \cdot Q_1(\theta)^{1-u_1} \cdot P_2(\theta)^{u_2} \cdot Q_2(\theta)^{1-u_2} \dots P_n(\theta)^{u_n} \cdot Q_n(\theta)^{1-u_n}$$

$$= \prod_{i=1}^n P_i(\theta)^{u_i} \cdot Q_i(\theta)^{1-u_i} \quad (4)$$

elde edilir.

Yerel bağımsızlık varsayımının bir sonucu, sabit bir θ yetenek seviyesindeki öğrencilerin test puanlarının sıklığı olarak aşağıdaki gibi verilir:

$$\begin{aligned} f(x | \theta) \\ = \sum_{\sum u_i = x} \prod_{i=1}^n P_i(\theta)^{u_i} \cdot Q_i(\theta)^{1-u_i} \end{aligned} \quad (5)$$

Burada x bir öğrencinin test puanıdır ve 0 ile n arasında değer alır.

Yerel bağımsızlık varsayımı θ 'nın tek boyutlu olduğu ve örtük alanın tek boyutlu olduğu varsayımlarının eşdeğer olması durumunda geçerlidir. İlk olarak ortak bir özellik ölçen bir test maddeleri kümesi ve sabit bir θ yetenek düzeyindeki öğrenciler için madde yanıtlarının istatistiksel olarak bağımsız olduğu düşünülmüş olsun. Sabitlenmiş yetenek düzeyi olan θ için maddeler istatistiksel olarak bağımsız değillerse, bu bazı öğrencilerin aynı yetenek düzeyindeki diğer öğrencilerden daha yüksek beklenen test puanı alacaklarını gösterir. Sonuç olarak öğrencilerin test puanlarını açıklamak için bir yetenekten daha fazlası gerekli olacaktır. Bu da maddelerin tek boyutlu olması gerektiği yönündeki orijinal varsayımın kirlenmesi demektir. İkinci olarak, yerel bağımsızlık varsayımı sabit bir yetenek düzeyindeki öğrenciler için madde yanıtlarının bağımsız olduğunu belirtir. Bu yüzden test maddeleri arasındaki ilişkiyi açıklamak için sadece tek bir yetenek gereklidir. Şunu farketmek önemlidir: Yerel bağımsızlık varsayımı test maddelerinin öğrenci gruplarının tümü üzerinde korelasyonsuz olduğunu belirtmez (Lord&Novick, 1968, s.361). Madde çiftleri arasındaki pozitif korelasyon test maddeleriyle ölçülen yetenekteki öğrenciler arasındaki farklılıklardan ileri gelir. Ama madde puanları yetenek düzeyi ile korelasyonsuzdur.

Yerel bağımsızlık ve örtük alanın tek boyutluluğu eşit varsayımlar olduğundan, yerel bağımsızlık varsayımına uyan test maddeleri kümesinin içeriği faktör analiz tekniklerinin kullanımında çalışılabilir. Aynı yetenek düzeyindeki öğrenciler için madde yanıtlarının istatistiksel bağımsızlığının kaba bir denetlemesi Lord tarafından önerilmiştir (1953a). Lord, dar bir yetenek aralığındaki öğrenciler için öğrencilerin madde yanıtlarını göz önüne almış ve her madde çifti için (kaykare) istatistiği ile madde yanıtlarının bağımsızlığının bir ölçüsünü sağlanabileceğini önermiştir. Eğer her (00,01,10,11) şeklindeki madde yanıt çiftinden elde edilen öğrencilerin oranları, öğrenci grubunun uç noktalarındaki bireyler için kestirilebilirse, iki madde üzerindeki madde yanıtları bağımsız olur. χ^2 (kaykare) değeri her madde çifti için hesaplanabilir ve anlamlılık için test edilebilir. Yetenek düzeylerinin farklı bölgelerinde olan tüm öğrenciler için bu süreç tekrarlanır (Ronald K. Hambleton & Swaminathan, n.d., 1991, s.23-24).

BÖLÜM IV

BULGULAR

SYT ölçeğinin geliştirilmesi çalışması kapsamında aşağıdaki çalışmalar yapılmıştır.

4.1. Örneklem Büyüklüğü

Bir madde analiz çalışmasında kullanılacak asgari öğrenci sayısı için kesin bir kural yoktur. Kuşkusuz, Lisansüstü Kayıt Sınavı veya ticari olarak yayınlanmış bir yetenek veya başarı testi gibi yaygın olarak kullanılacak bir test için madde analizi, belki de binlerce katılımcısı olan oldukça büyük, temsili bir örnekleme dayanmalıdır. Uzun zamandır uygulanan bir kural ise (Nunnally, 1967) cevaplayıcı sayısının madde sayısının 5 ile 10 kat fazlası olmasıdır. Test geliştiricisi bu kurala dayanıyorsa, en azından 20 madde ve 100 cevaplayıcı kullanmalıdır. Madde Tepki Teorisine dayanan parametre tahmini için gerekli olan örneklem boyutları, seçilen modele bağlı olarak 200 ile 1000 kişi arasında değişebilir (Crocker & Algina, 1986).

Nihai testin oluşturulması için hazırlanan Test1 366 11. Sınıf öğrencisine, Test2 ise 355 11. Sınıf öğrencisine uygulanmıştır. Örneklem büyüklüğü kuralına baktığımızda, test $20 \times 5 = 100$ cevaplayıcıdan fazla öğrenciye uygulanmış olup, örneklem büyüklüğü parametre hesaplamaları için yeterlidir.

Madde ve test istatistikleri elde edilmiş ve madde seçimleri yapıldıktan sonra nihai test oluşturulmuştur. Nihai test 1382 11. Sınıf öğrencisine uygulanmıştır. Uygulamanın yapılacağı öğrenci grubu Anadolu Liselerine %8 lik dilim ve üzerinde girmiş olan öğrencilerden seçilmiştir.

4.2. Maddelerin Analizi

Madde analizi, madde cevap dağılımının herhangi bir istatistiksel özelliğinin hesaplanması ve incelenmesidir. Dikotom olarak puanlanmış maddeler için en iyi bilinen tanımlayıcı madde güçlüğüdür (p), bu da maddeye doğru cevap verenlerin oranını ifade eder. Madde p-değerleri mutlaka bir madde kalitesi endeksi değildir;

bununla birlikte, norm referanslı testler için, .50 civarında p değerine sahip maddeler toplam puan farkının ve dolayısıyla güvenilirliğin en üst düzeye çıkarılmasına izin verir. Çoktan seçmeli maddeler kullanıldığında, .50'den büyük p değerleri toplam puan güvenilirliğini en üst düzeye çıkarmak için idealdir.

Madde ayıricılık gücü, bir madde üzerindeki performansın diğer bazı kriterlerdeki performansla ne kadar ilişkili olduğunu değerlendirmek için kullanılır. Bu kriterler, testteki toplam puan veya başka bir değişken olabilir. Toplam test puanı kullanıldığında, yüksek düzeyde ayırt edici maddelerin seçimi, test puanlarının iç tutarlılığının artmasına neden olur.

Madde güvenilirlik endeksi, madde standart sapmasının ve toplam test puanıyla madde korelasyonunun çarpımıdır ve madde geçerlilik endeksi, dış puanla madde puan standart sapmasının ve madde puan korelasyonunun çarpımıdır.

Madde seçiminin amacı, ilgili kritere göre ayırım yapan maddeleri seçerek test puanının güvenilirliğini veya geçerliliğini arttırmak olduğunda, bazen madde güvenilirlik endeksinin (veya madde geçerlilik endeksinin) madde ve kriter arasında basit korelasyon yerine kullanılması gerektiği önerilmiştir. Çünkü madde varyansı aslında belirli bir maddenin genel test puanının güvenilirliğine veya geçerliliğine olan katkısını arttırmaktadır (Crocker & Algina, 1986). Standart sapma ve varyans, sapma puanlarına dayanmaktadır. Sapma puanlarının mutlak değerleri ne kadar büyük olursa, standart sapma ve varyans o kadar büyük olacaktır (Crocker & Algina, 1986).

Madde güçlük katsayısı (p): Madde güçlük katsayısı "0" ile "1" arasında değerler alır; "1" e yaklaştıkça madde kolaylaşır, "0" a yaklaştıkça madde zorlaşır (Turgut & Baykul, 2011).

$$p = \frac{n(D)}{N}$$

p: Madde güçlük katsayısı

n(D): Maddeye doğru yanıt verenlerin sayısı

N: Testi alan toplam öğrenci sayısı

Madde ayıricılık gücü (madde geçerliliği): Madde ayıricılık gücü, maddenin ilgili davranışa sahip olanla olmayanı ne ölçüde ayırdığını gösterir. Madde ayıricılık gücü katsayısı bir korelasyon katsayısıdır ve (-1) ile (+1) arasında değerler alır. Teste konulacak maddeler için aşağıdaki değerler genel olarak ölçüt olarak kabul edilir.

Madde ayırıcılık gücü,

- 0.19 ve daha küçük olan maddeler teste konulmaz.
- 0.20-0.29 arasında olanlar teste düzeltilerek konulabilir.
- 0.30 ve daha büyük olan maddeler teste aynen konulabilir (Baykul, 2000) (Turgut & Baykul, 2011).

Madde standart sapması: Madde güçlük katsayısı(p) ile bu katsayının 1'den farkı olan $q(1-p)$ 'nin çarpımının kareköküne eşittir (Turgut & Baykul, 2011).

$$s = \sqrt{p \cdot q}$$

Bir maddenin güçlük indeksinin 0 ya da 1 olması durumunda o maddenin varyans ve standart sapması da 0 olur. Madde güçlük indeksleri uçlardan, yani 0 ve 1 den, 0,50 (orta güçlük düzeyi) değerine yaklaştıkça madde varyansı ve standart sapması da artar. Başka bir ifade ile, madde güçlük indeksinin 0,50 olması durumunda madde varyansı en büyük değeri olan 0,25'e, madde standart sapması da en en büyük değeri olan 0,50'ye ulaşır. Bir maddenin varyansı ya da standart sapması büyüdükçe, o madde ile ölçülmek istenilen özellik bakımından bireyler arasındaki farklılıkların ortaya konulma gücü artar. Başka bir deyişle madde varyansı ve standart sapması büyüdükçe, maddenin ölçülen özelliğe sahip olanlarla olmayanları birbirinden ayırt etme gücü artar. Bu bakımdan, eğer bir test, testi alan bireyleri belli bir özellik bakımından ayırt etmek amaçlı kullanılacaksa, madde varyansı ve standart sapması büyük olan, yani madde güçlük indeksi 0,50 civarındaki maddelerden oluşmalıdır (Atılğan, Kan, & Doğan, 2009).

Test1 uygulamasında Eşitlik Kurma ve Sayısal Akıl Yürütme alt testleri bulunmaktadır. Test 2 uygulamasında ise Görsel-Uzamsal Yetenek ve Geometri Testleri bulunmaktadır. Bu testlerde ölçülmesi hedeflenen davranışlar ve bu davranışları ölçen madde numaraları aşağıdaki tabloda verilmiştir.

Tablo 5: Test 1 ve Test 2'nin Davranış Dağılımı

| EŞİTLİK KURMA TESTİ | DAVRANIŞ |
|---|---|
| Davranış No 1 | Sayılarla İşlemler |
| Davranış No 2 | Denklem Kurma |
| Davranış No 3 | Problem Çözme |
| SAYISAL AKIL YÜRÜTME TESTİ | DAVRANIŞ |
| Davranış No 4 | Eşitlik Kurma |
| Davranış No 5 | Akıl yürütme |
| Davranış No 6 | Denklem kurma-problem çözme |
| GÖRSEL UZAMSAL YETENEK TESTİ | DAVRANIŞ |
| Davranış No 7 | Döndürme |
| Davranış No 8 | Katlama-Simetri |
| Davranış No 9 | Simetri-Öteleme |
| Davranış No 10 | Görsel Yetenek |
| GEOMETRİ TESTİ | DAVRANIŞ |
| Davranış No 11 | Geometrik Şekillerin Çevre ve Alan Hesabı |
| Davranış No 12 | Benzerlik |
| Davranış No 13 | Açılar |

Test 1 Eşitlik Kurma alt testi (ilk 10 soru) ve Sayısal Akıl Yürütme alt testinden(11.-20. Sorular) oluşmuştur. Test1'i 366 11. Sınıf öğrencisi cevaplamıştır.

Test 2 ise Görsel-Uzamsal Yetenek alt testi (ilk 10 soru) ve Geometri alt testinden (11.-20. Sorular) oluşmuştur. Test2'yi ise 355 11. Sınıf öğrencisi cevaplamıştır.

Doğrudan teste konabilecek maddelerin belirlenmesi, teste konulması mümkün olmayan maddelerin elenmesi amacıyla Test 1 ve Test 2'nin madde analizleri excel programında yapılarak sonuçlar aşağıda verilmiştir.

Aşağıdaki tabloda madde güçlük indeksleri, madde ayırıcılık gücü (madde geçerlik katsayısı), madde standart sapması istatistikleri yer almaktadır.

Tablo 6: Test 1 ve Test 2' nin Madde İstatistikleri

| Davranış No ve Davranış adı | Madde No | Madde Güçlüğü | Madde Ayırıcılık Gücü | Madde Standart Sapması |
|-----------------------------|------------|---------------|-----------------------|------------------------|
| Test1 1 | 1 | 0,19 | 0,17 | 0,40 |
| Sayılarla İşlemler | 2 | 0,53 | 0,43 | 0,50 |
| | 6 | 0,77 | 0,39 | 0,42 |
| | Test1 2 | 3 | 0,49 | 0,54 |
| Denklem Kurma | 4 | 0,35 | 0,36 | 0,48 |
| | 9 | 0,17 | 0,14 | 0,37 |
| | 10 | 0,20 | 0,20 | 0,40 |
| Test1 3 | 5 | 0,37 | 0,32 | 0,48 |
| Problem Çözme | 7 | 0,68 | 0,46 | 0,47 |
| | 8 | 0,10 | 0,14 | 0,30 |
| Test1 4 | 11 | 0,57 | 0,61 | 0,49 |
| Eşitlik Kurma | 15 | 0,22 | 0,30 | 0,42 |
| | 16 | 0,43 | 0,40 | 0,49 |
| Test1 | 12 | 0,65 | 0,47 | 0,48 |

| | | | | |
|------------------|----|------|------|------|
| 5 | 13 | 0,34 | | |
| Akıl yürütme | | | 0,31 | 0,47 |
| | 14 | 0,76 | 0,37 | 0,43 |
| Test1 | 17 | 0,48 | 0,58 | 0,50 |
| 6 | 18 | 0,35 | 0,59 | 0,47 |
| Denklem kurma- | 19 | 0,34 | 0,65 | 0,47 |
| problem çözme | 20 | 0,31 | 0,64 | 0,46 |
| Test2 | | | | |
| 7 | 1 | 0,55 | 0,35 | 0,50 |
| Döndürme | 2 | 0,72 | 0,33 | 0,45 |
| | 3 | 0,80 | 0,29 | 0,40 |
| Test2 | | | | |
| 8 | 4 | 0,77 | 0,36 | 0,42 |
| Katlama-Simetri | | | | |
| | 5 | 0,33 | 0,49 | 0,47 |
| | 10 | 0,84 | 0,31 | 0,36 |
| Test2 | | | | |
| 9 | 6 | 0,17 | 0,01 | 0,38 |
| Simetri-Öteleme | | | | |
| | 9 | 0,76 | 0,40 | 0,43 |
| Test2 | | | | |
| 10 | 7 | 0,56 | 0,18 | 0,50 |
| Görsel Yetenek | | | | |
| | 8 | 0,69 | 0,49 | 0,46 |
| Test2 | | | | |
| 11 | 11 | 0,63 | 0,51 | 0,48 |
| Geometrik | | | | |
| Şekillerin Çevre | 15 | 0,14 | 0,13 | 0,35 |
| ve alan hesabı | 16 | 0,27 | 0,31 | 0,45 |

| | | | | | |
|-----------|----|----|------|------|------|
| Test2 | 12 | 12 | 0,45 | 0,25 | 0,50 |
| Benzerlik | | 13 | 0,60 | 0,38 | 0,49 |
| | | 20 | 0,67 | 0,47 | 0,47 |
| Test2 | 14 | 14 | 0,36 | 0,25 | 0,49 |
| 13 | 17 | 17 | 0,16 | 0,12 | 0,37 |
| Açılar | 18 | 18 | 0,29 | 0,40 | 0,45 |
| | 19 | 19 | 0,13 | 0,09 | 0,34 |

Yukarıdaki tabloda koyu gri renk ile işaretlenen sorular güçlük ve ayırdedicilik bakımından istenen özelliklere sahiptir. Ayrıca standart sapma değerleri de 0,50 ye yakındır.

Test 1'deki ilk 10 madde Eşitlik Kurma alt testine ait olup bu testteki 2.,3.,4.,6. ve 7. maddelerin seçilmesi uygun görülmüştür. 11. İle 20. Maddeler arasındaki sorular Sayısal Akıl Yürütme alt testine ait olup bu testten 11.,17.,18.,19., ve 20. maddelerin seçilmesi uygun görülmüştür.

Test 2'de ise ilk 10 madde Görsel-Uzamsal Yetenek alt testine ait olup bu testteki 1.,2.,5.,9. ve 10. maddelerin seçilmesi uygun görülmüştür. 11. İle 20. Maddeler arasındaki sorular Geometri alt testine ait olup 11.,13.,16.,18., ve 20. maddelerin seçilmesi uygun görülmüştür.

Test 1 ve Test 2'nin genel test istatistikleri ise aşağıdaki tabloda verilmiştir.

Tablo 7: Test 1 ve Test 2 nin Genel Test İstatistikleri

| | Madde sayısı | Cevaplayıcı sayısı | KR-20 | Madde Güçlüğü Ortalaması | Ayırıcılık indeksi ortalaması |
|---------------|--------------|--------------------|-------|--------------------------|-------------------------------|
| Test 1 | 20 | 366 | 0,653 | 0,395 | 0,369 |
| Test 2 | 20 | 355 | 0,597 | 0,496 | 0,304 |

Test 1 ve Test 2 den madde seçimleri yapıldıktan sonra 20 maddelik asıl test oluşturulmuştur. Oluşturulan yeni testin analizleri yapılmış ve aşağıdaki tabloda verilmiştir.

Tablo 8: Nihai Test Genel Test İstatistikleri

| | Madde sayısı | Cevaplayıcı sayısı | KR-20 | Madde Güçlüğü Ortalaması | Ayırıcılık indeksi ortalaması |
|-------------------|--------------|--------------------|-------|--------------------------|-------------------------------|
| Nihai Test | 20 | 355 | 0,617 | 0,411 | 0,286 |

Nihai testin güvenilirlik hesaplaması 1036 cevaplayıcı ile yapılan uygulamada tekrar hesaplanmıştır.

4.3. Rasch Analizi Sonuçları

İkinci adımda ise Sayısal Yetenek ölçeğinin işleyişi aşağıdaki kriterlere göre değerlendirildi:

- Data-Model uyumu
- Kişi-model ve Test-Model uyumu testleri
- DMF araştırması
- Tek boyutluluk araştırması
- Yerel bağımsızlık varsayımının test edilmesi

İç ölçek geçerliliği ve kişi-yanıt geçerliliği araştırması için, WINSTEPS programı kullanılarak ortalama kare (MnSq) artıkları ve standart z-değerleri uygunluk istatistikleri, hem madde hem de kişiler için kullanılmıştır. Bu değerler Sayısal Yetenek Testi' ndeki gerçek yanıtlar ile Rasch modelinden beklenen yanıtlar arasındaki eşleşme derecesini gösterir. Uyum iyiliği istatistikleri infit ve outfit istatistikleri kullanılarak değerlendirildi. Mnsq uygunluk istatistiğinin beklenen değeri 1.0 dır ve maddenin uygunluk durumu için tercih edilir.

Sayısal Yetenek Testi' nin tek boyutluluğunu değerlendirmek için, artıkların temel bileşen analizi yapıldı. Burada kriter olarak model ile açıklanan varyansın birinci kontrastın özdeğerine oranına bakıldı. Bu oranın 3'ten büyük olması durumunda test tek boyutlu olarak kabul edildi (Conaghan, 2007).

Yerel bağımsızlık varsayımını değerlendirebilmek için aşağıdaki en yüksek maddeler arası standart artık korelasyonlarına bakıldı. . Bu değerler $|\cdot| \geq .30$ değerinden büyük olduğunda bu maddelerin yerel bağımsızlık varsayımını ihlal ettiği söylenir. Maddeler arası artık korelasyonları $|\cdot| \geq .30$ değerinin altında olduğunda ise bu maddeler için yerel bağımsızlık varsayımı sağlanmıştır denir (Bruin, 2012).

Son olarak SYT'nin madde kalibrasyonlarının cinsiyetler arasındaki stabilitesini değerlendirmek için bir dizi tek tip DMF analizi yapıldı. DMF'nin büyüklüğü, log-odds tahmin edicileri kullanan dikotom ölçekler için Winsteps programındaki Mantel-Haenszel istatistiği kullanılarak değerlendirildi. Ayrıca, madde yanlılığı olasılığını değerlendirmek için sonuçlar $p < 0.05$ ile rapor edildi.

4.3.1. Data-Model uyumu

Aşağıdaki tabloda SYT'deki maddelerin uyum istatistiklerini ve beklenen-gözlenen korelasyon değerlerini gösterilmiştir.

Tablo 9: Maddelerin Uyum İstatistikleri

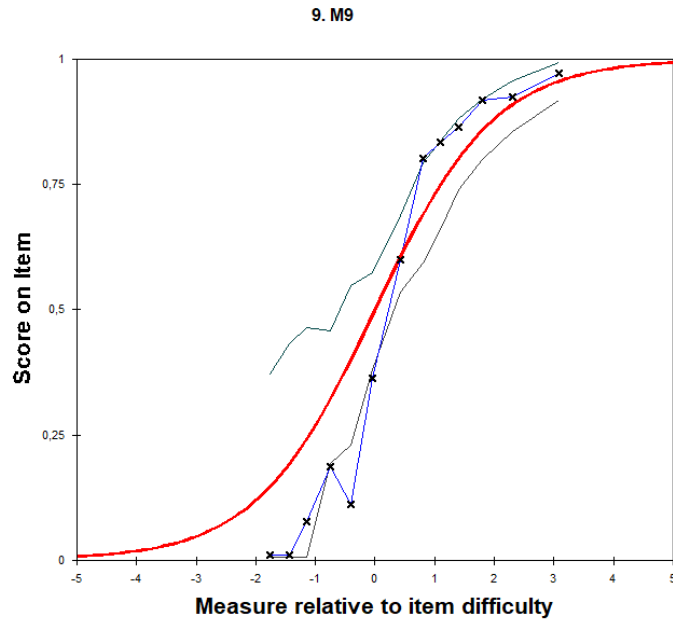
| INFIT | | OUTFIT | | Gözlenen eşleşme(%) | Beklenen eşleşme(%) | Gözlenen Korelasyon | Beklenen Korelasyon | Madde No |
|-------|------|--------|------|---------------------|---------------------|---------------------|---------------------|----------|
| MNSQ | ZSTD | MNSQ | ZSTD | | | | | |
| 1.37 | 8.1 | 1.65 | 7.1 | 70.3 | 78.5 | .11 | .41 | M17 |
| .97 | -.3 | 1.01 | .1 | 92.3 | 92.3 | .29 | .28 | M15 |
| .99 | -.1 | 1.05 | .4 | 90.5 | 90.6 | .29 | .30 | M4 |
| 1.01 | .2 | 1.05 | .4 | 87.0 | 86.8 | .33 | .34 | M14 |
| 1.07 | 1.4 | 1.05 | .5 | 81.9 | 82.7 | .33 | .38 | M12 |
| .99 | -.1 | .91 | -.8 | 84.6 | 84.3 | .38 | .37 | M5 |
| 1.10 | 3.3 | 1.16 | 3.6 | 67.1 | 70.4 | .39 | .47 | M13 |
| 1.05 | 1.3 | 1.04 | .6 | 76.0 | 77.4 | .39 | .42 | M11 |
| 1.07 | 2.5 | 1.13 | 3.1 | 70.2 | 70.4 | .40 | .47 | M1 |
| 1.05 | 1.6 | 1.09 | 1.9 | 70.4 | 70.9 | .42 | .46 | M3 |
| 1.01 | .3 | .99 | -.1 | 75.3 | 76.2 | .43 | .43 | M2 |
| 1.00 | -.1 | 1.00 | -.1 | 75.0 | 75.1 | .44 | .44 | M18 |
| .96 | -1.1 | .91 | -1.3 | 78.2 | 77.0 | .46 | .42 | M16 |
| .93 | -1.7 | .84 | -2.0 | 81.5 | 79.4 | .47 | .41 | M6 |
| 1.00 | -.1 | .98 | -.3 | 70.7 | 70.6 | .47 | .39 | M7 |
| .91 | -2.0 | .75 | -3.0 | 83.5 | 81.7 | .47 | .39 | M20 |
| .90 | -2.5 | .88 | -1.7 | 81.8 | 78.1 | .49 | .42 | M8 |
| .95 | -1.6 | .90 | -2.1 | 72.2 | 72.2 | .50 | .45 | M19 |
| .83 | -4.9 | .80 | -3.4 | 81.0 | 75.5 | .56 | .43 | M9 |

Eğer data Rasch modeline uyumlu ise beklenen ve gözlenen korelasyon değerleri eşit olur. Gözlenen korelasyon>Beklenen korelasyon ise madde yüksek ve düşük yetenekli öğrencileri yüksek ayırıcılık gücüne sahiptir. Tabloya göre 15.,5.,16.,6.,7.,20.8.,19.,9. Ve 10. Maddeler yüksek ayırıcılık gücüne sahiptir.

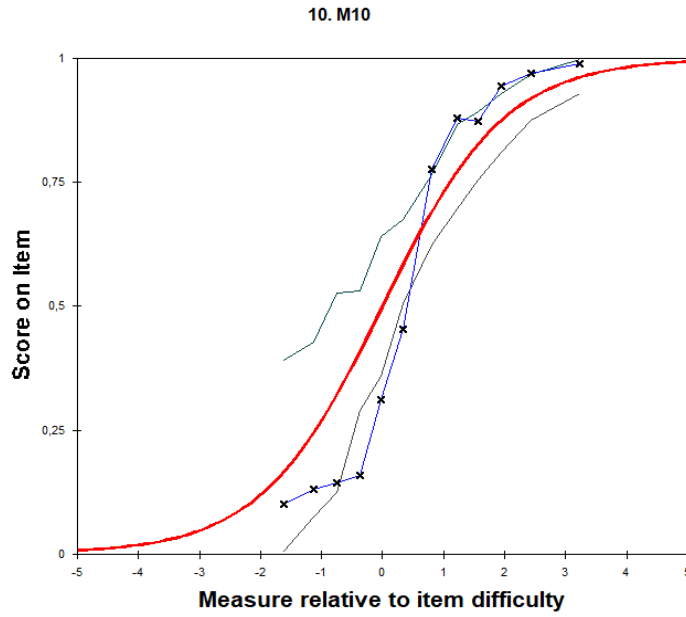
Gözlenen korelasyon<Beklenen korelasyon ise madde yüksek ve düşük yetenekli öğrencileri düşük ayırıcılık gücüne sahiptir. Tabloya baktığımızda 17.,4.,14.,12.,13.,11.,1.,2. Maddeler düşük ayırıcılık gücüne sahiptir.

INFIT ve OUTFIT meansquare istatistiklerine baktığımızda modele en uyumsuz madde, madde 17'dir (infite mnsq=1.37 ve outfit mnsq=1.65). Ölçekteki diğer tüm maddelerin infit ve outfit mnsq değerleri incelendiğinde hepsinin 1'e yakın olduğu görülmektedir. Yani maddeler model ile iyi uyum sağlamışlardır.

Şekil 7: Madde 9'un Beklenen Madde Karakteristik Eğrisi

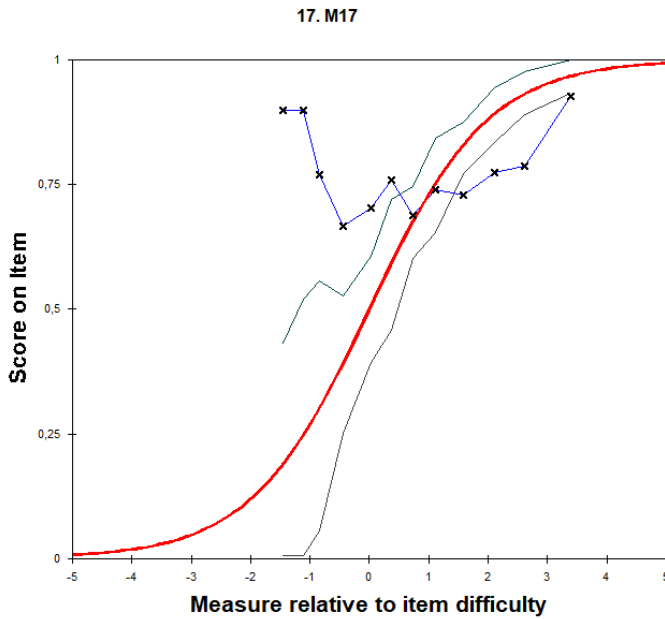


Şekil 8: Madde 10'un Beklenen Madde Karakteristik Eğrisi



Madde 9 ve madde 10'un beklenen madde karakteristik eğrilerine baktığımızda, madde uyum istatistikleri (outfit mnsq değerleri) 1'e yakın olmasına rağmen, bu maddelerin düşük yetenekli öğrencilere beklenenden daha kolay geldiği ve yüksek yetenekli öğrenciler için de beklenenden daha zor geldiği görülmektedir.

Şekil 9: Madde 17'nin Beklenen Madde Karakteristik Eğrisi



Madde 17 nin beklenen madde karakteristik eğrisine baktığımızda ise çok açıkça görülmektedir ki bu madde ortalama yetenek düzeyinin altındaki öğrenciler için beklenenden daha kolay ve bu düzeyin üzerindeki yüksek yetenekli öğrenciler için beklenenden daha zor gelmiştir. Bu durum data model uyumunu olumsuz yönde etkilemektedir.

Data model uyumunu incelerken bakılacak diğer bir istatistik beklenen ve gözlenen eşleşme yüzdeleridir. Bu yüzdeler birbirine eşit olduğunda data modele uyuyor demektir. Tablo incelendiğinde 17. Madde dışında diğer maddelerin beklenen ve gözlenen eşleşme yüzdeleri ya birbirine çok yakın ya da eşittir.

4.3.2. Kişi-Model Uyumu ve Test-Model Uyumu Testleri

1382 11. Sınıf öğrencisine uygulanan 20 soruluk sayısal yetenek testi sonuçları analiz edilmiş ve datadan uyumsuzluk gösteren öğrenciler çıkarılmıştır. Bu elemeler sonucunda 1036 öğrencinin datası analiz edilmiş olup öğrenci ve test uyum sonuçları aşağıdaki tabloda verilmiştir.

Tablo10: Öğrenci-Model Uyum İstatistikleri

| | Güvenirlilik | | INFIT | | OUTFIT | | Ayrırcılık | |
|----------------|--------------|-------|-------|------|--------|------|------------|----------|
| | Gözlenen | Model | mnsq | Zstd | mnsq | zstd | Gözlenen | Beklenen |
| | beklentisi | | | | | | | |
| Ortalama | .66 | .68 | 1.00 | .1 | .99 | .1 | 1,41 | 1,47 |
| Standart Sapma | .94 | .95 | .17 | | .50 | | | |

Öğrenciler için olan INFIT Ve OUTFIT uyum istatistiklerine baktığımızda sırasıyla 1,00 ve .99 olduğu görülmektedir. Bu istatistiklerin 1'e yakın olması mükemmel uyumu göstermektedir. Standart sapma model beklentisi .95 olarak raporlanmaktadır. Standart sapma infit mnsq ve outfit mnsq değerlerinin sırasıyla .17 ve .50 'dir.

Tablo 11: *Test-Model Uyum İstatistikleri*

| | GÜVENİRLİK | | INFIT | | OUTFIT | | AYIRICILIK | |
|----------|-------------------|--------------|--------------|-------------|---------------|-------------|-------------------|-----------------|
| | Gözlenen | Model | mnsq | Zstd | mnsq | zstd | Gözlenen | Beklenen |
| ORTALAMA | .99 | .99 | 1.00 | -.1 | .99 | -.1 | 9,98 | 10,14 |
| STANDART | .85 | .86 | .11 | | .20 | | | |
| SAPMA | | | | | | | | |

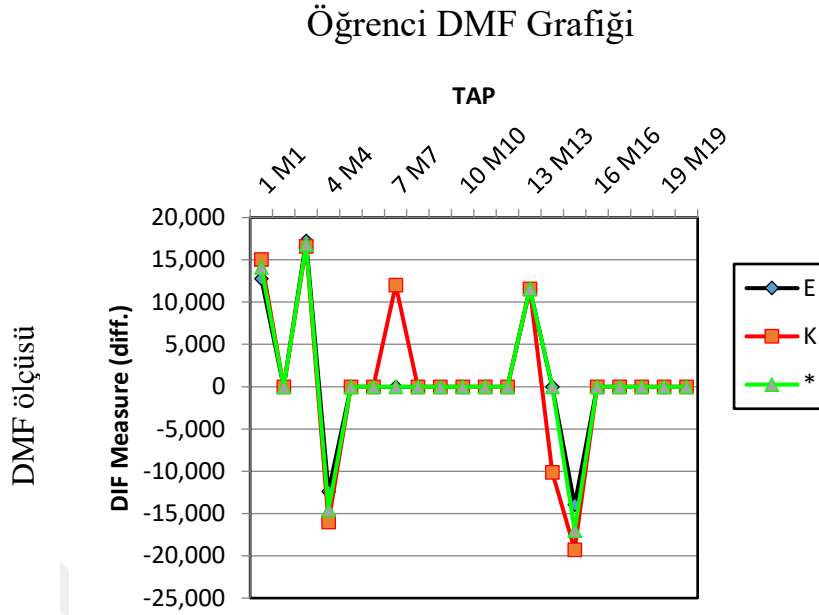
Testin geneli için olan INFIT ve OUTFIT istatistiklerine baktığımızda ise sırasıyla 1.00 ve .99 olduğu görülmektedir. Standart sapma model beklentisi .86 (1'e yakın) olarak raporlanmaktadır. Standart sapma infit mnsq ve outfit mnsq değerlerinin sırasıyla .11 ve .20 'dir.

4.3.3. DMF Araştırması

Bu çalışmada cinsiyete göre sayısal yetenekte farklılaşma olup olmadığı araştırıldığından, kızlar ve erkekler arasındaki farkın maddeler boyunca aynı olup olmadığı veya cinsiyetin maddeler boyunca farklılaşıp farklılaşmadığına bakılmalıdır. Bir araçtan yapılan bir ölçme tüm katılımcıların alt grupları için geçerli olmalıdır. Farklılıklar gözlenirse, o zaman maddenin Değişen Madde Fonksiyonu (DMF) gösterdiği söylenir ve madde, alt gruplar için aynı madde olarak değerlendirilmemelidir. Araştırmadaki veri setinde katılımcılar 2018-2019 öğretim yılında 11. Sınıfta okuyan kızlar ve erkeklerden oluşmaktadır. Testteki maddeler, bu iki cinsiyete göre değişen madde fonksiyonu bakımından kontrol edilmelidir.

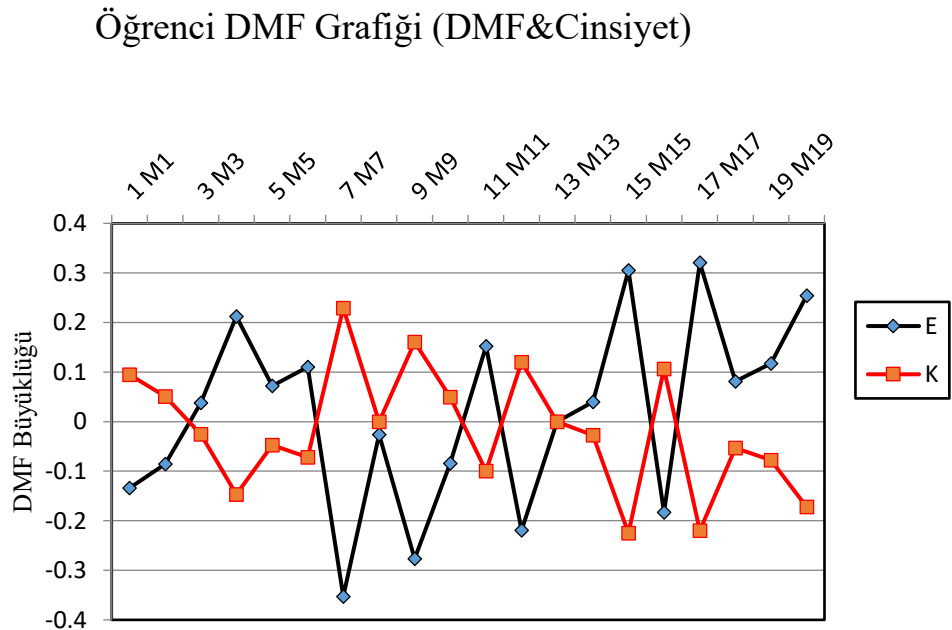
DMF'yi kontrol etmenin iyi bir görsel yolu, iki cinsiyet için her bir yetenek düzeyinde bir yanıt kategorisinin gözlemlenen ortalama puanını ayrı ayrı çizmek ve ortalama puanları karşılaştırmaktır.

Şekil 10:DMF Ölçüsü Grafiği



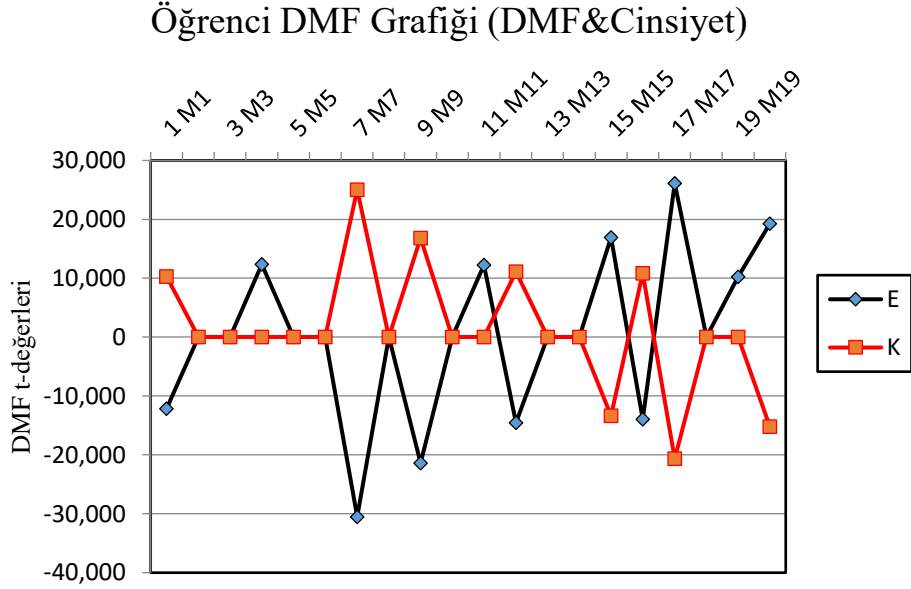
Bu grafik her maddenin iki öğrenci grubu için olan logit güçlüklerini gösteriyor. Madde7 ve 1 kızlar için erkeklere göre daha zor görünüyor. Madde 4,14 ve 15 kızlara daha kolay gelmiştir. Madde 15 kızlar için en kolay maddedir ya da başka bir deyişle kızlar bu maddeyi yanıtlamada erkeklere göre daha üstün yeteneklidir. Diğer maddeler her iki grup için de aynı güçlükte görünmektedir.

Şekil 11:DMF Büyüklüğü Grafiği



DMF büyüklüğünü gösteren grafiğe baktığımızda burada her maddenin güçlüğü ile ilişkili iki grup için madde DMF'nin logits büyüklükleri görünmektedir.

Şekil 12: DMF t-değerleri Grafiği



DMF t-değerleri grafiği de DMF büyüklüğünü gösteren grafikteki her bir DMF büyüklüğü le ilişkili olan istatistiksel anlamlılıkları(t değerlerini) göstermektedir. Madde 2,3,4,5,6,8,10,11,13,14,18,19 hariç diğer maddelerin DMF büyüklükleri farkı istatistiksel olarak anlamlı görünmektedir. 1,7,9,12,16. Maddelerin güçlükleri kızlar için erkeklere göre istatistiksel olarak anlamlı derecede daha zordur (t-değerleri farkı 2'den büyüktür).15,17 ve 20. maddelerin güçlükleri ise istatistiksel olarak anlamlı derecede erkeklere göre kızlardan daha zor gelmiştir. Kesin anlamlılık testleri için aşağıdaki tabloya bakalım.

Tablo 12: DMF Anlamlılık Testi Sonuçları

| Madde No | Cinsiyet | Öğrenci Sayısı | DMF Puanı | DMF ölçüsü | DMF Büyüklüğü | DMF t değeri | Olasılık(p) |
|----------|----------|----------------|-------------|------------|---------------|--------------|-------------|
| M1 | E | 430 | .03 | 1.28 | -.13 | -1.21 | .2252 |
| M1 | K | 606 | -.02 | 1.51 | .09 | 1.03 | .3045 |
| M2 | E | 430 | .01 | .01 | -.09 | -.68 | .5000 |
| M2 | K | 606 | -.01 | .15 | .05 | .52 | .6034 |
| M3 | E | 430 | -.01 | 1.72 | .04 | .35 | .7300 |
| M3 | K | 606 | .00 | 1.66 | -.03 | -.28 | .7831 |
| M4 | E | 430 | -.02 | -1.24 | .21 | 1.23 | .2175 |
| M4 | K | 606 | .01 | -1.60 | -.15 | -.98 | .3275 |
| M5 | E | 430 | -.01 | -.65 | .07 | .49 | .6209 |

| | | | | | | | |
|-----|---|-----|-------------|-------|------|-------|-------|
| M5 | K | 606 | .01 | -.76 | -.05 | -.41 | .6855 |
| M6 | E | 430 | -.01 | -.12 | .11 | .85 | .3949 |
| M6 | K | 606 | .01 | -.30 | -.07 | -.68 | .4944 |
| M7 | E | 430 | .06 | .62 | -.35 | -3.05 | .0024 |
| M7 | K | 606 | -.05 | 1.20 | .23 | 2.50 | .0125 |
| M8 | E | 430 | .00 | -.12 | -.03 | -.20 | .8404 |
| M8 | K | 606 | .00 | -.09 | .00 | .00 | 1.000 |
| M9 | E | 430 | .04 | -.10 | -.28 | -2.4 | .0330 |
| M9 | K | 606 | -.03 | .33 | .16 | 1.68 | .0930 |
| M10 | E | 430 | .01 | -.05 | -.08 | -.66 | .5124 |
| M10 | K | 606 | -.01 | .08 | .05 | .50 | .6155 |
| M11 | E | 430 | -.02 | .12 | .15 | 1.23 | .2209 |
| M11 | K | 606 | .02 | -.13 | -.10 | -.98 | .3277 |
| M12 | E | 430 | .02 | -.78 | -.22 | -1.45 | .1470 |
| M12 | K | 606 | -.02 | -.44 | .12 | 1.11 | .2675 |
| M13 | E | 430 | .00 | 1.16 | .00 | .00 | 1.000 |
| M13 | K | 606 | .00 | 1.16 | .00 | .00 | 1.000 |
| M14 | E | 430 | .00 | -.94 | .04 | .25 | .7998 |
| M14 | K | 606 | .00 | -1.01 | -.03 | -.22 | .8266 |
| M15 | E | 430 | -.02 | -1.40 | .31 | 1.69 | .0914 |
| M15 | K | 606 | .01 | -1.93 | -.23 | -1.34 | .1817 |
| M16 | E | 430 | .03 | -.17 | -.18 | -1.40 | .1632 |
| M16 | K | 606 | -.02 | .12 | .11 | 1.09 | .2777 |
| M17 | E | 430 | -.05 | .18 | .32 | 2.61 | .0094 |
| M17 | K | 606 | .03 | -.36 | -.22 | -2.07 | .0393 |
| M18 | E | 430 | -.01 | .30 | .08 | .67 | .5023 |
| M18 | K | 606 | .01 | .17 | -.05 | -.54 | .5869 |
| M19 | E | 430 | -.02 | .73 | .12 | 1.02 | .3062 |
| M19 | K | 606 | .01 | .53 | -.08 | -.83 | .4070 |
| M20 | E | 430 | -.03 | -.21 | .25 | 1.93 | .0548 |
| M20 | K | 606 | .02 | -.63 | -.17 | -1.52 | .1286 |

Yukarıdaki tabloda exel grafiğinde rapor edilen sayıları görüyoruz. Eğer bir maddenin bir grup lehine ya da aleyhine yanlı olup olmadığını anlamak istiyorsak DMF puanı değerine bakarız. Eğer DMF puanı değeri pozitif ise madde yanlılığı o grubun lehinedir, eğer negatifse madde yanlılığı o grubun aleyhinedir ($p < .05$).

Yukarıdaki tabloda DMF puanı değerlerine baktığımızda erkekler için pozitif değere sahip olan maddeler 1,2,7,9,10,12 ve 16 ncı maddelerdir. Bu maddeler erkekler lehine yanlı görünmektedir.

Erkekler için negatif DMF puanı değerine sahip maddeler ise 3,4,5,6,11,15,17,18,19,20. maddelerdir ve bu maddeler erkekler aleyhine yanlı görünmektedir. Daha ayrıntılı bilgi için aşağıdaki olasılık (p) tablosuna bakılır.

Aşağıdaki tablo “Bu madde istatistiksel olarak DMF içermez” şeklindeki H_0 hipotezini test eden sonuçları göstermektedir.

Tablo 13:DMF İçeren Maddeler

| Öğrenci sınıflaması (K/E) | DIF Ki-Kare Değeri | Serbestlik Derecesi | Olasılık | Madde Numarası |
|---------------------------|--------------------|---------------------|----------|----------------|
| 2 | 2.5257 | 1 | .1120 | 1 |
| 2 | .7249 | 1 | .3945 | 2 |
| 2 | .1949 | 1 | .6589 | 3 |
| 2 | 2.4795 | 1 | .1153 | 4 |
| 2 | .4085 | 1 | .5227 | 5 |
| 2 | 1.1908 | 1 | .2752 | 6 |
| 2 | 15.4304 | 1 | .0001 | 7 |
| 2 | .0406 | 1 | .8404 | 8 |
| 2 | 7.3660 | 1 | .0066 | 9 |
| 2 | .6814 | 1 | .4091 | 10 |
| 2 | 2.4568 | 1 | .1170 | 11 |
| 2 | 3.3330 | 1 | .0679 | 12 |
| 2 | .0000 | 1 | 1.000 | 13 |
| 2 | .1123 | 1 | .7375 | 14 |
| 2 | 4.6334 | 1 | .0314 | 15 |
| 2 | 3.1230 | 1 | .0772 | 16 |
| 2 | 11.0003 | 1 | .0009 | 17 |
| 2 | .7452 | 1 | .3880 | 18 |
| 2 | 1.7345 | 1 | 1.7390 | 19 |
| 2 | 5.9971 | 1 | .0143 | 20 |

H_0 hipotezi madde 7,9,15,17 ve 20 için reddediyoruz ($p<.05$). 7. ve 9. maddeler erkekler lehine yanlılık gösterirken, 15,17 ve 20. maddeler kızlar lehine yanlılık göstermektedir. Fakat bu maddeler dışındaki diğer maddeler için bu H_0 hipotezini kabul ediyoruz ($p>.05$).

4.3.4.Tek Boyutluluk İncelemesi

Bu adımda, ölçeğin tek bir temel yapıyı ölçüp ölçmediğini belirlemek için ölçeğin tek boyutluluğu incelendi. Rasch programları genellikle artıkların temel bileşen analizini sağlar. Bu, maddelerin yerel bağımsızlığının test edilmesini sağlar. Bu test, Rasch faktörü göz önüne alındığında, rastgele ilişkilerden başka, maddeler arasında başka ilişki olmaması gerektiğini ima eder. Bunun için bir test Smith

tarafından önerilmiştir. Bu test, artıkların içindeki madde örüntüsünü alır, maddeler ve birinci artık faktör arasındaki korelasyonu inceler ve bu kalıpları 2 madde alt kümesini (yani, pozitif ve negatif olarak ilişkili maddeler) tanımlamak için kullanır. Bu 2 madde grubu daha sonra ayrı kişi tahminleri yapmak için kullanılır ve her bir kişi için bu tahminlerdeki fark için bağımsız bir t testi kullanılarak, bu tür testlerin -1.96 ile 1.96 arasındaki oranının% 5'i geçmemesi gerekir. Binominal testin, gözlenen anlamlı test sayısı için bir güven aralığı hesaplanır ve bu değer ölçeğin tek boyutlu olması için beklenen% 5'lik değeri ile örtüşmelidir (Conaghan, 2007).

Sayısal yetenek testinin tek boyutluluk özelliği incelendiğinde model ile açıklanan ham varyans özdeğerinin, 1. Kontrast içindeki açıklanamayan ham varyans özdeğerine oranı $7.0636/2.0705=3.4115$ olarak bulunur. Bu oran 3'ten büyük olduğundan testin tek bir örtük özelliği yani sayısal yeteneği ölçtüğünü göstermektedir. Sayısal Yetenek Testinin tek boyutlu olduğu sonucuna varılmıştır.

Tablo 14: *Özdeğer Birimi İçindeki Standart Artık Varyans Tablosu*

| | Özdeğer | Gözlenen | Beklenen |
|---|----------------|-----------------|-----------------|
| Gözlenen toplam ham Varyans | 27.0636 | % 100 | % 100 |
| Model ile açıklanan ham varyans | 7.0636 | % 26.1 | % 26.1 |
| Kişi ile açıklanan ham varyans | 3.6635 | % 13.5 | % 13.5 |
| Maddeler ile açıklanan ham Varyans | 3.4001 | % 12.6 | % 12.6 |
| Açıklanamayan ham varyans(toplam) | 20.00 | % 73.9 | % 73.9 |
| 1.Kontrast içindeki açıklanamayan varyans | 2.0705 | % 7.7 | |

| | | |
|---|--------|------|
| 2. Kontrast içindeki açıklanamayan varyans | 1.6570 | %6.1 |
| 3.kontrast içindeki açıklanamayan varyans | 1.2964 | %4.8 |
| 4. kontrast içindeki açıklanamayan varyans | 1.1576 | %4.3 |
| 5. Kontrast içindeki açıklanamayan varyans | 1.1375 | %4.2 |

4.3.5. Yerel Bağımsızlık Varsayımının Değerlendirilmesi

Yerel bağımsızlık varsayımını değerlendirebilmek için aşağıdaki en yüksek maddeler arası standart artık korelasyonlarını gösteren tabloya bakılır. Maddeler arası korelasyon $|\cdot| \geq .30$ 'den küçükse yerel bağımsızlık varsayımının sağlandığı söylenir (Bruin, 2012).

Tablo 15: *En Yüksek Standart Artık Korelasyonları*

| Korelasyon | Madde No | Madde No |
|------------|----------|----------|
| .55 | 9 | 10 |
| .25 | 8 | 10 |
| .23 | 8 | 9 |
| .17 | 19 | 20 |
| .13 | 16 | 20 |
| -.16 | 9 | 19 |
| -.16 | 7 | 16 |
| -.15 | 7 | 19 |
| -.15 | 8 | 19 |
| -.15 | 10 | 17 |
| -.15 | 9 | 13 |
| -.15 | 3 | 18 |
| -.15 | 10 | 13 |
| -.14 | 7 | 20 |

| | | |
|------|----|----|
| -.14 | 2 | 18 |
| -.14 | 3 | 20 |
| -.13 | 9 | 11 |
| -.13 | 10 | 11 |
| -.13 | 9 | 16 |
| -.13 | 10 | 19 |

Tabloya bakıldığında sadece 9-10 maddeleri arasındaki artık korelasyonun 0.55 olduğu görülmektedir. Bu değer $|.30|$ değerinden büyük olduğundan bu maddeler yerel bağımsızlık varsayımını ihlal ederler. Diğer maddeler arası artık korelasyonları $|.30|$ değerinin altında olduğundan bu maddeler için yerel bağımsızlık varsayımı sağlanır. $|.30|$ değerinden büyük olan korelasyon sadece 9. ve 10. maddeler arasında olduğundan testin yerel bağımsızlık varsayımını karşılanır (Bruin, 2012).

BÖLÜM V

SONUÇ, TARTIŞMA VE ÖNERİLER

Bu bölümde, elde edilen bulgular ışığında sonuçlar genel olarak özetlenmiş ve bu sonuçlar doğrultusunda tartışma yapılarak araştırmacılara öneriler sunulmuştur.

5.1 Sonuç

Bu çalışmada, Sayısal Yetenek farklılaşması cinsiyete göre incelenmiştir. Bunun için önce her bir alt testi 10'ar maddeden oluşan eşitlik kurma, sayısal akıl yürütme, geometri ve görsel uzamsal yetenek maddelerinden oluşan toplam 40 maddelik bir test oluşturulmuştur. Bu test yaklaşık olarak 400 kişilik 2017-2018 eğitim öğretim yılında 11. sınıfta okuyan öğrencilerden oluşan örnekleme uygulanmıştır. Bu uygulama sonrasında madde güçlükleri ve ayırdedicilikleri hesaplanmıştır. Aritmetik Akıl Yürütme, Eşitlik Kurma, Geometri ve Görsel-Uzamsal Yetenek yapılarını en iyi ölçen her alt testteki 5'er madde seçilip, teste son şekli verilmiştir. Teste son şekli verildikten sonra 20 maddeden oluşan sayısal yetenek testi 2018-2019 eğitim öğretim yılında Bornova Anadolu Lisesi, Cihat Kora Anadolu Lisesi, 15 Temmuz Şehitler Anadolu Lisesi, Atatürk Lisesi, Karşıyaka Lisesi, Atakent Anadolu Lisesi okullarının 11. Sınıfında okuyan 1382 öğrenciye uygulanmıştır. İlk analizler yapıldıktan sonra modele uyumayan öğrenciler datadan temizlenmiştir.

Yapılan Rasch analizi sonucunda data-model uyumu, kişi-model uyumu, test model uyumu, DMF araştırması, yerel bağımsızlık ve tek boyutluluk varsayımlarının denetlenmiştir.

DMF araştırması analizleri sonucunda, Sayısal yetenek farklılaşması cinsiyete göre incelenmiş olup, eşitlik kurma-sayısal akıl yürütme alanlarındaki 2 maddenin erkekler lehine DMF içerdiği ve geometri-görsel uzamsal alanlarındaki 3 maddenin ise kızlar lehine DMF içerdiği tespit edilmiştir.

SYT'nin Rasch modele uyum araştırması sonucunda datanın modele uyduğu ve model varsayımlarının karşılandığı tespit edilmiştir.

5.2. SYT'nin Psikometrik Özelliklerinin Tartışılması

Bu tez çalışmasında, Sayısal Yetenekteki farklılaşma cinsiyete göre incelenmiştir. Bunun için önce her bir alt testi 10'ar maddeden oluşan toplam 40 maddelik bir test oluşturuldu. Bu test yaklaşık olarak 400 kişilik 2017-2018 eğitim öğretim yılında 11. sınıfta okuyan öğrencilerden oluşan örnekleme uygulandı. Bu uygulama sonrasında madde güçlükleri ve ayırdedicilikleri hesaplanarak madde seçimi yapıldı. Buna göre Aritmetik akıl yürütme ve eşitlik kurma alt testlerinden oluşan Test1'deki 20 adet maddeden; madde güçlükleri 0.35 ile 0.77 arasında, madde ayırdedicilikleri 0.36 ile 0.65 arasında ve madde standart sapmaları 0.42 ile 0.50 arasında değişen 10 madde seçilmiştir. Geometri ve Görsel-uzamsal yetenek sorularından oluşan Test2'deki 20 maddeden ise madde güçlükleri 0.27 ile 0.84 arasında, ayırdedicilikleri 0.31 ile 0.51 arasında ve madde standart sapmaları 0.36 ile 0.50 arasında değişen 10 madde seçilmiştir. Güvenirlik ve geçerlilik çalışmaları yapıldığında, Test1'in Kr-20 güvenirlilik katsayısı 0.653 ve Test2'nin güvenirlilik katsayısı 0.597 olarak bulunmuştur. Sonuç olarak ölçmek istediğimiz yapıları (Aritmetik Akıl Yürütme, Eşitlik Kurma, Geometri ve Görsel-Uzamsal Yetenek) en iyi ölçen her alt testteki 5'er madde seçilip, teste son şekli verilmiştir. Oluşturulan bu nihai testin 355 öğrenci üzerinden yapılan Kr-20 güvenirlilik hesaplamalarında güvenirlilik katsayısı 0.61, madde güçlük ortalaması 0.411 ve madde ayırdedicilik ortalaması 0.286 olarak bulunmuştur. Güvenirlilik katsayısı düşük görünse de daha sonra yapılan Rasch analizinde 1036 öğrenci üzerinden yapılan test güvenirlilik hesaplamalarında test güvenirliliği 0.99 olarak bulunmuştur. Örneklem büyüklüğü arttığında testin genel güvenirliliğinin arttığı görülmüştür.

Teste son şekli verildikten sonra 20 maddeden oluşan sayısal yetenek testi 2018-2019 eğitim öğretim yılında Bornova Anadolu Lisesi, Cihat Kora Anadolu Lisesi, 15 Temmuz Şehitler Anadolu Lisesi, Atatürk Lisesi, Karşıyaka Lisesi, Atakent Anadolu Lisesi okullarının 11. Sınıfında okuyan 1382 öğrenciye uygulandı. İlk analizler yapıldıktan sonra modele uyumayan öğrenciler datadan temizlendi. Bu öğrencilerin datadan silinmesi işlemi kişi model uyumu tablosunda outfit mnsq istatistiği değerleri 1.5 ten büyük olanların elenmesi ve tüm maddelere doğru cevap veren ya da hepsine yanlış cevap verenlerin elenmesi şeklinde gerçekleştirildi. Eleme

sonrası 1036 öğrenci üzerinde analizler gerçekleştirilerek rapor edildi. Rasch analizlerindeki kişi-model uyumuna bakıldığında outfit mnsq değerinin 0.99 ve infit mnsq değerinin de 1.00 olduğu görülmüş olup, bu da kişi-model uyumunun mükemmel olduğunu göstermektedir.

5.3. SYT'nin Rasch Modelinin Varsayımlarını Karşılıyıp Karşılımadığının Tartışılması

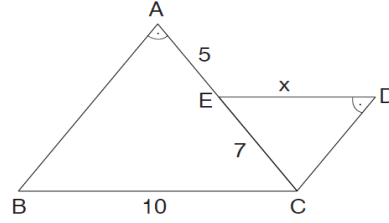
Yapılan Rasch analizi sonucunda data-model uyumu, kişi-model uyumu, test model uyumu, DMF araştırması, yerel bağımsızlık ve tek boyutluluk varsayımlarının denetlenmiştir. Özalp Ateş (2015)'in Ölçeklerde yapı geçerliliğinin incelenmesi çalışmasında, ölçeğin sadece boyut yapısını inceleyen faktör analizi yöntemlerine kıyasla, içsel yapı geçerliliğinin madde uyumu, birey uyumu, madde yanlılığı gibi farklı özellikler bakımından da değerlendirildiği Rasch Analizi'nin beraber kullanılmasını önermiştir (Özalp Ateş, 2015). Bu çalışmadaki bulgular sadece Rasch analizlerine dayalı olup faktör analizi gibi diğer yöntemler test geliştirme aşamasında kullanılmamıştır.

Bulgulara göre 15.,5.,16.,6.,7.,20.,8.,19.,9. ve 10. maddeler yüksek ayırıcılık gücüne sahiptir. 17.,4.,14.,12.,13.,11.,1.,2. maddeler ise düşük ayırıcılık gücüne sahiptir. INFIT ve OUTFIT meansquare istatistiklerine baktığımızda modele en uyumsuz maddenin madde 17 olduğunu görülmüştür (infit mnsq=1.37 ve outfit mnsq=1.65). Bu maddenin ölçekten çıkarılması önerilir. Ölçekteki diğer tüm maddelerin infit ve outfit mnsq değerleri incelendiğinde hepsinin 1'e yakın olduğu görülmüştür. Yani maddeler model ile iyi uyum sağlamışlardır. Ayrıca data model uyumunu incelemek için beklenen ve gözlenen eşleşme yüzdelerine bakılmıştır. Bu yüzdeler birbirine eşit olduğunda datanın modele uyumlu olduğu söylenir.17. Madde dışında diğer maddelerin beklenen ve gözlenen eşleşme yüzdeleri ya birbirine çok yakın ya da eşittir. Madde 9 ve madde 10'un beklenen madde karakteristik eğrilerine baktığımızda, madde uyum istatistikleri(outfit mnsq değerleri) 1'e yakın olmasına rağmen, bu maddelerin düşük yetenekli öğrencilere beklenenden daha kolay geldiği ve yüksek yetenekli öğrenciler için de beklenenden daha zor geldiği görülmektedir. Yani sadece beklenen ve gözlenen eşleşme yüzdelerine bakılması, data-model uyumu

ile ilgili yorum yapmada yetersizdir. Daha sağlıklı kararlar için madde karakteristik eğrileri de incelenmelidir.

17. Madde bir geometri-benzerlik sorusu olup aşağıdaki gibidir. Bu maddenin testten çıkarılmasının uyum istatistiklerini olumlu yönde etkileyeceği düşünülmektedir.

17.



$BC \parallel ED$

$m(\widehat{BAC}) = m(\widehat{EDC})$

$|AE| = 5\text{cm}, |EC| = 7\text{cm}, |BC| = 10\text{cm}$

İse şekilde verilenlere göre $|ED| = x$ kaç cm' dir?

A) 6 B) 8 C) $\frac{42}{5}$ D) $\frac{50}{7}$ E) 9

Öğrenciler için olan INFIT Ve OUTFIT uyum istatistiklerine baktığımızda sırasıyla 1,00 ve .99 olduğu görülmektedir. Bu istatistiklerin 1'e yakın olması mükemmel uyumu göstermektedir. Standart sapma model beklentisi .95 olarak raporlanmıştır. Standart sapma infit mnsq ve outfit mnsq değerlerinin sırasıyla .17 ve .50 olması bir miktar uyumsuzluk olduğunu düşündürmektedir.

Testin geneli için olan INFIT ve OUTFIT istatistiklerine baktığımızda ise sırasıyla 1.00 ve .99 olduğu görülmüş ve standart sapma model beklentisi .86(1'e yakın) olarak raporlanmıştır. Standart sapma infit mnsq ve outfit mnsq değerlerinin sırasıyla .11 ve .20 olması test-data uyumun iyi olduğunu düşündürmektedir.

Sayısal yetenek testinin tek boyutluluk özelliği incelendiğinde model ile açıklanan ham varyans özdeğerinin, 1. Kontrast içindeki açıklanamayan ham varyans özdeğerine oranı $7.0636/2.0705=3.4115$ olarak bulunmuştur. Bu oran 3'ten büyük olduğundan testin tek bir örtük özelliği yani sayısal yeteneği ölçtüğünü göstermektedir. Sayısal Yetenek Testinin tek boyutlu olduğu sonucuna varılmıştır.

Yerel bağımsızlık varsayımını değerlendirmek için maddeler arasındaki artıkların korelasyonuna bakılmış ve sadece 9-10 maddeleri arasındaki artık korelasyonun 0.55 olduğu görülmüştür. Bu değerler $|.30|$ değerinden büyük

olduğundan bu maddeler yerel bağımsızlık varsayımını ihlal ederler. Diğer maddeler arası artık korelasyonları $|\cdot 30|$ değerinin altında olduğundan bu maddeler için yerel bağımsızlık varsayımı sağlanmıştır. $|\cdot 30|$ değerinden büyük olan korelasyon sadece 9. ve 10. maddeler arasında olduğundan testin yerel bağımsızlık varsayımını karşıladığı söylenmiştir. Yerel bağımsızlık varsayımını ihlal eden madde çiftlerini(9-10) testten elyerek ya da bu maddeleri birleştirip tek bir madde olarak teste koyarak bu ihlalin önüne geçilebilir.

5.4. SYT Maddelerinin Cinsiyet Gruplarına Göre Değişen Madde Fonksiyonu(DMF) İçerip İçermediğinin Tartışılması

DMF araştırma sonuçları için öğrenci DMF grafiklerine bakıldığında, bu grafiklerde her maddenin iki öğrenci grubu için olan logit güçlükleri, DMF büyüklükleri ve DMF t-değerleri gösterilmiştir. Madde 7 ve madde 1 kızlar için erkeklere göre daha zor görülmektedir. Madde 4,14,15 ve 16 da kızlara daha kolay gelmiştir. Madde 15 kızlar için en kolay maddedir ya da başka bir deyişle kızlar bu maddeyi yanıtlamada erkeklere göre daha üstün yeteneklilerdir. Diğer maddeler her iki grup için de aynı güçlükte görünmektedir. DMF t-değerleri için, madde 2,3,5,6,9,12,13,17 hariç diğer maddelerin DMF büyüklükleri farkı istatistiksel olarak anlamlı görünmektedir. 1,7,9,12,16. maddeler kızlar için erkeklere göre istatistiksel olarak anlamlı derecede daha zordur(t-değerleri farkı 2'den büyüktür).15,17 ve 20. Maddelerin güçlükleri ise istatistiksel olarak anlamlı derecede erkeklere göre kızlardan daha zor gelmiştir.

Kesin anlamlılık testleri için DMF anlamlılık testi sonuçlarına bakılmıştır. Buna göre DMF puanı değerlerine bakıldığında erkekler için pozitif değere sahip olan maddeler 1,2,7,9,10,12 ve 16 ıncı maddelerdir. Bu maddeler erkekler lehine yanlı görünmektedir. Erkekler için negatif DMF puanı değerine sahip maddeler ise 3,4,5,6,11,15,17,18,19,20. maddelerdir ve bu maddeler erkekler aleyhine yanlı görünmektedir. Daha ayrıntılı bilgi için ise “Bu madde istatistiksel olarak DMF içermez” şeklindeki H_0 hipotezini test eden sonuçları gösteren olasılık(p) değerleri tablosu incelenmiş, 7,9,15,17 ve 20'in DMF içerdiği tespit edilmiştir($p < .05$). 7. ve 9. maddeler erkekler lehine yanlılık gösterirken, 15,17 ve 20. maddeler kızlar lehine

yanlılık gösterdiği tespit edilmiştir. Fakat bu maddeler dışındaki diğer maddeler için “Bu madde istatistiksel olarak DMF içermez” şeklindeki H_0 hipotezi kabul edilmiştir ($p > .05$). Erkekler lehine yanlılık gösteren maddeler aşağıdaki gibidir:

7.

| | Arda | Mert | Zeynep | Polen | Orhan |
|---------|------|------|--------|-------|-------|
| Ağırlık | 27 | 45 | 51 | 63 | 69 |

Yukarıdaki tabloda Arda, Mert, Zeynep, Polen ve Orhan'ın ağırlıkları kg cinsinden verilmiştir. Bu beş kişi bir miktar elmayı aşağıdaki kurallara göre paylaşıyorlar.

- İlk olarak elmaları bu beş kişi aralarında eşit paylaşıyorlar.
- Daha sonra bu kişiler ağırlıkça kendinden ağır olanlara her 1 kg farkı için 3 elma veriyor.
- Paylaşım sonucunda Arda'nın 20 elması oluyor.

Paylaşım sonucunda kimin elmalarının sayısı ilk paylaşımındaki elmaların sayısı ile aynıdır?

- A) Arda B) Mert C) Zeynep
D) Polen E) Orhan

Bir baba dikdörtgen şeklindeki tarlasını Ayşe, Fatma, Ali ve Veli isimli çocuklarına miras olarak 4 parçaya ayırıp aşağıdaki gibi paylaşmıştır. Çocuklar da tarlaya A, B, C ve D ürünlerini ekmişlerdir.

| | | |
|-------|-----|------|
| Ayşe | Ali | Veli |
| Fatma | | |

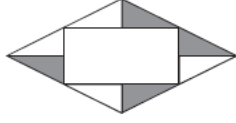
- Ayşe'nin tarlasının alanı, Fatma'nın tarlasının alanına eşittir.
- Veli'nin tarlasının alanı Ali'ninkinin 2, Ayşe'ninkinin 4 katı büyüklüktedir.
- Ayşe tarlasına A ürününü ektiğinde 600 kg, B ürününü ektiğinde 360 kg, C ürününü ektiğinde 180 kg ve D ürününü ektiğinde ise 100 kg ürün almaktadır.
- Tüm tarlaların verimi eşittir.
- Her tarlaya farklı ürün ekilmekte ve boş tarla kalmamaktadır.

9. Verilen bilgilere göre tarlalardan **en fazla** toplam kaç kg ürün alınabilir?

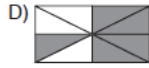
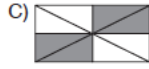
- A) 3400 B) 4600 C) 4800
D) 5000 E) 5100

Kızlar lehine yanlılık gösteren maddeler de aşağıdaki gibidir:

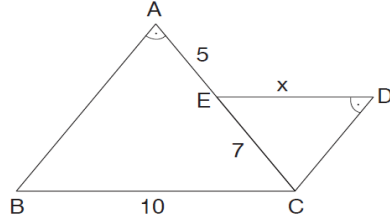
15.



Yukarıda açılmış biçimi verilen şeffaf zarf kapatıldığında aşağıdakilerden hangisi elde edilir?



17.



$BC \parallel ED$

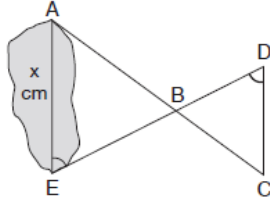
$$m(\widehat{BAC}) = m(\widehat{EDC})$$

$$|AE| = 5\text{cm}, |EC| = 7\text{cm}, |BC| = 10\text{cm}$$

İse şekilde verilenlere göre $|ED| = x$ kaç cm' dir?

- A) 6 B) 8 C) $\frac{42}{5}$ D) $\frac{50}{7}$ E) 9

20.



Bir kamp rehberi şeklinde görülen bataklığın üzerinden geçen A ve E noktaları arasındaki x cm uzunluğundaki bir köprünün uzunluğunu hesaplamak istiyor.

$|AB| = 1800$ cm, $|EB| = 1400$ cm,

$|BD| = 700$ cm ve $|CD| = 800$ cm dir.

AC ve DE doğru parçaları da B noktasında kesişmektedir.

$m(\widehat{AEB}) = m(\widehat{CDB})$ ise $|AE| = x$ kaç cm dir?

A) 1200

B) 1600

C) 1700

D) 1800

E) 2000

Bu maddeler incelendiğinde 7. ve 9. maddelerin aritmetik akıl yürütme/Eşitlik kurma alt testlerine ait olduğu ve bu maddeleri erkeklerin kızlara göre daha iyi yaptığı görülüyor. 15. Maddenin görsel uzamsal yetenek alt testine, 17. ve 20. maddelerin ise geometri alt testine ait olduğu görülüyor ve kızlar bu maddelere doğru yanıt vermede erkeklere göre daha başarılıdırlar.

5.5. Sayısal Yetenekte Cinsiyete Göre Farklılaşmanın Olup Olmadığının Tartışılması

Sayısal yetenek testlerinde cinsiyet temelli bir farklılaşmanın olup olmadığını araştırmak önemlidir. Meslek seçimi ve kurumlara yerleşme süreçlerinde kız ve erkek bireylerin girdikleri sınavlardaki testlerin kızların veya erkeklerin lehine ya da aleyhine olmaması gerekir. Hayatımız boyunca, özellikle de okul yıllarında herkes çok sayıda teste tabi tutulur. Bu testlerden bazıları yüksek öneme sahiptir, yani yaşamları etkileyen kararlar için temel oluştururlar. Yüksek öneme sahip sınavların çoğunda, cinsiyet, ırk veya etnik grup, sosyo-ekonomik durum ve coğrafi bölgenin bir fonksiyonu olarak farklılık gösteren grup farklılıkları veya ortalama farklılıklar vardır.

Çelik 2019'da yüksek lisans tezi kapsamında yaptığı çalışmada, Uluslararası Öğrenci Değerlendirme Programı (PISA) 2015 matematik alt testindeki maddelerin istatistiki bölge birimleri ve cinsiyete göre DMF içerip içermediğini incelemiştir. Araştırmasında olası DMF'nin tespit edilmesinde Madde Tepki Kuramı'na (MTK) göre Rasch modeli esas almıştır. Uluslararası Öğrenci Değerlendirme Programı 2015 matematik alt testindeki maddelerin eğitim bölgelerine dayalı analizlerden elde edilen sonuçları incelendiğinde, tüm maddelerde anlamlı düzeyde DMF bulgusuna rastlamıştır. Araştırma kapsamına alınan matematik alt testine ait 55 maddenin cinsiyete göre DMF incelemesinde beş maddenin DMF içerdiğini tespit etmiştir (Çelik, 2019). Bu çalışmada da DMF içeren 5 madde tespit edilmiş olup Mustafa Çelik'in PISA matematik alt testi için yaptığı DMF araştırmasıyla benzerlik göstermektedir.

Ayan (2011) yüksek lisans tez araştırmasında, Uluslararası Öğrenci Değerlendirme Programı (PISA) fen okuryazarlığı alt testinde yer alan maddelerle, cinsiyet değişkenine göre, değişen madde fonksiyonu (DMF) analizi yürütmüştür. Bu çalışmada DMF analizinde lojistik regresyon (LR) ve Mantel Haenszel (MH) yöntemleri kullanmış ve ardından bu iki teknikten elde edilen sonuçları karşılaştırmıştır. Cinsiyete göre MH yöntemi ile yapılan DMF analizi sonuçlarına göre incelenen 34 adet maddenin 4 tanesinin B düzeyinde DMF içermekte olduğunu, PISA 2009 Fen okuryazarlığı alt testinde 3 tanesi kızlar 1 tanesi ise erkekler lehine çalışmak üzere toplamda 4 adet maddenin orta düzeyde DMF göstermekte olduğu sonucuna ulaşmıştır. Cinsiyete dayalı LR yöntemi ile yapılan incelemede ise yine 4 adet TBO DMF gösterdiği tespit edilen madde belirlemiştir (Ayan, 2011). SYT'nin DMF araştırması sadece Rasch analizi ile yapıldığından diğer DMF belirleme tekniklerine göre aynı sonuçların elde edilip edilmeyeceği bilinmemektedir.

T.Öğretmen ve N.Doğan, 2004'te yayınladıkları bir çalışmada Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme sınavı Matematik alt testindeki maddelerin madde yanlılığı analizlerini madde tepki kuramı çerçevesinde incelemiştir. Araştırmadan elde edilen sonuçlar, OKÖSYS Matematik alt testinin, cinsiyet gruplarına göre testin her maddesinin yanlılık gösterdiğini ortaya koymuştur. Özellikle geometri soru grubunun hemen hepsinin yanlılık değerlerinin yüksek ve kızların lehine olduğunu bulmuşlardır. SYT'nin DMF içeren maddelerine baktığımızda da geometri

ve görsel-uzamsal yetenek testlerindeki üç maddenin kızlar lehine DMF içerdiği tespit edildi. Araştırmadaki sonuçlar bu yönüyle Öğretmen ve Doğan'ın araştırması ile paralellik göstermektedir (Öğretmen & Doğan, 2004).

Literatürdeki ilgili araştırmalara baktığımızda Julia Sherman, matematik arka planı kontrol edilmiş bilişsel becerileri ve matematiğe karşı tutumları benzer 135 kız ve 75 erkek üzerinde yaptığı çalışmada, erkeklerin kızlara göre matematikte 11. sınıfta anlamlı olarak daha iyi performans gösterdiğini tespit etmiştir. Özellikle Mekansal görselleştirme konusunda erkekler lehine cinsiyete bağlı bir fark olduğunu raporlamıştır. SYT'nin DMF analizi sonuçlarına baktığımızda ise geometri ve görsel uzamsal yetenek alanlarında Sherman'ın sonucundan farklı olarak kızlar lehine anlamlı fark çıkmıştır (Gallagher & Kaufman, 2000).

Hyde, Janet S., Fennema, Elizabeth, Lamon ve Susan J. matematik performansındaki cinsiyet farklılıklarına ilişkin 1963-1988 yılları arasında yayınlanmış 100 adet çalışmanın bir meta-analizini gerçekleştirmişlerdir. Araştırmacılar, sürekli olarak erkeklerin matematik testlerinde kızlardan daha iyi performans gösterdiğine karar vermiştir. Yaş eğilimlerinin incelenmesi sonucunda, kızların ilkökul ve ortaokulda hesaplamada testinde hafif bir üstünlük gösterdiğini, ilkökul ve ortaokulda ise problem çözme alanında cinsiyet üstünlüğü olmadığını; Lise ve kolejde ise erkekler lehine olan farklı sonuçlar ortaya çıkmıştır. Matematik performansındaki cinsiyete dayalı başarı farklılıklarının küçük olduğu ve bununla birlikte, problem çözümede lisedeki kızların performansının erkeklerinkinden düşük olduğu sonucuna varmışlardır. Bu yönüyle SYT'nin bulgularına baktığımızda erkeklerin lehine DMF içeren maddelerin benzer şekilde problem çözme alanında olduğu görülmektedir (Hyde, J. S., Fennema, E., & Lamon, S. J., 1990).

CP Benbow, ve JC Stanley tarafından yürütülen çalışmada ise 9927 entelektüel yetenekli ortaokul öğrencisi üzerinde yürütülen çalışmada, matematiksel akıl yürütme yeteneğinde (Scholastic Aptitude Test-Matematik testi) erkek öğrenciler lehine önemli bir cinsiyet farkı bulunmuştur (Benbow & Stanley, 1980).

Başka bir araştırmada TIMSS matematik testlerindeki büyük cinsiyet farklılıklarının orta öğretimin bitimine kadar var olduğu görülmüştür. Ortaöğretim öğrencileri için cinsiyete göre sonuçlar ortaöğretim öncesi sınıflarda olanlar ile farklılık göstermiştir. Ülkelerin çoğunda, matematik okuryazarlığı ve ileri matematikte

erkekler kızlardan daha fazla ortalama başarı oranına sahip olduğu sonucuna varılmıştır.

Sayısal yetenekteki cinsiyet farklılıkları üzerine yapılan tüm arařtırmalara bakıldığında bu tez alıřmasındaki bulguların literatürdeki konu ile ilgili yapılan alıřmalara zenginlik katacağı düşünölmektedir.

5.6. Öneriler

Ulusal ve uluslararası uygulanan testler ile öđrencilerin, başarı düzeyleri ve belirli bir alandaki yeterlilik düzeyleri belirlenmeye alıřılmaktadır. Bu kapsamda yapılan testlerde yetenek kestiriminin en dođru řekilde yapılması önem arz etmektedir. Türkiye’de yapılan sınavların bir ön denemesi yapılmamaktadır. Dolayısıyla yanlılık gösteren maddelerin testten ıkarılması mümkün olmamaktadır. Yanlılık gösteren maddelerin teste dahil edilmemesi öđrencilerin başarı ve yeterlilik düzeyi belirlenmesinde önemli rol oynamaktadır. Bu tür yanlılık gösteren maddeleri kestiren modellerin kullanılması ve testin uyumsuzluk yaratan maddelerden arındırılarak uygulanması önerilmektedir.

Bu arařtırmada, modele uyumsuzluk gösteren maddenin 17. Madde olduđu ortaya konmuřtur. Bu maddenin testten atılarak testin tekrar uygulanması önerilmektedir. Ayrıca yerel bađımsızlık varsayımını ihlal eden madde iftinin birleřtirilerek tek bir madde olarak teste konulup analizlerin tekrar edilmesi önerilebilir.

Bu arařtırmada Rasch modeline göre DMF incelemesi yapılmıř ve sonuçlar buna göre raporlanmıřtır. Daha geniř bir raporlama için diđer DMF inceleme yöntemleri kullanılarak maddelerin DMF ierip iermedikleri arařtırılabilir.

Nihai testin oluřturulması için madde güçlükleri ve ayırdediciliklerine bakılmıřtır. Bu ařamada aımlayıcı ve dođrulayıcı faktör analizi de yapılarak daha ayrıntılı bir inceleme yapılması önerilebilir.

Bu alıřmada SYT’nin alt testlerindeki, eřitlik kurma-sayısal akıl yürütme alanlarındaki 2 maddenin erkekler lehine DMF ierdiđi ve geometri-görsel uzamsal alanlarındaki 3 maddenin ise kızlar lehine DMF ierdiđi tespit edilmiřtir. Literatürde maddelerin DMF göstermesinin sebebinin kız ve erkek öđrencilerin farklı ilgi alanları

ve yeteneklere sahip olmasından kaynaklandığı veya genetik olarak farklı düşünme yapılarına sahip olmalarından kaynaklandığı gibi birçok araştırma vardır. Bu çalışmada DMF içeren maddeler tespit edilmiş fakat nedenleri üzerinde durulmamıştır. Bu maddelerin DMF içermelerinin nedenleri araştırılabilir.

Bu çalışma SYT verisi ve cinsiyet değişkeni ile sınırlı tutulmuştur. Araştırmaya başka değişkenler de eklenebilir veya kültürel ya da yaş grupları ile ilgili çalışmalar da yapılabilir. Ayrıca SYT sadece İzmir ilindeki 11. Sınıf öğrencilerinden ilk %8 lik dilim ve üzeri puan alan öğrencilere uygulanmıştır. Uygulama için örneklem büyüklüğü daha geniş tutulabilir ve yüzdeler dilim sınırlaması konulmayabilir.

Son olarak DMF çalışmalarının testlerin geçerlik çalışmaları kapsamında tüm ulusal sınavlar için rutin olarak yapılması sağlanmalıdır.

KAYNAKLAR

- Adams, R. J., & Rowe, K. J. (1988). Item bias. *Educational research, methodology, and measurement: an international handbook*, 398–403.
- Atılğan, H., Kan, A., & Doğan, N. (2009). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Ayan, C. (2011). *PISA 2009 Fen okuryazarlığı alt testinin değişen madde fonksiyonu açısından incelenmesi*. Yüksek lisans tezi, Hacettepe üniversitesi, Ankara.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: fact or artifact *science*, 210 (4475), 1262 LP – 1264.
- Bruin, D. (May 2012). [Examining dimensionality within the Rasch framework.] Unpublished raw data.
- Burnett, S. A., Lane, D. M., & Dratt, L. M. (1979). Spatial visualization and sex differences in quantitative ability. *intelligence*. [https://doi.org/10.1016/0160-2896\(79\)90003-5](https://doi.org/10.1016/0160-2896(79)90003-5)
- Büyükatak, E. (2016). *Pisa 2012 deki matematiğe yönelik duyuşsal özelliklerin bölge, okul türü ve cinsiyete göre sınıflama doğruluğunu incelenmesi*. Yüksek lisans tezi, Hacettepe Üniversitesi, Ankara.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. SAGE Publications.
- Casey, B., Nuttal, R., Byrnes, J., & Pajares, F. (n.d.) (2005). *Gender Differences in mathematics*. 102, 143-144. DC: Cambridge University Press.
- Cohen, A.S., & Ibarra, R.A. (2004). Gender differences in mathematics: *examining gender-related differential item functioning using insights from psychometric and multicontext theory*, Ann M. Mallagher(Eds). New York, DC: Cambridge University Press.
- Conaghan, P. G. (2007). The Rasch Measurement Model in Rheumatology : *What is it and why use it ? when should it be applied , and what should one look for in a rasch paper*, 57(8), 1358–1362. <https://doi.org/10.1002/art.23108>.

- Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. *Introduction to Classical and Modern Test Theory*, 527. Ohio. DC: Cengage Learning.
- Çelik, M. (2019). *PISA 2015 Matematik alt testinin eğitim bölgeleri ve cinsiyete göre değişen madde fonksiyonunun incelenmesi*. Yüksek lisans tezi, Gaziantep Üniversitesi, Gaziantep.
- Dadaş, Ö.F., (2018). *Sırasız eşik değerlerinin birleştirilmesinde farklı stratejilerin rasch modeline uyum üzerindeki etkisi*. Yüksek lisans tezi, Ege Üniversitesi, İzmir.
- Devine, P. J., & Raju, N. S. (1982). Extent of overlap among four item bias methods. *educational and psychological measurement*, 42(4), 1049–1066.
<https://doi.org/10.1177/001316448204200412>
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P.W. Holland & H.Wainer (Eds.), *Differential item functioning* (pp. 35–66).
- Dwyer, C. A., Gallagher, A., Levin, J., & Morley, M. E. (2003). What is quantitative reasoning defining the construct for assessment purposes. *Educational Testing Service*, (November), 54.
- Eccles, J. S., & Jacobs, J. E. (1986). Social Forces shape math attitudes and performance. *signs: journal of women in culture and society*, 11(2), 367–380.
- Gallagher, A. M., & Kaufman, J. C. (2000). Gender differences in mathematics: an integrative psychological approach. *Feminism & Psychology*. Ann M. Mallagher(Eds). New York, DC: Cambridge University Press.
- Halpern, D F. (2013). *Sex Differences in Cognitive Abilities: 4th Edition*. Taylor & Francis.
- Halpern, Diane F. (2012). Sex differences in cognitive abilities, 1–458.
- Halpern, Diane F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The Science of sex differences in science and mathematics. *psychological science in the public interest, supplement*, 8(1), 1–51.
- Hambleton, R K, Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. *Contemporary Sociology*, 21(2), 289.

- Hambleton, Ronald K., & Swaminathan, H. (n.d.), 1991. *Item response theory principles and applications*.
- Hanbleton, R. K., Swaminathan, H., & Rogers, D. J. (n.d.), 1991. *Fundamentals of item response theory*.
- Holland, P. W., & Thayer, D. T. (1986a). Differential item functioning and the Mantel-Haenszel procedure. *ETS Research Report Series, 1986(2)*, i–24.
- Holland, P. W., & Thayer, D. T. (1986b). Differential item functioning and the mantel-haenszel procedure. *ETS Research Report Series, 1986(2)*.
- Hyde, J.S., Fennema, E., & Lamon, S.J. (1990). Gender differences in mathematics performance: a meta-analysis. *psychological bulletin*.
- Kline, P. (2000). *The handbook of psychological testing. Filozofska Istrazivanja* (Vol. 25).
- Köse, M. (2012). *Pisa 2003, 2006 ve 2009 Türkiye uygulaması matematik ortak maddelerindeki başarıların incelenmesi*. Yüksek lisans tezi, Hacettepe Üniversitesi, Ankara.
- Leahey, E., & Guo, G. (2001). Gender differences in mathematical trajectories. *Social Forces, 80(2)*, 713–732.
- Lim, R. G., & Drasgow, F. (1990). Evaluation of two methods for estimating item response theory parameters when assessing differential item functioning. *Journal of Applied Psychology, 75(2)*, 164–174.
- Mayes, R. L., Peterson, F., & Bonilla, R. (2013). Numeracy advancing education in quantitative literacy quantitative reasoning learning progressions for environmental science: developing a framework quantitative reasoning learning progressions for environmental science: developing a framework. *Numeracy, 6(1)*.
- Mellenbergh, G. J. (1989). Item bias and item response theory. *International journal of educational research, 13(2)*, 127–143.
- Murphy, K. R., & Davidshofer, C. O. (2005). *Psychological testing: principles and applications*. Pearson/Prentice Hall.
- Osterlind, S. J. (1983). *Test item bias*. SAGE Publications.

- Öğrenci Seçme ve Yerleştirme Merkezi (2017-Ağustos). LYS 2017 sonuçlarına ait sayısal bilgiler. <https://www.osym.gov.tr/TR,13261/2017-lys-sonuclarina-iliskin-sayisal-bilgiler.html>.
- Öğretmen, T., & Doğan, N. (2004). OKÖSYS Matematik alt testine ait maddelerin yanlılık analizi. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 5(8), 2–12.
- Öğretmen T., (1995). *Differential item functioning (DIF) analysis of the verbal ability section of the first stage of the university entrance examination in Turkey*. Yüksek lisans tezi, Ortadoğu Teknik Üniversitesi, Ankara.
- Özalp Ateş, F.S., (2015). *Ölçeklerde yapı geçerliliğinin değerlendirilmesinde faktör analizi ve Rasch analizi yaklaşımları*. Yüksek lisans tezi, Ankara üniversitesi, Ankara.
- Perry, R. J., Samuel, V. T., Petersen, K. F., Shulman, G. I., Haven, N., & Haven, N. (2015). HHS Public Access, 510(7503), 84–91.
- Price, E. (1987). Document resume ed 281 883 tm 870 291.
- Raju, N. S. (1990a). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14(2), 197–207.
- Raju, N. S. (1990b). Determining the significance of estimated signed and unsigned areas between two item response functions. *Applied Psychological Measurement*, 14(2), 197–207.
- Rudner, L. M., Getson, P. R., & Knight, D. L. (1980). Biased item detection techniques. *Journal of Educational Statistics*, 5(3), 213–233.
- Sanders, B., Soares, M. P., & D'Aquila, J. M. (1982). The sex difference on one test of spatial visualization: a nontrivial difference. *Child Development*, 53(4), 1106.
- Shepard, L. A., Camilli, G., & Williams, D. M. (1985). Validity of approximation techniques for detecting item bias. *Journal of Educational Measurement*, 22(2), 77–105.
- Smith, E., (2019). "Practical Rasch Measurement - Further Topics" Paper presented at statistic.com an online course from The Institute for Statistics Education. Retrieved June 28, 2019, from <https://www.statistics.com/practical-rasch-measurement-further/#requirement>.

- Tennat, A.& Conaghan, P.G., (2007). The Rasch Measurement Model in Rheumatology : *What is it and why use it ? when should it be applied, and what should one look for in a rasch paper*, 57(8), 1358–1362.
<https://doi.org/10.1002/art.23108>.
- Thomas, H. (1993). A Theory explaining sex differences in high mathematical ability has been around for some time. *Behavioral and Brain Sciences*, 16(01), 187.
<https://doi.org/10.1017/S0140525X00029575>
- Toker, F. (1968). *Zeka kuramlari*. Ankara: Milli Egitim Bakanligi Talim ve Terbiye Dairesi Arastirma ve Degerlendirme Buros.
- Turgut, M. F., & Baykul, Y. (2011). *Eđitimde olęme ve deęerlendirme*. Pegem Akademi.
- Wu, M., & Adams, R. (2007). Applying the rasch model to psycho-social measurement a practical approach.
- Yenal, E. (1995). *Öęrenci Seęme Sınavı Sayısal Testi `nin madde yanlılıęı analizi*.Yüksek lisans tezi, Ortadoęu Teknik Üniversitesi, Ankara.
- Yu, C.H. (2017), A simple guide to the item response theory (IRT) and Rasch modeling.
- Yüksel C., (2012). *Ölęeklerde saptanan madde işlev farklılıęının karma Rasch modelleri ile incelenmesi*. Doktora tezi, Ankara Üniversitesi, Ankara.
- Zumbo, B. (1999). A handbook on the theory and methods of differential item functioning (DIF). *Ottawa: National Defense Headquarters*, 1–57.

EKLER

Ek 1: Sayısal Yetenek Testi(SYT)

LÜTFEN BU DÖKÜMANI DİKKATLİCE OKUMAK İÇİN ZAMAN AYIRINIZ

Sizi Yasemin Yardım Palamutlu tarafından yürütülen "Sayısal Yetenek Farklılaşmasının Cinsiyete Göre İncelenmesi: Bir Değişen Madde Fonksiyonu(DMF) Çalışması" başlıklı araştırmaya davet ediyoruz. Bu araştırmaya katılıp katılmama kararını vermeden önce, araştırmanın neden ve nasıl yapılacağını bilmeniz gerekmektedir. Bu nedenle bu formun okunup anlaşılması büyük önem taşımaktadır. Eğer anlayamadığınız ve sizin için açık olmayan şeyler varsa, ya da daha fazla bilgi isterseniz bize sorunuz.

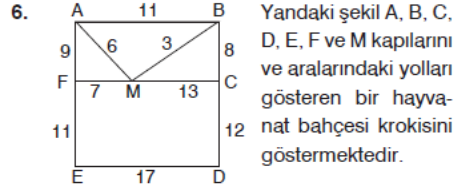
Bu çalışmaya katılmak tamamen gönüllülük esasına dayanmaktadır. Çalışmaya katılmama veya katıldıktan sonra herhangi bir anda çalışmadan çıkma hakkında sahipsiniz. Çalışmayı yanıtlamanız, araştırmaya katılım için onam verdiğiniz biçiminde yorumlanacaktır. Size verilen formlardaki soruları yanıtlarken kimsenin baskısı veya telkini altında olmayın. Bu formlardan elde edilecek kişisel bilgiler tamamen gizli tutulacak ve yalnızca araştırma amacı ile kullanılacaktır.

iletisim: yaseminyyp@gmail.com

Yasemin YARDIM PALAMUTLU
Sosyal Bilimler Enstitüsü Eğitim Fakültesi
Eğitim Bilimleri Bölümü
Ölçme ve Değerlendirme Anabilim Dalı
(Tezli Yüksek Lisans Programı)
Yüksek Lisans Öğrencisi

SAYISAL YETENEK TESTİ

- B \neq 0 olmak üzere iki basamaklı AB pozitif tamsayıları kümesi üzerinde F bağlantısı $F(AB) = \frac{AB}{A}$ şeklinde tanımlıyor.
Örneğin $F(26) = \frac{26}{2} = 13$
Buna göre $F(AB) = 12$ eşitliğini sağlayan kaç tane AB sayısı vardır?
A) 2 B) 3 C) 4 D) 5 E) 6
- 67 yolcunun bulunduğu bir metro istasyonunda 35 tane bayan vardır. Gözlüklü yolcuların sayısı 16'dır. Bayan yolculardan en az 2 tanesinin, erkek yolculardan en az 1 tanesinin gözlüklü olduğu bilinmektedir.
Gözlüklü olmayan erkek yolcu sayısı en az kaç tane olabilir?
A) 13 B) 15 C) 17 D) 18 E) 19
- $x > 4$ olmak üzere x tane kişi aşağıdaki kurallara göre bir oyun oynamaya başlamışlardır.
1. Herhangi bir elde oyunu kazanan kişiye kaybedenlerin her biri o ele katılan kişi sayısı kadar çay ısmarlar.
2. Bir elde oyunu kazanan kişi oyundan çıkar ve oyun geride kalan kişilerle oynamaya devam eder.
4. elde oyunu kazanan kişiye, kaybedenler toplam 72 tane çay ısmarladığına göre oyunun başlangıcındaki kişi sayısı kaçtır?
A) 11 B) 12 C) 13 D) 14 E) 15
- 0, 2, 4, 6, 8 rakamları birer kez kullanılarak oluşturulan birbirinden farklı tüm 5 basamaklı sayılar büyükten küçüğe doğru sıralanıyor.
Bu sıralamada en büyük sayı ile en küçük sayı arasındaki fark kaçtır?
A) 55662 B) 54952 C) 65962
D) 66652 E) 65952
- Bir araç 1 litre benzin ile ortalama 13 km yol gitmektedir.
Benzinin litre fiyatı 5,27 TL ise bu aracın 2652 km yol gitmek için kaç TL lik benzine ihtiyacı vardır?
A) 1075,08 B) 1175,08 C) 1165,04
D) 1065,08 E) 1075,04



- Hayvanat bahçesine A, B, C, D, E, F ve M kapılarının herhangi birisinden girilebilir.
- Krokide gösterilen yolları kullanırken geçilen bir yoldan tekrar geçilemez, fakat istenilen yönde ve şekilde hareket edilebilir.
- Girilen kapı çıkış için kullanılamaz.
- Bir noktadan diğerine ulaşıldığında krokide gösterilen kadar hayvan görülmüş olur.

Hayvanat bahçesine girip 20 hayvan görüp hayvanat bahçesinden çıkmış olan biri hangi kapıyı ne girişte ne de çıkışta kullanmış olamaz?

- A) A B) M C) E D) D E) F

7.

| | Arda | Mert | Zeynep | Polen | Orhan |
|---------|------|------|--------|-------|-------|
| Ağırlık | 27 | 45 | 51 | 63 | 69 |

Yukarıdaki tabloda Arda, Mert, Zeynep, Polen ve Orhan'ın ağırlıkları kg cinsinden verilmiştir. Bu beş kişi bir miktar elmayı aşağıdaki kurallara göre paylaşıyorlar.

- İlk olarak elmaları bu beş kişi aralarında eşit paylaşıyorlar.
- Daha sonra bu kişiler ağırlıkça kendinden ağır olanlara her 1 kg farkı için 3 elma veriyor.
- Paylaşım sonucunda Arda'nın 20 elması oluyor.

Paylaşım sonucunda kimin elmalarının sayısı ilk paylaşımındaki elmaların sayısıyla aynıdır?

- A) Arda B) Mert C) Zeynep
D) Polen E) Orhan

(8, 9 ve 10. soruları aşağıdaki bilgilere göre cevaplayınız.)

Bir baba dikdörtgen şeklindeki tarlasını Ayşe, Fatma, Ali ve Veli isimli çocuklarına miras olarak 4 parçaya ayırıp aşağıdaki gibi paylaşmıştır. Çocuklar da tarlaya A, B, C ve D ürünlerini ekmişlerdir.

| | | |
|-------|-----|------|
| Ayşe | Ali | Veli |
| Fatma | | |

- Ayşe'nin tarlasının alanı, Fatma'nın tarlasının alanına eşittir.
- Veli'nin tarlasının alanı Ali'ninkinin 2, Ayşe'ninkinin 4 katı büyüklüktedir.
- Ayşe tarlasına A ürününü ektiğinde 600 kg, B ürününü ektiğinde 360 kg, C ürününü ektiğinde 180 kg ve D ürününü ektiğinde ise 100 kg ürün almaktadır.
- Tüm tarlaların verimi eşittir.
- Her tarlaya farklı ürün ekilmekte ve boş tarla kalmamaktadır.

8. Ali'nin ve Veli'nin aynı miktarda ürün aldığı bir yılda Ali ve Veli tarlalarına sırasıyla aşağıdakilerden hangisinde verilen ürünleri ekmişlerdir?

- A) C - D B) B - C C) C - B
D) D - C E) A - B

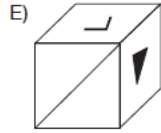
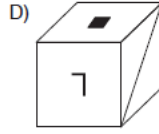
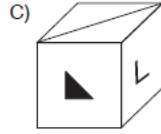
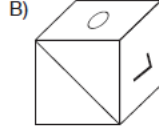
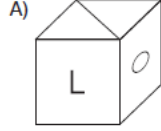
9. Verilen bilgilere göre tarlalardan **en fazla** toplam kaç kg ürün alınabilir?

- A) 3400 B) 4600 C) 4800
D) 5000 E) 5100

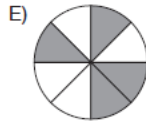
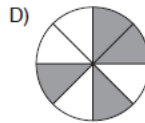
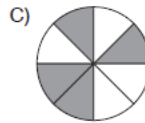
10. Verilen bilgilere göre tarlalardan **en az** toplam kaç kg ürün alınabilir?

- A) 1600 B) 1700 C) 1720
D) 1800 E) 1860

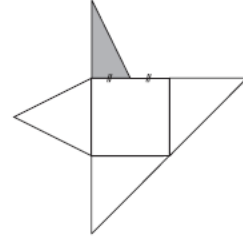
11. Yandaki küpün her yüzünde farklı bir şekil olduğuna göre, aşağıdaki küplerden hangisi bu küpün eşi olabilir?



12. Yandaki şekil saat yönünde 180° döndürüldüğünde aşağıdaki şekillerden hangisi oluşur?



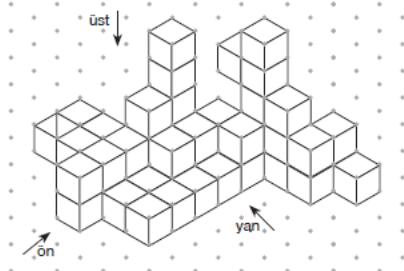
- 13.



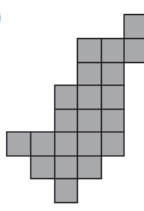
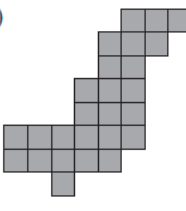
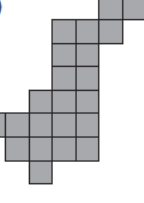
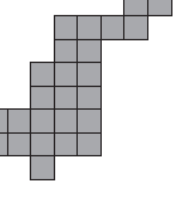
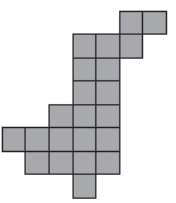
Yukarıdaki şeklin bir küpün açılmış hali olacak şekilde tamamlanabilmesi için her biri taralı üçgenin alanına denk olan kaç parçaya daha ihtiyaç vardır?

- A) 9 B) 10 C) 12 D) 13 E) 14

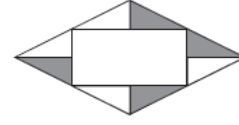
14.



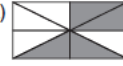
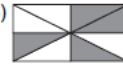
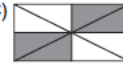
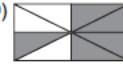
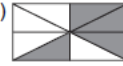
Yanda verilen şeklin üstten görüntüsü aşağıdakilerden hangisidir?

- A) 
- B) 
- C) 
- D) 
- E) 

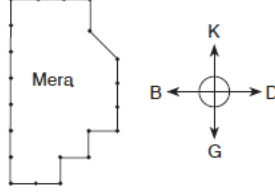
15.



Yukarıda açılmış biçimi verilen şeffaf zarf kapatıldığında aşağıdakilerden hangisi elde edilir?

- A) 
- B) 
- C) 
- D) 
- E) 

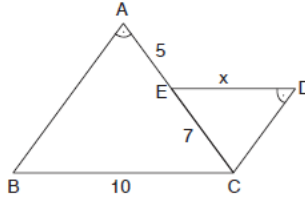
16. Aşağıdaki şemada, çitle çevrili bir mera görülmektedir. Çitin bir bölümü hariç, kuzey - güney veya doğu - batı yönünde düzdür. Ardaşık çit direkleri arası, bir diagonal (eğik) kısım haricinde, 10 m'dir.



Aşağıdakilerden hangisi mera çevresini (Ç) metre cinsinden en iyi tanımlar?

- A) $\text{Ç} > 210$ B) $\text{Ç} = 210$ C) $\text{Ç} < 210$
D) $\text{Ç} > 230$ E) $\text{Ç} = 240$

17.



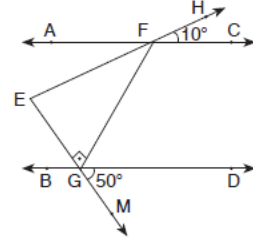
$BC \parallel ED$
 $m(\widehat{BAC}) = m(\widehat{EDC})$
 $|AE| = 5\text{cm}$, $|EC| = 7\text{cm}$, $|BC| = 10\text{cm}$
ise şekilde verilenlere göre $|ED| = x$ kaç cm'dir?

- A) 6 B) 8 C) $\frac{42}{5}$ D) $\frac{50}{7}$ E) 9

18. $y = x$ doğrusu için aşağıdakilerden hangisi doğru değildir?

- A) $y = x$ doğrusu ile $y = x + 5$ doğrusu paraleldir.
B) $y = x$ doğrusu ile $y = -2x$ doğrusu diktir.
C) $y = x$ doğrusunun eğimi 1 dir.
D) Orjinden geçer.
E) $y = x$ doğrusu ile $y = -x$ doğrusu arasındaki açı 90° dir.

19.

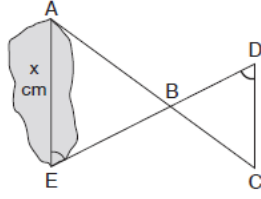


AC ve BD doğruları paraleldir.
 $m(\widehat{HFC}) = 10^\circ$ $m(\widehat{DGM}) = 50^\circ$, \widehat{EFG} üçgeni bir dik üçgendir ve $|EG| = 10\text{ cm}$ 'dir.

Verilenlere göre $|EF|$ uzunluğu kaç cm'dir?

- A) 12
B) Verilen bilgiler yetersizdir.
C) 5
D) 15
E) 20

20.



Bir kamp rehberi şeklinde görülen bataklığın üzerinden geçen A ve E noktaları arasındaki x cm uzunluğundaki bir köprünün uzunluğunu hesaplamak istiyor.

$|AB| = 1800$ cm, $|EB| = 1400$ cm,


$|BD| = 700$ cm ve $|CD| = 800$ cm dir.

AC ve DE doğru parçaları da B noktasında kesişmektedir.

$m(\widehat{AEB}) = m(\widehat{CDB})$ ise $|AE| = x$ kaç cm dir?

- A) 1200 B) 1600 C) 1700
D) 1800 E) 2000

Ek 2:SYT'nin uygulama izni


T.C.
KARŞIYAKA KAYMAKAMLIĞI
İlçe Milli Eğitim Müdürlüğü

Sayı : 47133873-604.01.01-E.2973031
Konu : Yasemin YARDIM
PALAMUTLU'nun Araştırma İzni

11.02.2019

.....MÜDÜRLÜĞÜNE
KARŞIYAKA

İlgi: a) İl Milli Eğitim Müdürlüğü' nün, 08/02/2019 tarih ve 2837306 sayılı yazısı.
b) MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü'nün 22/08/2017 tarihli ve 12607291 sayılı yazısı (Genelge 2017/25)
c) Ege Üniversitesi Rektörlüğü'nün 04/12/2018 tarihli ve 102436 sayılı yazısı.
d) 08/02/2019 tarihli ve 2763690 sayılı Valilik Onayı.

Ege Üniversitesi Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme tezli yüksek lisans öğrencisi Yasemin YARDIM PALAMUTLU'nun "Sayısal Yetenek Farklaşmasının Cinsiyete Göre İncelenmesi: Bir Değişen Madde Fonksiyonu (DMF) Çalışması" konulu tez çalışması için kullanacağı ölçekleri okulunuzda uygulanma isteği, (d) Valilik Onayı ile uygun görülmüştür.

Söz konusu ölçeklerin uygulanmasının, Müdürlüğümüze bağlı tüm okullarda 2018-2019 eğitim yılında eğitim öğretimi aksatmayacak ve eğitim kurumu yöneticilerinin uygun gördüğü şekilde, araştırma yapılmadan önce araştırmanın yapılacağı okullar tarafından "Milli Eğitim Bakanlığına Bağlı Her Tür Okul ve Kurumlarda Yapılmasına İzin Verilen Araştırma Uygulamasında, Olabilecek Zararları Karşılama Taahhüdü" adlı ekin araştırmacı tarafından doldurulması gerekmektedir.

Bilgilerinizi ve gereğini rica ederim.

Mehtap GEMİCİ
Müdür a.
Şube Müdürü

EKLER:
İlgi Yazı ve Ekleri (12 sayfa)

DAĞITIM
Karşıyaka Lisesi
15 Temmuz Şehitler And. Lisesi
Atakent Anadolu Lisesi
Cihat Kora And. Lisesi

Bilgi için: Birgül YAZAR Memur (76)