

T.C.
GAZİANTEP ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANABİLİM DALI

**2012-SEVİYE BELİRLEME SINAVININ RASCH
MODELİNE GÖRE CİNSİYET DEĞİŞKENİ AÇISINDAN
YANLILIĞININ İNCELENMESİ**

YÜKSEK LİSANS TEZİ

Azmi TÜRKAN

Gaziantep
Ocak 2014

T.C.
GAZİANTEP ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANABİLİM DALI

**2012-SEVİYE BELİRLEME SINAVININ RASCH
MODELİNE GÖRE CİNSİYET DEĞİŞKENİ AÇISINDAN
YANLILIĞININ İNCELENMESİ**

YÜKSEK LİSANS TEZİ

Azmi TÜRKAN

Tez Danışmanı: Doç. Dr. Bayram ÇETİN

Gaziantep
Ocak 2014

T.C.
GAZİANTEP ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
EĞİTİM BİLİMLERİ ANABİLİM DALI

**2012-Seviye Belirleme Sınavının Rasch Modeline Göre
Cinsiyet Değişkeni Açısından Yanlılığının İncelenmesi**

Azmi TÜRKAN

Tez Savunma Tarihi: 16.01.2014

Eğitim Bilimleri Enstitüsü Onayı

Doç. Dr. M.Fatih ÖZMANTAR

Enstitüsü Müdürü

Bu tezin Yüksek Lisans tezi olarak gerekli şartları sağladığımı onaylarım.

Prof.Dr. Zeynep HAMAMCI

Enstitü ABD Başkanı

Bu tez tarafımda okunmuş, kapsamı ve niteliği açısından bir Yüksek Lisans tezi olarak kabul edilmiştir.

Doç. Dr. Bayram ÇETİN

Tez Danışmanı

Bu tez tarafımızca okunmuş, kapsam ve niteliği açısından bir Yüksek Lisans tezi olarak kabul edilmiştir.

Jüri Üyeleri:

İmzası

Doç. Dr. Bayram ÇETİN (Jüri Başkanı)

Doç. M. Fatih ÖZMANTAR

Yrd. Doç. Dr. Yeşim ÖZER ÖZKAN

ÖN SÖZ

Bu noktalara gelebilmemde emeđi olan bařta tez danıřanım Doç. Dr. Bayram ETİN olmak üzere eđitim ve öđretimde hayatımdaki bütün hocalarıma ve her türlü yardımlarını esirgemeyen aile sonsuz teřekkürler.

Azmi TÜRKAN

Ocak 2014

ÖZET

2012-SEVİYE BELİRLEME SINAVININ RASCH MODELİNE GÖRE CİNSİYET DEĞİŞKENİ AÇISINDAN YANLILIĞININ İNCELENMESİ

TÜRKAN, Azmi
Yüksek Lisans Tezi, Eğitim Bilimleri ABD
Tez Danışmanı: Doç. Dr. Bayram ÇETİN
Ocak 2014, 81 sayfa

Bir testte bulunması gereken en önemli özelliklerin başında geçerlik ve güvenilirlik gelmektedir. Testin geçerlik ve güvenilirliğini sağlamanın bir basamağı test maddelerinin yanlılığını incelemektir. Bu çalışmanın amacı 2012 Seviye Belirleme Sınavı (SBS) maddelerinin cinsiyet değişkenine göre Rasch Modeli açısından yanlılığını incelemektir. Çalışma kapsamında 2012-SBS *Türkçe, Matematik, Fen ve Teknoloji* ve *Sosyal Bilgiler* alt testleri incelenmiştir. Çalışmanın örneklemini 2012 yılında SBS'ye katılmış olan 1075546 kişi arasından tabakalı örneklem orantılı seçim yöntemi ile rastgele seçilen 216363 kişi oluşturmuştur. Verilerin analizinde Winsteps 3.8.1 paket programı kullanılmıştır. Yanlılığı belirlemek amacıyla Rasch modeli temel alınarak Değişen Madde Fonksiyonu (DMF) içeren maddeler incelenmiştir. Belirlenen DMF içeren maddelerin madde etkisinden mi, yanlılıktan mı kaynaklandığını tespit etmek için uzman görüşüne başvurulmuştur. Yapılan istatistiksel işlemler sonucunda iki tane Matematik alt testi ve bir tane Sosyal Bilgiler alt testi olmak üzere toplam üç maddenin Değişen Madde Fonksiyonu içerdiği tespit edilmiştir. Bu üç maddenin yanlılığı ile ilgili uzman görüşlerine göre ise maddelerde yanlılık bulunmamaktadır. Diğer bir ifadeyle DMF içeren maddelerin tamamında madde etkisi bulunmaktadır. Sonuç olarak 2012 SBS *Türkçe, Matematik, Fen ve Teknoloji* ve *Sosyal Bilgiler* alt testlerinin yanlı soru içermediği, madde yanlılığı bakımından geçerli ve güvenilir bir test olduğu saptanmıştır.

Anahtar Kelimeler:

Seviye Belirleme Sınavı, Yanlılık, Değişen Madde Fonksiyonu, Rasch Modeli

ABSTRACT**AN INVESTIGATION OF ITEM BIAS IN 2012-LEVEL DETERMINATION
EXAM ACCORDING TO RASCH MODEL IN TERMS OF GENDER
VARIABLE**

TÜRKAN, Azmi

Master's Thesis, Department of Educational Sciences

Advisor: Assoc. Prof. Dr. Bayram ÇETİN

January 2014, 81 pages

Validity and reliability are two of the most important characteristics of a test. One of the steps to ensure validity and reliability of the test is to investigate item bias. The purpose of the current study is to investigate the differential item functioning of the items of 2012 Level Determination Exam according to Rasch Model in terms of gender variable. To this end, *Turkish, Mathematics, Science and Technology* and *Social Studies* sub-tests of 2012 Level Determination Exam (LDE) were investigated. The sample of the study is comprised of 216363 students chosen at random from among 1075546 students who attended LDE via proportional stratified sampling method. Winsteps 3.8.1 packet program was used to analyze data. To determine item bias, items including Differential Item Functioning (DIF) were examined based on Rasch Model. A field specialist's views were taken to reveal whether the items determined to have included DIF stemmed from item impact or item bias. As a result of the statistical analyses, it was found that only three items, two in the Mathematics sub-test and one in the Social Studies sub-test, had differential item functioning (DIF) in total. According to the field specialist's views, the items did not include item bias. In other words, all three items including DIF had item impact. It was concluded that *Turkish, Mathematics, Science and Technology* and *Social Studies* sub-tests of 2012 Level Determination Exam (LDE) did not include differential item functioning, and that LDE was a valid and reliable test in terms of item bias.

Keywords:

Level Determination Exam, Item Bias, Differential Item Functioning, Rasch Model

İÇİNDEKİLER

ÖN SÖZ.....	i
ÖZET.....	ii
ABSTRACT	iii
İÇİNDEKİLER.....	iv
TABLolar LİSTELERİ	vi
ŞEKİLLER LİSTESİ.....	vii
KISALTMALAR LİSTESİ.....	ix
BİRİNCİ BÖLÜM.....	1
GİRİŞ	1
1.1.PROBLEM DURUMU.....	1
1.2.ARAŞTIRMA PROBLEMİ.....	3
1.3.ARAŞTIRMANIN ALT PROBLEMLERİ	3
1.4.ARAŞTIRMANIN AMACI	4
1.5.ARAŞTIRMANIN ÖNEMİ	4
1.6.SINIRLILIKLAR.....	4
İKİNCİ BÖLÜM.....	6
KAYNAK ÖZETLERİ	6
2.1.MADDE VE TEST YANLILIĞI	6
2.2.YANLILIĞIN TARİHÇESİ.....	6
2.3.DEĞİŞEN MADDE FONKSİYONU (DMF)	7
2.4.DMF İLE YANLILIK ARASINDAKİ FARK.....	8
2.5.YANLILIK BELİRLEME YÖNTEMLERİ.....	10
2.5.1.Klasik Test Kuramı ve DMF Belirleme Yöntemleri	10
2.5.2.Madde Tepki Kuramı Test Kuramı ve DMF Belirleme Yöntemleri	12
2.6.SEVİYE BELİRLEME SINAVI (SBS).....	14

2.7.YURTİÇİNDE YAPILAN ÇALIŞMALAR.....	16
2.8.YURTDIŞINDA YAPILAN ÇALIŞMALAR.....	18
ÜÇÜNCÜ BÖLÜM.....	22
MATERYAL VE YÖNTEM.....	22
3.1.ARAŞTIRMANIN MODELİ.....	22
3.2.EVREN VE ÖRNEKLEM.....	23
3.3.VERİLERİN ELDE EDİLMESİ.....	23
3.4.VERİLERİN ÇÖZÜMLENMESİ.....	24
DÖRDÜNCÜ BÖLÜM.....	26
BULGULAR VE TARTIŞMA.....	26
4.1.BULGULAR.....	26
4.1.2.Normallik.....	26
4.1.2.Tek Boyutluluk.....	29
4.1.3.Yerel Bağımsızlık.....	35
4.1.4.Türkçe Alt Testine İlişkin Bulgular.....	36
4.1.5.Matematik Alt Testine İlişkin Bulgular.....	42
4.1.6.Fen ve Teknoloji Alt Testine İlişkin Bulgular.....	48
4.1.7.Sosyal Bilgiler Alt Testine İlişkin Bulgular.....	53
4.1.8. Değişen Madde Fonksiyonu (DMF) İçeren Maddelere İlişkin Uzman Görüşlerine ait Bulgular.....	59
4.2.TARTIŞMA.....	60
BEŞİNCİ BÖLÜM.....	68
SONUÇ VE ÖNERİLER.....	68
5.1.SONUÇ.....	68
5.2.ÖNERİLER.....	68
KAYNAKLAR.....	70
EKLER.....	77
Ek-1:Verilerin Milli Eğitim Bakanlığında Alınmasına Dair Dilekçe.....	77
Ek-2: Yanlığa İlişkin Uzman Görüşü.....	78
ÖZGEÇMİŞ (VITAE).....	81

TABLOLAR LİSTELERİ

Tablo 1: SBS-2012 alt testleri soru sayısı ve soru ağırlık katsayısı.....	15
Tablo 2: Kitapçık ve cinsiyet türüne göre öğrenci evren dağılımı	22
Tablo 3: Kitapçık ve cinsiyet türüne göre öğrenci örneklem dağılımı.....	23
Tablo 4: Testlere ilişkin betimsel istatistikler.....	27
Tablo 5: Test türlerinin birey ve madde güvenilirlik değerleri.....	36
Tablo 6: Türkçe alt testine ilişkin madde güçlük düzeyi, uyum indeksi ve nokta çift serili korelasyon değerleri.....	36
Tablo 7: Türkçe alt testine ilişkin değişen madde fonksiyonu kontrast değerleri.....	40
Tablo 8: Matematik alt testine ilişkin madde güçlük düzeyi, uyum indeksi ve nokta çift serili korelasyon değerleri.....	43
Tablo 9: Matematik alt testine ilişkin değişen madde fonksiyonu kontrast değerleri	46
Tablo 10: Fen ve teknoloji alt testine ilişkin madde güçlük düzeyi, uyum indeksi ve nokta çift serili korelasyon değerleri.....	48
Tablo 11: Fen ve teknoloji alt testine ilişkin değişen madde fonksiyonu kontrast değerleri	51
Tablo 12: Sosyal bilgiler alt testine ilişkin madde güçlük düzeyi, uyum indeksi ve nokta çift serili korelasyon değerleri.....	54
Tablo 13: Sosyal bilgiler alt testine ilişkin değişen madde fonksiyonu kontrast değerleri	57

ŞEKİLLER LİSTESİ

Şekil 1: Değişen madde fonksiyonu içermeyen madde.....	9
Şekil 2: Tek biçimli değişen madde fonksiyonu içeren madde.....	9
Şekil 3: Tek biçimli olmayan değişen madde fonksiyonu içeren madde.....	10
Şekil 4: Türkçe alt testine ilişkin puan dağılımı.....	27
Şekil 5: Matematik alt testi puan dağılımı	28
Şekil 6: Fen ve teknoloji alt testi puan dağılımı.....	28
Şekil 7: Sosyal bilgiler alt testi puan dağılımı	29
Şekil 8: Türkçe alt testinin tek boyutluluğuna ilişkin doğrulayıcı faktör analiz sonuçları	30
Şekil 9: Matematik alt testinin tek boyutluluğuna ilişkin doğrulayıcı faktör analiz sonuçları	32
Şekil 10: Fen ve teknoloji alt testinin tek boyutluluğuna ilişkin doğrulayıcı faktör analiz sonuçları	33
Şekil 11: Sosyal bilgiler alt testinin tek boyutluluğuna ilişkin doğrulayıcı faktör analiz sonuçları.....	34
Şekil 12: Türkçe alt testi madde güçlük ve öğrenci dağılımı.....	39
Şekil 13: Türkçe maddelerinin cinsiyete göre DMF değişimi.....	42
Şekil 14: Matematik alt testi madde güçlük ve öğrenci dağılımı.....	45
Şekil 15: Matematik maddelerinin cinsiyete göre DMF değişimi.....	47
Şekil 16: Fen ve teknoloji alt testi madde güçlük ve öğrenci dağılımı.....	50

Şekil 17: Fen ve teknoloji maddelerinin cinsiyete göre DMF değişimi.....	53
Şekil 18: Sosyal bilgiler alt testi madde güçlük ve öğrenci dağılımı.....	56
Şekil 19: Sosyal bilgiler maddelerinin cinsiyete göre DMF değişimi.....	59
Şekil-20: Matematik alt testi 4.soru	62
Şekil-21: Matematik alt testi 19.soru	63
Şekil-22: Sosyal bilgiler alt testi 3.soru	66

KISALTMALAR LİSTESİ

SBS	: Seviye Belirleme Sınavı
DMF	: Değişen Madde Fonksiyonu
KTK	: Klasik Test Kuramı
LR	: Lojistik Regresyon
MH	: Mantel-Haenszel
MTK	: Madde Tepki Kuramı
YGS	: Yükseköğretim Giriş Sınavı
LYS	: Lisans Yerleştirme Sınavı
OKS	: Ortaöğretim Kurum Sınavı
ÖSS	: Öğrenci Seçme Sınavı
MNSQ(Mean-square)	: Ortalama Kareler Toplamı
PTMEA	: Nokta Çift Serili

BİRİNCİ BÖLÜM

GİRİŞ

2012 Seviye Belirleme Sınavı (SBS) Türkçe, Matematik, Fen Bilgisi ve Sosyal Bilgiler alt testlerinin yanlışlık durumlarının incelendiği çalışmanın birinci bölümünde öncelikle araştırmanın problem durumu, amacı, önemi, sınırlılıkları ve sayıltıları ile ilgili açıklamalara yer verilmiştir. Çalışmanın ikinci bölümünde ilgili alanyazında yer alan bilgiler doğrultusunda araştırmanın kuramsal temeli oluşturulmuştur. Çalışmanın üçüncü bölümünde kullanılan materyaller ve yöntemler açıklanarak son bölümde ise araştırmadan elde edilen sonuçlara yer verilmiştir.

1.1.PROBLEM DURUMU

Bireylerin var olan özelliklerinin saptanması veya kazandırılmak istenen özelliklerin bireye kazandırılıp kazandırılmadığı, kazandırıldıysa ne derecede kazandırıldığı eğitim-öğretimin vazgeçilmez öğelerinden biridir. Bu bağlamda ölçme ve değerlendirme kavramları eğitim-öğretim sisteminin devamını sağlayan önemli bir unsur olarak karşımıza çıkmaktadır. Eğitimde, uygulanan programın ne derecede etkili olduğu, kullanılan metotların işevurukluluk derecesini saptama, birey başarısı ve öğrenmede yaşanan zorlukları ortaya çıkarma, bireylerin becerileri doğrultusunda yönlendirme gibi temel değerlendirmelerin hepsi, ölçme sonuçlarına dayanmaktadır (Baykul, 2010).

Ölçme terimi kavramsal olarak incelendiğinde, ilgili literatürde farklı tanımlar olduğu anlaşılmaktadır. En genel anlamda ölçme, bir nesneyi, olguyu veya niteliği gözlemleyip gözlem sonuçlarını sayı veya sembollerle ifade etmektir (Kan, Atılgan ve Doğan; 2006; Turgut ve Baykul, 2010 ve Tekin, 2000). Ayrıca Crooks (1988), her bireyi bir birinden ayıran özelliklerin olduğunu, bu özelliklerin bireyleri

birbirinden farklı kıldığını ve bu farklılığın ölçme ve değerlendirme kavramlarına bir temel oluşturduğunu savunmaktadır. Yani ölçme bireylerin farklılıklarının sayı veya sembollerle ifade edilmesidir. Değerlendirme ise yapılan ölçme işlemi sonuçlarının bir ölçüt ile karşılaştırılıp bir yargıya varılmasıdır (Turgut ve Baykul, 2010; Semerci, 2009). Diğer bir ifade ile değerlendirme, ölçümler sonucu elde edilen sayı veya sembollerin anlamlandırılması olarak düşünülebilir.

Eğitimde ölçme ve değerlendirmenin temel amacı bireyler hakkında uygun ve isabetli kararlar almaktır. Bu kararların alınmasında farklı ölçme yöntemleri kullanılabilir. Gözlem, görüşme (mülakat), vak'a kayıtları (anekdot), biyografi (yaşam öyküsü), otobiyografi (özyaşam öyküsü), vb test dışı tekniklerin yanı sıra testler en sık kullanılan teknikler arasında gösterilmektedir (Yılmaz, 2009). Eğitimde test, ölçülmek istenen özelliği ölçen araç olarak nitelendirilebilir. Turgut ve Baykul'a (2010) göre eğitimde test genel anlamda 'her türlü ölçme aracı' özel anlamda ise 'çoktan seçmeli test ve doğru-yanlış testleri' olarak belirtilmektedir.

Testler, bireylerin yetenek, başarı, ilgi gibi özelliklerini tespit etmede kullanılan araçların en başında gelmesi nedeniyle sahip olması gereken bazı özellikler vardır. Micheel ve Karnes'e (1950) göre bir testte bulunması gereken özellikler: kapsamlı, ayırt edici, objektif, kullanışlı, geçerli ve güvenilir olmalıdır. Tekin (2000), Yılmaz (1998) ve Turgut'a (1995) göre ise testlerde bulunması gereken özellikler; kullanışlılık, geçerlilik ve güvenirlidir. Kullanışlılık bir testin uygulama ve puanlama kolaylığı, mali açıdan ekonomikliği olarak nitelendirilirken geçerlilik ölçülmek istenen özelliğin dışında başka şey katmadan (Micheel ve Karnes, 1950 ve Tekin, 2000) testin amaca hizmet derecesi (Yılmaz, 1998) olarak nitelendirilmiştir. Güvenirlik, ölçülmek istenen özelliğin sonuçlarının tesadüfi hatalardan arınık olma derecesidir (Turgut ve Baykul, 2010 ; Tekin, 2000 ve Demircioğlu, 2009). Testlerde bulunan hatalar geçerlilik ve güvenirlilik kavramları ile yakından alakalıdır. Testlerde bulunan hata türleri; sabit, tesadüfi ve sistematik hata olmak üzere üçe ayrılmaktadır. Sabit hata, her ölçümde miktarı değişmeyen hatalardır. Şansla ortaya çıkan, kaynağı belirsiz ve etki yönü yordanamayan hatalara tesadüfi hata denirken, bütün miktarı ve kaynağı belli olan her ölçüm sonucunda sistemli bir şekilde ortaya çıkan hata türlerine ise sistematik hata denir (Tekin, 2000).

Sistemik hatalar, ölçmeyi yapan kişilere bağlı ve her ölçme için yordanabilir niteliktedirler (Özbek, 2008). Bir sınav puanlanırken puanlayıcının cinsiyete göre az ya da çok puan vermesi yada yazısı güzel olan öğrencilere fazla puan vermesi gibi grup bazında yapılan sistemik hataları, yanlılık olarak tanımlamaktadırlar (Tekin, 2000; Turgut, 1995; Camilli ve Shepard, 1994). Bu bağlamda bir sistemik hata kaynağı olan yanlılık, testin geçerlik ve güvenilirliğini büyük ölçüde etkilemektedir.

Yanlılık, test puanlarında geçerlik kavramıyla doğrudan ilişkili olduğundan, test puanları belirli bir grup lehine avantaj ya da dezavantaj sağlamamalıdır. Bu doğrultuda yanlılık çalışmalarının yapılması, bireylerin geleceği için büyük önem arz eden seçme ve yerleştirme sınavları açısından kaçınılmaz olarak görünmektedir. Bu nedenle seviye belirleme sınavı da bireylerin geleceğinin belirlenmesinde önemli bir sınav olduğundan, bu sınav ile ilgili yanlılık çalışmalarının yapılması gerekliliği ortaya çıkmaktadır.

1.2.ARAŞTIRMA PROBLEMİ

2012 Seviye Belirleme Sınavı Türkçe, Matematik, Fen Bilgisi ve Sosyal Bilgiler alt test maddeleri cinsiyet açısından yanlılık içermekte midir?

1.3.ARAŞTIRMANIN ALT PROBLEMLERİ

1. 2012 Seviye Belirleme Sınavı Türkçe alt test maddeleri, Rasch Modeline göre cinsiyet açısından Değişen Madde Fonksiyonu (DMF) içermekte midir?

2. 2012 Seviye Belirleme Sınavı Matematik alt test maddeleri, Rasch Modeline göre cinsiyet açısından Değişen Madde Fonksiyonu (DMF) içermekte midir?

3. 2012 Seviye Belirleme Sınavı Fen Bilgisi alt test maddeleri, Rasch Modeline göre cinsiyet açısından Değişen Madde Fonksiyonu (DMF) içermekte midir?

4. 2012 Seviye Belirleme Sınavı Sosyal Bilgiler alt test maddeleri, Rasch Modeline göre cinsiyet açısından Değişen Madde Fonksiyonu (DMF) içermekte midir?

5. Değişen Madde Fonksiyonu (DMF) içeren maddeler, uzman görüşü doğrultusunda yanlılık göstermekte midir?

1.4.ARAŞTIRMANIN AMACI

Araştırmanın amacı, 2012 Seviye Belirleme Sınavı (SBS) Türkçe, Matematik, Fen Bilgisi ve Sosyal Bilgiler alt testi maddelerinin, Rasch modeline göre cinsiyet açısından Değişen Madde Fonksiyonu (DMF) içerip içermediği; içeriyorsa, DMF içeren maddelerin yanlılık gösterip göstermediğini tespit etmektir.

1.5.ARAŞTIRMANIN ÖNEMİ

Bireylerin gelecekleri ile ilgi karar vermede kullanılmış olan SBS maddelerinin geçerlik ve güvenilirlik düzeylerinin en üst düzeyde olması gerekmektedir. Bu tür seçme ve yeterlik belirleme amaçlı testlere çeşitli hatalar karışabilmektedir. Bu hataların bir kısmı kontrol altına alınmazken, bir kısmı ise kontrol altına alınabilmektedir. Sistemik hatalar kontrol altına alınabilen hata türleri arasındadır. Sistemik hatalar yanlılık olarak tanımlanabilmektedir. Yanlılığın kontrol altına alınması testin geçerlik ve güvenilirliğini doğrudan etkileyen bir faktör olduğundan, yanlılık çalışmaları bireylerin geleceklerini belirleyen sınavlar açısından hayati öneme sahiptir. Seçme ve yeterlik belirleme amacıyla kullanılan testler, aynı yetenek düzeyindeki bireyler için eşit cevaplama olasılığına sahip olmalıdır. Diğer bir anlatımla testlerde bulunan maddeler, bireylerin çeşitli özelliklerine göre yanlılık göstermemelidir.

Bireyleri seçme ve yerleştirme amacıyla kullanılan sınavlardan ilki ve en önemlilerinden biri olan SBS sınavlarının geçerli ve güvenilir olması, bireylerin geleceklerinin şekillenmesi ve yönlendirilmesi açısından hayati öneme sahiptir. Bu çalışmanın SBS ve benzeri sınavlara geçerlik ve güvenilirlik anlamında katkı sağlaması düşünülmektedir. Ayrıca SBS'nin farklı alt testlerinde farklı cinsiyet gruplarından gelen yanıtlayıcılara pratik bilgi aracılığıyla avantaj sağlayacak maddelerin tespit edilmesi ve gereken önlemlerin alınması açısından katkı sağlaması umulmaktadır.

Bu katkıların yanı sıra ülkemizde yanlılıkla ilgili birçok çalışma bulunmakla beraber Rasch modeline dayalı olarak madde yanlılığı tespit etme çalışmalarına rastlanılmaması anlamında ilgili literatüre katkı sağlaması düşünülmektedir.

1.6.SINIRLILIKLAR

- Araştırmada elde edilen veriler, 2012 Seviye Belirleme Sınavı verilerinden seçilen örneklem ile sınırlıdır.

- Arařtırmada elde edilen Deęiřen Madde Fonksiyonu (DMF) ieren maddeler, cinsiyete gore tespit edilen sonularla sınırlıdır.
- Arařtırmada elde edilen DMF'lerin yanlılık ierip iermedięi hakkındaki yorumlar, alıřılan 10 uzmanın grřu ile sınırlıdır.

İKİNCİ BÖLÜM KAYNAK ÖZETLERİ

2.1.MADDE VE TEST YANLILIĞI

Test yanlılığı, farklı alt gruplardan gelen aynı yetenek düzeyine sahip bireylerin testin özelliklerinden ya da testin temel amacına uygun olmayan koşullardan kaynaklanan herhangi bir grup lehine ya da aleyhine farklılaşmasıdır (Zumbo, 1999, Schumacker, 2005 ve Hambleton ve Rogers, 1996). Tanımdan da anlaşılacağı gibi belirli bir grup lehine ve belirli bir kural dahilinde farklılaşma söz konusu olduğundan, yanlılık sistematik hata olarak algılanmaktadır. Bu bağlamda test yanlılığı bir testin geçerlilik ve güvenilirliğini önemli düzeyde etkilemektedir (Kristjansson, Alyeswort, Mcdowell ve Zumbo, 2005).

Ölçme çalışmaları bireyler hakkında bir kanıya varıp belirli durumları yordayabilmek için kullanıldığından testin geçerli ve güvenilir olması büyük bir önem teşkil etmektedir. Test geçerliliğini ve güvenilirliğini artırmanın bir yolu madde yanlılığının belirlenmesidir. Aynı yetenek düzeyindeki bireylerin, belirli özelliklerinden dolayı bir maddeyi doğru cevaplama olasılıklarının farklılaşması olarak tanımlanan madde yanlılığı (Zumbo, 1999), özellikle seçme ve yerleştirme amacıyla yapılan sınavlarda bireylerin geleceğini etkileme açısından önemlidir (Bekçi, 2007).

2.2.YANLILIĞIN TARİHÇESİ

Test yanlılığı 1900'lü yıllarda yapılan etnik köken ve üst düzey dil becerisi gerektiren IQ testlerinde, Afro-Amerikanlar aleyhine sonuçlanan, etnik ve kültürel yanlılık olduğu tespit edilen çalışmalarla ön plana çıkmıştır (Cole ve Zieky, 2001). 1910 yılında Alfret Binet'in sosyoekonomik düzeyleri düşük bireyler üzerinde yaptığı çalışmada kültürel farklılığa bağlı olarak puan farklılaşması ile yanlılık kavramı ortaya çıkmıştır. Yapılan çalışmada bireylerin yeteneklerinden ya da zihinsel kapasitelerinden ziyade dil farklılığı, çevresel etmenler ya da aile

ortamlarının farklılığının yanlılığa neden olduğu tespit edilmiştir. Standart testlerdeki yanlılık çalışmaları yeni bir çalışma alanı olmayıp Binet ve Simon'nun 1905 yılında kendi zeka testlerinin ilk versiyonlarının tanıtımı ile gündeme gelmiştir. İşçi sınıfı çocukları ve yüksek gelirli ailelerin çocuklarına uygulanan bu çalışmanın sonucunda, gruplar arasında bariz bir farklılığın olduğu anlaşılmıştır. Zeka testlerinin daha sonra geliştirilen versiyonlarında, tespit edilen bu kültürel yanlılık, sosyal sınıf farklılıklarını ortaya çıkaran maddelerin elenmesiyle kontrol altına alınmıştır. Chicago'lu psikolog Kenneth Eells, 1951 yılında yaptığı çalışma ile yanlılıkla ilgili çağdaş araştırmamın ilk adımlarını atmıştır. Çalışmasında test maddelerine odaklanarak çeşitli kültürel gruplar arasındaki IQ farklılıklarını incelemiştir. Psikometri alanında saygın bir dergi olan “*Educational and Psychological Measurement*” da yayımlanan “An Investigation of Item Bias” çalışması madde yanlılık konusunda modern çalışmalardan bir diğeridir. Bu çalışmada madde yanlılığı Anova tekniği ile tespit edilmeye çalışılmıştır (Cleary ve Hilton, 1968).Yöntemsel hatalara rağmen, ANOVA stratejisi bugün madde yanlılığı çalışmalarında çokça uygulanmaya devam etmektedir. Madde yanlılığı konusundaki diğer önemli bir yöntem de Scheuneman tarafından ileri sürülmüştür. Scheuneman çalışmasında ki-kare yöntemini önermiştir. Ancak geliştirilen bu yöntem eksikliklerinden dolayı eleştirilmiştir. California Üniversitesinden Berkeley, yaptığı bir kaç seri çalışma ile madde yanlılığı tespitinde Scheuneman'ın ortaya atmış olduğu Ki-Kare yöntemini geliştirerek bu istatistiksel yönteme önemli derecede katkıda bulunmuştur (Osterlind, 1983). Kullanılan bu teknikler kendinden önceki çalışmaların eksikliklerini gidermek ve daha gelişmiş yöntemler ortaya koymak için yapılmıştır. Ancak daha sonraki dönemlerde yapılan çalışmalarda bireylerin zeka seviyelerini ölçen IQ testlerindeki farklılığın yanlılıktan kaynaklandığı gibi kalıtsal ve çevresel faktörlerden de etkilenebileceği düşünülmüştür. Bu bağlamda günümüzde de farklılığın kaynağını tespit etmek için yanlılık çalışmaları yapılmaktadır.

2.3.DEĞİŞEN MADDE FONKSİYONU (DMF)

Değişen madde fonksiyonu (DMF), bir testte cevaplayıcıların maddelere verdikleri cevaplar arasındaki değişimin istatistiksel olarak belirlenmesi şeklinde ifade edilebilir (Yurdugül, 2003; Gierl, Khaliq ve Boughton, 1999 ve Taylor ve Lee, 2012). Değişen madde fonksiyonu, iki grup (odak ve referans) arasında, bütün madde güçlük düzeyleri açısından bir grup lehine daha zor veya daha kolay olması

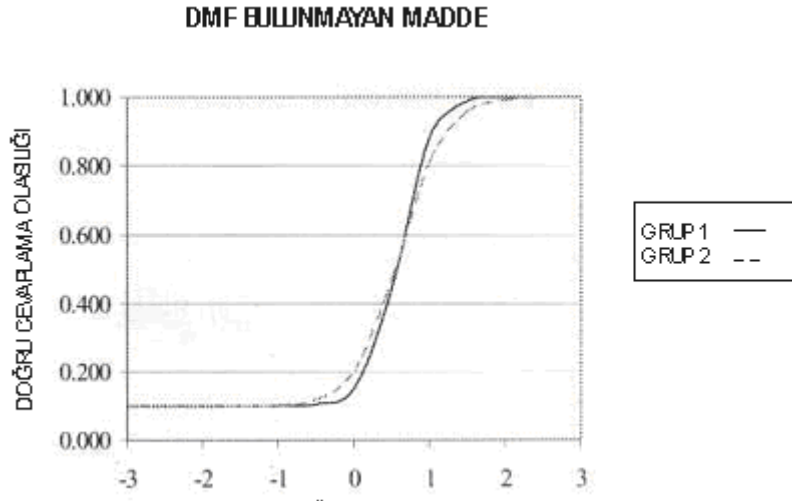
durumudur (Kristjansson, ve Diğerleri, 2005). Zumbo'ye (1999) göre gruplar arası bu farklılığın nedeni madde etkisinden (item impact) veya DMF'dan kaynaklanabilmektedir. Madde etkisi farklı başarı düzeyine sahip gruplar arasındaki gerçek farklılık olarak ifade edilirken, DMF aynı seviyedeki bireyler arasındaki farklılık olarak ifade edilir. Test yanlılığını belirlemenin en iyi yöntemi DMF belirlemektir. Örneğin başarı düzeyi düşük ve yüksek olan iki sınıfa uygulanan bir matematik testi sonuçlarının başarı düzeyi yüksek olan sınıf lehine çıkması olası bir sonuç olduğundan bu madde etkisi olarak gerçek değişim şeklinde nitelendirilebilir. Ancak başarı düzeyi aynı olan iki sınıfa uygulanan matematik testinin sonuçlarının bir grup lehine çıkması DMF olarak yorumlanabilmektedir. DMF Kurnaz'a (2006) göre farklı dillerden çeviri, maddelerin içeriğinden kaynaklanan farklılıklar, dil ve kültür yapı farklılığı, testin yapısından kaynaklı, kullanılan istatistiksel yöntemlerin farklılığından dolayı ortaya çıkabilmektedir. Geçerlik ve güvenilirliğin azalmasının bir kaynağı olarak düşünülen DMF'ye neden olan bu olumsuzlukların ortadan kaldırılabilmesi için yapılan çevirinin kültüre ve testin yapısına uygunluğu, dilinin anlaşılabilirliği gibi DMF kaynaklarının ortadan kaldırılması gerekir.

2.4.DMF İLE YANLILIK ARASINDAKİ FARK

DMF ölçülen özellik bakımından benzer olan ancak sosyo-ekonomik, etnik köken, cinsiyet gibi değişkenlerden kaynaklı istatistiksel sonuçların gruplardan herhangi birinin lehine ya da aleyhine çalışmasıdır (Hambleton, Swaminathan ve Rogers, 1991). Yani testin uygulandığı bireylerin ölçülen özellik bakımından benzer olduğu kabul edilmektedir. Tanımdan da anlaşılacağı gibi DMF bir istatistiksel işlem sürecidir. Öncelikle aynı yetenek düzeyine sahip alt gruplarda Değişen Madde Fonksiyonunu (DMF) içeren maddeler tespit edilir (Gök, Kelecioğlu ve Doğan, 2010; Roeber, 2005). Belirlenen DMF'lerin ölçme işleminden yani gerçek bir farktan (madde etkisinden) mı kaynaklandığını yoksa grubu oluşturan bireylerin yapısından mı kaynaklandığını tespit etmek için uzman görüşüne başvurulur. Sürecin bu aşamasında farklılık gruplardan kaynaklanıyor ise, tespit edilen bu DMF'ye yanlılık denmektedir (Camilli ve Shepard, 1994; Zumbo, 1999). Diğer bir ifade ile DMF, yanlılık çalışmasının istatistiksel bir basamağıdır.

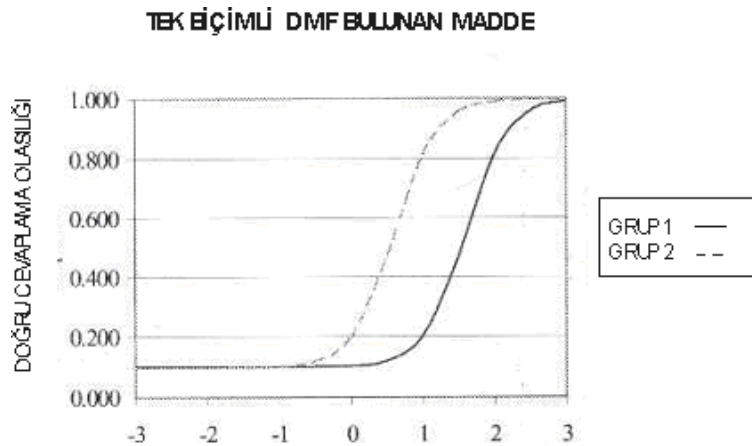
İlgili literatürde DMF, Tek Biçimli DMF ve Tek Biçimli Olmayan DMF olmak üzere ikiye ayrılmaktadır (Mellenbergh, 1982; Swaminathan ve Rogers, 1990).

Ölçülen özellik bakımından madde cevaplarında gruplar arası istatistiksel olarak bir farklılık yoksa bu madde DMF içermemektedir. Şekil-1 DMF içermeyen bir madde grafiğini temsil etmektedir.



Şekil-1: Değişen Madde Fonksiyonu İçermeyen Madde

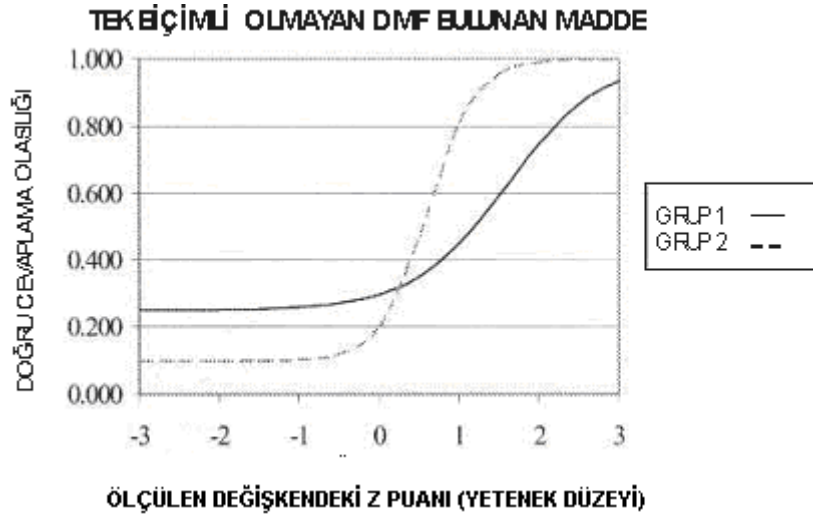
Beceri düzeyleri aynı olan iki grup arasında bir maddeyi doğru cevaplama olasılığı her düzeyde sabit bir şekilde farklılık gösteriyorsa bu maddenin tek biçimli DMF içerdiği söylenebilir. İki grup arasındaki farklılık sabit bir oranla değişmektedir. Şekil-2'de iki grubun her düzeyinde, bir maddeyi doğru cevaplama olasılığı sabit değişen, yani tek biçimli DMF içeren madde grafiği verilmiştir.



Şekil-2: Tek Biçimli Değişen Madde Fonksiyonu İçeren Madde

Şekil-2’de tek biçimli DMF içeren maddeye ilişkin karakteristik eğrisinde gruplar arasındaki farklılaşma paralel olarak değişmiştir.

Bireylerin test maddelerini cevaplama olasılığı düzeylere göre bir sabit değişim göstermiyorsa bu maddeler tek biçimli olmayan DMF içermektedir. Yani bireylerin düzeylerine göre cevaplama olasılığı arasında bir tutarlılık yoktur. Tek biçimli olmayan Değişen Madde Fonksiyonu (DMF) Şekil-3’te gösterilmiştir.



Şekil-3: Tek Biçimli Olmayan Değişen Madde Fonksiyonu İçeren Madde

Şekil-3’te gruplar arasındaki düzeylere göre bireylerin soruları doğru cevaplama olasılıkları arasında dalgalanma görülmektedir. Tek biçimli olmayan DMF içeren bir maddeye ilişkin farklı grupların madde karakteristik eğrisi kesişmektedir.

2.5.YANLILIK BELİRLEME YÖNTEMLERİ

Yanlılık belirlemenin istatistiksel basamağı olan DMF belirleme Klasik Test Kuramı ve Madde Tepki Kuramına göre tespit edilmektedir.

2.5.1.Klasik Test Kuramı ve DMF Belirleme Yöntemleri

Bireylerin davranışlarının ya da herhangi bir özelliğinin ölçülmesinde gerçek puanının kestirilmesi ölçmeye karışan çeşitli hatalardan dolayı pek mümkün değildir. Yani Klasik Test Kuramına (KTK) göre ölçme sonucu bireylerin gerçek puanlarına hata kavramının eklenmesi ile gözlenen puanlar kestirilmektedir (Baykul, 2010). Ölçmedeki yetenek düzeyi, madde ayırteçiciliği, madde güçlüğü, dağılımı,

güvenirlik gibi özellikleri gruba bağlı olarak değişmektedir. Bireylerin yetenekleri testten teste değişkenlik gösterirken testin istatistiksel özellikleri gruptan gruba değişmektedir (Özer, 2012). Deveci (2008) KTK'nın sınırlılıklarını; madde istatistiklerinin gruptan gruba değişmesi, test puanlarının gruptan gruba değiştiği için karşılaştırma yapılamaması, testi alan her yetenek grubunun puanın kestirilememesi ve güvenilirliğin bireyler bazında yapılmadan sadece grup bazında kestirilmesi olarak sıralamıştır.

Klasik test kuramında DMF belirleme yöntemleri arasında en çok ön plana çıkan iki yöntem: Lojistik Regresyon (LR) yöntemi ve Mantel-Haenszel (MH) yöntemidir.

Lojistik Regresyon

Hem tek biçimli hem tek biçimli olmayan DMF belirleme yönteminde kullanılan Lojistik Regresyon, bağımlı değişkenin iki değer aldığı bağımsız değişkenin ise sürekli bir değişken olduğu durumlarda kullanılır. LR tekniği ile DMF belirlenirken, madde etkisi değişkeni anlamlı olduğunda maddenin DMF içerdiğini bir göstergesidir. Aynı zamanda grup-madde etkileşim değişkeni anlamlı olduğunda da tek biçimli olmayan değişen madde fonksiyonu olduğu sonucuna varılır. Bu yöntemde elde edilen standartlaştırılmış regresyon katsayısı R^2 üç farklı düzeyde DMF derecesini verir. $R^2 < 0,035$ ise DMF yoktur ya da ihmal edilebilir düzeydedir, $0,035 \leq R^2 < 0,070$ ise orta düzeyde DMF vardır, $R^2 \geq 0,070$ ise önemli düzeyde DMF vardır. Bir maddenin DMF içeren (B veya C düzeyi) bir madde olarak sınıflandırılabilmesi için ki-kare değerinin 0,05' e eşit veya 0,05' den küçük olması ve R^2 değerinin de en az 0,035 olması gerekmektedir (Zumbo,1999).

Mantel-Haenszel (MH) Tekniği

Mantel-Haenszel (MH) tekniği ki-kare üzerine kurulu bir test istatistiğidir (Agresti, 1984). α (odds oranı) non-parametrik olarak odak ve referans gruptaki performans farklılığının derecesini ölçer. MH'nin kolay yorumlanabilmesi için logaritmik bir dönüştürme ile delta katsayısı elde edilir. MH tekniği ile DMF belirlenirken; $\Delta -MH=0$ ya da $\alpha =1$ ise madde de DMF olmadığı, $\Delta -MH<0$ ya da $\alpha >1$ ise madde de referans grubunun lehine DMF olduğu, $\Delta -MH > 0$ ya da $\alpha <1$ ise maddede odak grubunun lehine DMF olduğu şeklinde yorumlanır (Nandakumar,

1993). Bununla birlikte MH analizi sonucunda ki-kare deęerinin anlamlı çıkması da maddenin DMF içerdiğini ifade eder. DMF düzeyini belirlerken; $|\Delta - MH| < 1$ ise maddede ihmal edilebilir düzeyde (A düzeyi); $1 \leq |\Delta - MH| < 1,5$ ise maddede orta düzeyde (B düzeyi); $|\Delta - MH| \geq 1,5$ ise maddede önemli düzeyde (C düzeyi) DMF olduđu söylenir (Zieky, 1993).

2.5.2. Madde Tepki Kuramı Test Kuramı ve DMF Belirleme Yöntemleri

Madde Tepki Kuramı (MTK), bir ölçme işleminde madde parametrelerinin gruptan bağımsız, grup yetenek düzeyleri ya da grubun istatistiksel işlemlerinin testten bağımsız olması olarak nitelendirilebilir. MTK'da bireylerin maddeleri cevaplama davranışı, yetenek denilen faktör ile kestirilmeye çalışılmaktadır. Yani bireyin gözlenemeyen davranışlarını gözlemlerle matematiksel olarak ilişkilendirilmesi olarak tanımlanabilir. MTK kapsamında bireyin bir cevaplama davranışı ile o maddeye cevap vermesini sağlayan ilişki, madde karakteristik eğrisi olarak nitelendirilmektedir. KTK'nın aksine MTK'da madde özellikleri bireyin yeteneğinden bağımsız; bireyin yetenek düzeyi de ölçme aracından bağımsızdır. Yani madde ve yetenek parametreleri değişmezdir (Güler, 2008).

Madde Tepki Kuramı kapsamında birçok DMF belirleme yöntemi bulunmaktadır. Bunlardan bazıları alan ölçümleri ve olabirlik, ki-kare ve Rasch DMF analizi yöntemleridir.

Olabilirlik oran ölçümü

Raju tarafından öne sürülen olabilirlik oran ölçümünde, odak ve referans gruplarının madde karakteristik eğrileri arasında kalan alana bakılarak, maddeler arasında istatistiksel olarak anlamlı farklılıkların olup olmadığı incelenir. İncelenen madde karakteristik eğrileri bize maddenin tek biçimli yada tek biçimli olmayan DMF içermesi veya DMF içermemesi hakkında fikir verir (Camilli ve Shepard, 1994).

Ki-kare yöntemi

Cromwell'e göre ki-kare yönteminde odak grup ile referans grup parametre değerleri arasındaki kovaryans değeri hesaplandıktan sonra odak grubun parametreleri referans grup parametre değerlerine dönüştürülür. Yapılan bu

istatistiksel işlemden sonra dönüştürülmüş parametre ve kovaryans değerlerinden yola çıkılarak ki-kare yöntemi ile DMF tespit edilir.

Rasch modeli ile DMF analizi

Rasch modeli MTK kapsamında tek parametrelidir. Bu modelin temel sayıtları: ölçülen yeteneğin grupta normal dağılım göstermesi, testi oluşturan maddelerin tek boyutlu olması, aynı yetenek düzeyindeki maddelerin cevaplanma olasılıklarının birbirinden bağımsız olması, şans faktörü sıfır olmasıdır (Rasch, 1961).

Rasch modelinde, her aday bir yeteneğe (β), her madde de bir güçlüğü (δ) sahiptir. Eğer bir test DMF içeriyorsa referans grubun madde güçlüğü odak grubun madde güçlüğünden farklı olacaktır. Rasch analizi, eşit aralıklı bir ölçekte her bir grupta, her katılımcı için yetenek kestirimi yapar. Her bir şüpheli madde performansını incelemek için Rasch tahminleri referans grubun her bir üyesi için:

$$\beta - \delta_R = \ln(P_{R1} / P_{R0}),$$

her bir odak grup üyesi için:

$$\beta - \delta_F = \ln(P_{F1} / P_{F0})$$

parametrelerle elde edilebilir.

Bu parametrelerdeki :

P_{R1} : referans grubunda maddeyi doğru bulan deneklerin oranı

P_{R0} : referans grubunda maddeyi yanlış bulan deneklerin oranı

P_{F1} : odak grupta maddeyi doğru bulan deneklerin oranı

P_{F0} : odak grupta maddeyi yanlış bulan deneklerin oranı

Rasch modeli aynı testte eşit ham puanlarına sahip bütün katılımcıları, aynı yetenek düzeyine yerleştirdiği için, her iki grupta eşit puanlayan katılımcıların şüpheli madde performanslarını karşılaştırmak mümkündür. Bu;

$$\delta_F - \delta_R = \ln(P_{R1} / P_{R0}) - \ln(P_{F1} / P_{F0}) = \ln(a_k)$$

Rasch modeli için iki grubun şüpheli maddelerini verir. Yanlılığın anlamlılığını belirlemek için:

$$t = (\delta_F - \delta_R) / SE(\delta_F - \delta_R),$$

testi yapılır. Bu testteki SE, farklılığın standart hatasını belirtmektedir (Taylor ve Lee, 2012) .

2.6. SEVİYE BELİRLEME SINAVI (SBS)

Değişen madde fonksiyonu ve yanlılık belirleme ile ilgili yapılan tüm istatistiksel işlemlerin temel amacı kullanılan testlerin geçerlilik ve güvenilirlik seviyelerini arttırmaktır. Bu bağlamda madde yanlılığı çalışmaları, bireylerin geleceği açısından büyük önem arz eden seçme ve yerleştirme gerektiren sınavlar için kaçınılmaz olarak görünmektedir. Türkiye’de bireyleri seçme ve yerleştirmeye yönelik birçok sınav yapılmaktadır. Bu sınavların en önemlileri ortaöğretimden sonra bireylerin üniversiteye girmeleri için yapılan sınavlar ve ilköğretimden sonra bireylerin ortaöğretime yerleşmeleri için yapılan sınavlardır.

Ortaöğretimden sonra üniversiteye yerleşmek için yapılan sınavlar Yükseköğretime Giriş Sınavı (YGS) ve Lisans Yerleştirme Sınavı (LYS) olmak üzere iki kademedен oluşmaktadır. Bu sınavlar bireylerin Türkçeyi kullanma gücü, matematiksel işlemlerden yararlanma becerisi, sosyal bilimlerle ilgili temel kavramlar ve ilkelerle ilgili düşünebilmesi, fen bilimlerle ilgili temel kavramlar ve ilkelerle ilgili düşünebilme ve dil becerilerinin niteliğini ölçmeyi amaçlamaktadır.

Ortaöğretime öğrenci seçme ve yerleştirme amacıyla Milli Eğitim Bakanlığı 2007 yılına kadar 8. sınıf öğrencilerinin kazanım düzeylerini belirlemek için Ortaöğretim Kurumlar Sınavı (OKS) uygulamıştır. 2008 yılından sonra seçilmiş ortaöğretim kurumlarına (Anadolu Lisesi, Fen Lisesi, Anadolu Öğretmen Lisesi vb.) öğrenci alımı için 6., 7. ve 8. sınıflara Seviye Belirleme Sınavı (SBS) uygulanmaya başlanmıştır. SBS 2010 yılından sonra kademeli olarak 6. ve 7. sınıflarda uygulanmadan kaldırılmıştır (MEB, 2011).

Seviye Belirleme sınavı öğrencinin sahip olması gereken kazanımlar doğrultusunda yorumlama gücü, analiz etme, eleştirel düşünme, problem çözme, mantıksal bağlantı kurma gibi becerileri kazandırmak amacıyla temel dersler olan Türkçe, Matematik, Fen ve Teknoloji, Sosyal Bilgiler (Din Kültürü ve Ahlak Bilgisi) ve Yabancı Dil derslerinden oluşmaktadır. Sınava giren bireyler, sınavda gösterdikleri başarı ve okuldaki bütün derslerden alınan başarı puanları (Yılsonu Başarı Puanı) ile birlikte hesaplanarak bir sonuç elde edilmektedir. Bu sonuçlar doğrultusunda bireyler ortaöğretim kurumlarına yerleştirilmektedir.

Seviye belirleme sınavı, müfredata uygun olarak Türkçe, Matematik, Fen ve Teknoloji, Sosyal Bilgiler ve Yabancı Dil olmak üzere 5 dersten oluşmaktadır. Bu derslerin soru dağılımı Tablo-1’de sunulmuştur.

Tablo-1: SBS-2012 alt testleri soru sayısı ve soru ağırlık katsayısı

Dersler	Soru Sayısı	Soru Ağırlık Katsayısı
Türkçe	23	4
Matematik	20	4
Fen ve Teknoloji	20	3
Sosyal Bilgiler	20	3
Yabancı Dil (İngilizce, Almanca, Fransızca, İtalyanca)	17	1

Tablo-1’de görüldüğü gibi en yüksek soru ağırlık katsayısına Türkçe ve Matematik alt testleri sahipken, en düşük soru ağırlık katsayısına ise Yabancı Dil testi sahiptir (MEB, 2012).

Tablo-1’de belirtilen katsayı ve ders dağılımına göre, belirlenen puanlar doğrultusunda her yıl yüz binlerce aday sınava girip belirli bir ortaöğretime yerleştirilmektedir. Yapılan sınavlarda akademik başarının tespit edilmesi amaçlandığından dolayı sınava karışan hatanın minimize edilmesi büyük önem taşımaktadır. Diğer bir ifade ile bu tür sınavların geçerlilik ve güvenilirlik çalışmalarının yapılması gerekmektedir. Bu bağlamda sınavın herhangi bir yanlılık içerip içermediğini incelemek sınavın geçerlilik ve güvenilirliğinin yükseltilmesi için önemli bir unsurdur.

2.7.YURTIÇİNDE YAPILAN ÇALIŞMALAR

Kurnaz (2006) yapmış olduğu çalışmasında, Peabody Resim Kelime Testinde yer alan maddelerin sosyo-ekonomik düzey ve cinsiyet değişkenlerine göre değişen madde fonksiyonunu araştırmak için 592 katılımcının verdiği yanıtları kullanmıştır. DMF içeren maddelerin tespiti için Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) tekniklerinden faydalanılıp DMF içeren maddeler çeldirici cevap analizi ile incelenmiştir. Araştırmada elde edilen bulgulara göre Mantel-Haenszel tekniği ile 16 maddenin cinsiyet değişkenine göre 33 maddenin ise sosyo-ekonomik düzey değişkenine göre DMF içerdiği tespit edilmiştir. Bunun yanı sıra Lojistik Regresyon tekniğine göre 14 maddenin cinsiyet, 33 maddenin sosyo-ekonomik düzeye göre DMF içerdiği tespit edilmiştir. Elde edilen DMF maddelerinin, testin güvenilirliğini düşürdüğü ortaya çıkmıştır. Elde edilen bulgulara göre Mantel-Haenszel tekniğinin çeldirici yanlılığına daha duyarlı olduğu sonucuna varılmıştır.

Bekçi (2007), ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavında yer alan maddelerin cinsiyete ve okul türüne göre değişen madde fonksiyonu (DMF) içerip içermediğini belirlemek için 777137 kişinin katıldığı sınavda Türkçe, Matematik, Fen Bilgisi ve Sosyal Bilgiler alt testlerini incelemiştir. Araştırmacı Sosyal Bilgiler alt testinin DMF içermediği, okul türüne göre Türkçe alt testinin iki sorusunda, Fen Bilgisi alt testinin bir sorusunda ve Matematik alt testinin ise dört sorusunda testin DMF'li madde içerdiği tespit etmiştir. Araştırmacı DMF'li maddeleri Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) yöntemleriyle yaptığı istatistiksel işlemler sonucunda tespit etmiştir. Bu veriler ışığında OKS testlerinde yer alan maddelerin uzman görüşleri doğrultusunda, yanlılık içermediğini gözlemlemiştir.

Doğan ve Öğretmen (2008) DMF'nu tespit etmede kullanılan Ki-kare, Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) tekniklerini karşılaştırarak uygulamada ortaya çıkan benzerlik ve farklılıkları belirlemek amacıyla, 2003 yılında yapılan Ortaöğretim Kurumları Seçme ve Yerleştirme (OKÖSYS) sınavına katılan 600000 öğrencinin sınav sonuçları veri olarak kullanılmıştır. Çalışmanın örneklemini yansız olarak seçilen 3345 öğrenci oluşturmaktadır. Çalışmada sorulan sorulardan Fen Bilgisi alt testinin cinsiyete göre değişen madde fonksiyonu içerip içermediğine bakılmıştır. Araştırma sonuçlarına göre en fazla DMF'li madde sayısı MH tekniği ile elde edilmiştir. İşaretsiz Ki-kare tekniği DMF'li madde sayısı tespit etmede ikinci

sıradadır. İşaretli Ki-kare ve LR tekniği ile hiç bir madde için DMF tespit edilmemiştir. Sonuç olarak, kullanılan bu üç teknik Ki-kare değerlerinin büyüklüğü bakımından benzerlik gösterirken, DMF içeren madde sayısı bakımından farklı sonuçlar verdiği tespit edilmiştir.

Gök, Kelecioğlu ve Doğan (2010) Değişen Madde Fonksiyonu (DMF) belirlemede kullanılan Mantel-Haenszel ve Lojistik Regresyon tekniklerini karşılaştırmak için 2005 yılında yapılan Ortaöğretim Kurumları Sınavı (OKS) Matematik ve Fen Bilgisi alt testlerini okul türü ve cinsiyet değişkenine göre farklılaşıp farklılaşmadığını incelemişlerdir. Araştırma kapsamında sınava katılan 777137 kişi SPSS paket programıyla tesadüfi olarak belirlenen katılımcıların %2'si örnekleme oluşturmuştur. Yapılan istatistiksel işlemler sonucunda elde edilen bulgular; Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) tekniklerinin karşılaştırılmasıyla, Matematik ve Fen bilgisi alt testlerinin okul türü ve cinsiyet değişkenlerine göre DMF içeren madde sayısında genel olarak düşük düzeyde bir uyum görülmüştür.

Kalaycıoğlu ve Kelecioğlu (2011) yapmış oldukları çalışmada 2005 yılı Öğrenci Seçme Sınavında (ÖSS) yer alan maddelerin DMF içerip içermediğini, içeriyorsa bu maddelerin yanlı olup olmadıklarını tespit etmişlerdir. Testler için DMF belirlemede Mantel-Haenszel (MH) ve Lojistik Regresyon (LR) teknikleri kullanılmıştır. DMF içeren maddelerin yanlılıklarını tespit etmek için uzman görüşüne başvurulmuştur. Çalışmanın örneklemini 2005 yılında ÖSS'ye giren 599330 lise son sınıf öğrencisi oluşturmuştur. Katılımcıların, her biri 45 maddeden oluşan Türkçe, Sosyal Bilimler, Matematik ve Fen Bilimleri alt testlerine verdikleri cevaplar veri olarak kullanılmıştır. Araştırmanın sonucunda, cinsiyete göre Türkçe alt testini oluşturan maddelerin DMF içermediği sonucuna varılmıştır. Sosyal Bilimler alt testinde yedi, matematik ve Fen Bilimleri alt testlerinde ise üç maddenin DMF içerdiği ve bu maddelerden Fen Bilimleri testinde yer alan maddelerden bir tanesinin yanlılık gösterdiği belirlenmiştir.

Yurtiçinde yapılan çalışmalar incelendiğinde, yapılan çalışmaların bir çoğunda DMF belirleme yöntemi olarak Mantel-Haenszel ve Lojistik Regresyon yöntemleri kullanılmıştır. Bu çalışmalarda her iki yöntem karşılaştırılması ve DMF tespit etmede bir birlerine göre üstünlükleri araştırılmıştır. Yapılan bazı çalışmalarda

ise ülke genelinde yapılan seçme ve yerleştirme sınavlarının alt testlerindeki maddelerin yanlılıkları incelenmiştir. Bu testlerden bazılarında yanlı madde tespit edilirken bazılarında tespit edilmemiştir. Yurtiçinde yapılan çalışmalarda Rasch modeline dayalı olarak DMF tespit edilen çalışmalara pek rastlanılmamıştır.

2.8.YURTDIŞINDA YAPILAN ÇALIŞMALAR

Hanna (1986) Matematik başarısında cinsiyete göre farklılığı incelemek için 8. sınıf düzeyinde Ontario, Kanada, okullarında bir araştırma yapmıştır. Araştırmanın örneklemini 2511 okul arasından seçilen 130 okul oluşturmuştur. Araştırmaya 1773 erkek ve 1750 bayan olmak üzere toplam 3523 kişi katılmıştır. Araştırmada veri toplama aracı olarak International Association for the Evaluation of Educational Achievement'nın (IEA) Second International Mathematics Study (SIMS) testi kullanılmıştır. Veri toplama aracı olarak kullanılan test, 5 seçenekli çoktan seçmeli test olup toplam beş konu alanından oluşmaktadır. Bu konu alanları 58 maddelik *Aritmetik*, 31 maddelik *Cebir*, 42 maddelik *Geometri*, 17 maddelik *Olasılık ve İstatistik* ve 26 maddelik *Ölçme* testlerinden oluşmuştur. Araştırma sonucunda, yüksek olmamakla beraber sadece *Geometri ve Ölçme* testlerinde erkekler lehine bir fark ortaya çıkmıştır. Aradaki bu fark çok büyük olmamakla beraber İstatistiksel olarak 0.01 düzeyinde anlamlı bulunmuştur.

Linacre ve Wright'ın 1987 yılında yapmış oldukları çalışmada, Değişen Madde Fonksiyonu içeren madde tespitinde kullanılan Mantel-Haenszel (MH) yöntemi ve Rasch modelini karşılaştırmayı amaçlamışlardır. Mantel-Haenszel istatistikleri özetlenip tahmin ettikleri parametreler belirlenmiştir. MH yöntemi, yanlılığı yani DMF içeren maddeleri tespit etme ve nicelleştirme yapmak için kullanılır. Çalışmada her iki yaklaşımın kuramsal özellikleri karşılaştırılmış ve aynı varsayımları gerektirdikleri gösterilmiştir. Sonuç olarak MH yönteminin istatistiksel olarak Rasch modeline kıyasla daha alt basamakta olduğu anlaşılmıştır. Mantel-Haenszel yöntemi, Rasch modelinin uygulandığı yöntemin dolaylı bir şekilde uygulandığı bir yöntemdir.

Zwick ve Erikcan 1989 yılında yapmış oldukları çalışmalarında, *National Assessment of Educational Progress (NAEP)* ölçeğinin Tarih alt testinde bulunan DMF maddelerini Mantel-Haenszel yöntemi ile tespit etmişlerdir. DMF analizi 11. sınıfa devam eden 7743 kişilik veri gurubu ile yapılmıştır. Analiz sonuçlarına

bakıldığında bazı maddelerin, siyahi, Hispanik (Latin kökenli ispanyollar) ve kadınlar açısından zor yanıtlanabilecek türden oldukları anlaşılmıştır. Yani kimi maddelerin bu grup aleyhine DMF içerdiği tespit edilmiştir. Çalışmada *National Assessment of Educational Progress (NAEP)* ölçeğinin Tarih alt testinde etnik köken, renk ve cinsiyete göre yanlı maddelerin olduğu tespit edilmiştir.

Gamer ve Engelhard (1999), ulusal ortaöğretim bitirme sınavlarında uygulanan, çoktan seçmeli ve yapılandırılmış matematik testi maddelerinde, erkek ve kadın performansları arasındaki yanlılığı incelemeyi amaçlamışlardır. Araştırmacılar bu farklılığı Rasch modelini temel alarak ortaya çıkarmaya çalışmıştır. Araştırma verileri, toplam 3952 tane 11.sınıf öğrencisine *1994 Georgia High School Graduation Testi* uygulanarak toplanmıştır. Araştırma sonuçları, ortalamalar ve DMF indekslerinde, kadınların *Cebirle* ilgili çoktan seçmeli maddelerde istatistiksel olarak anlamlı ve tutarlı bir avantaja sahip olduğunu göstermektedir. *Geometri ve Ölçme, Sayılar ve Hesaplama, Veri Analizi, Orantısal Muhakeme* konularında erkeklerin daha az tutarlı bir avantaja sahip oldukları gözlemlenmiştir. Sekiz maddenin ikisinde erkeklerin ortalamasının kadınların ortalamasından daha yüksek olduğu tespit edilmiştir. Ancak erkekler ve kadınlar yeteneğe göre kıyaslandığında, performanslarında anlamlı farklılık olduğu ve bu farklılığın kadınlar lehine olduğu saptanmıştır.

Le (1999) çalışmasında, 1988 yılında yapılan the National Education Longitudinal Study (NELS: 88) başarı testinin tarih alt testine ilişkin cinsiyet açısından DMF içeren maddelerin tespit edilmesini amaçlamıştır. Araştırmanın örneklemini, farklı okullara devam eden 432 tane 11.sınıf öğrencisi oluşturmuştur. Tarih testine ilişkin başarı testi toplam 30 sorudan oluşmaktadır. Bu maddelerden 17'sinin *erkeksi* maddeleri, 12'sinin ise *kadınsı* maddeleri yansıttığı düşünülmüştür. Bir maddenin ise nötr olarak belirlenmiştir. Araştırma sonuçları incelendiğinde, Tarih testine ilişkin başarı testinde, maddelerin bir kısmının erkekler lehine çalıştığı ama bu avantajın sınırlı sayıda maddede olduğu anlaşılmıştır.

Zenisky, Hambleton ve Robin'in (2004) yapmış oldukları çalışma, büyük ölçekli fen testlerinin cinsiyete göre DMF içeren madde tespiti, maddelerin DMF içerme eğilimleri ve DMF içeren maddelerin hangi açıdan(içerik mi, Bilişsel olarak mı, Madde tipinden kaynaklı mı, madde içeriğinden mi veya görseellikten kaynaklı

mı) DMF içerdiğini incelemeyi amaçlamıştır. Çalışma grubu ilköğretim, ortaöğretim ve lise kademelerine devam eden ve her seviyeden ikişer grup olmak üzere, toplam 6 gruptan oluşmuştur. Her gruptan 60000 kişi olmak üzere toplam 360000 kişilik bir veri seti oluşturulmuştur. Her grup için verilen testler 32 ile 42 madde arası olmak üzere çoktan seçmeli ve açık uçlu maddelerden oluşmuştur. Çalışma sonuçlarına göre genel olarak çoktan seçmeli test maddelerinin erkekler lehine meyilli olduğu, açık uçlu maddelerinin ise kadınlar lehine meyilli olduğu anlaşılmıştır. Sonuçlar incelendiğinde erkek ve kadınlar arasında en dengeli dağıldığı görülen test *hayat bilgisi* testidir. *Hayat Bilgisi* testini; *teknoloji ve yeryüzü* testi ve *uzay bilimleri* testleri takip etmiştir. Bir başka bulgu ise görsel öge içeren çoktan seçmeli maddeler erkeklere avantaj sağlarken; kadın ve erkek grubunda şekil çizilmesi istendiği açık uçlu sorularda bir farklılığa rastlanmamıştır.

Arıffın, Idrıs ve Ishak (2010)'ın MyGSI (*Malaysian Generic Skills Instrument*) ölçeğinin cinsiyet ve etnik kökenlerine göre, DMF içerip içermediğini tespit etmek için yapmış oldukları çalışmaları, *Kebangsaan Malaysia* Üniversitesinde 12 bölümden 1262 kişinin katılımıyla yapılmıştır. MyGSI ölçeği toplam 102 madde olmak üzere 13 alt boyuttan oluşmaktadır. Ölçekte DMF içeren maddeler, Rasch modeli kullanılarak tespit edilmiştir. Verilerin analizi için Winsteps yazılımı 3.64.2 versiyonu kullanılmıştır. Cinsiyet açısından DMF incelendiğinde, bir maddenin erkekler lehine çalıştığı gözlemlenmiştir. Etnik köken açısından ise farklı kökenler arasında farklı maddeler bakımından DMF olduğu tespit edilmiştir.

Taylor ve Lee'nin (2012) yapmış oldukları çalışma, 4. 7. ve 10.sınıf okuma ve matematik testlerinden oluşmaktadır. Bu testler ulusal ölçüt dayanıklı testlerden maddeler alınarak oluşturulmuştur. Testler çoktan seçmeli ve yapılandırılmış maddelerden oluşmaktadır. DMF içeren maddelerin cinsiyet açısından incelenmesi için Polysibtest ve Rasch modelleri kullanılmıştır. Çoktan seçmeli maddelerde, Rasch modeline göre DMF içeren madde sayısı aynı anda uygulanmış madde yanlılık yöntemine göre daha fazla orandadır.

Taylor ve Lee çalışmalarında, okuma ve matematik çoktan seçmeli testleri, genel olarak erkekler lehine çalışmışken yapılandırılmış maddeler ise kadınlar lehine çalışmıştır. İçerik analizleri sonucunda, DMF içeren maddelerin tipik olarak metin yorumlamaları veya örtük anlamlı maddelerde tespit edildiği görülmektedir.

Testlerde mantıksal yorum yapma ve bilgi içeren maddeleri analiz etmeyle ilgili maddelerin erkeklere avantaj sağladığı gözlemlenmektedir. Kadınlar lehine çalışan çoğu maddenin ise, öğrencilerin kendi yorumlarını ve analizlerini yapmak istedikleri maddelerden oluştuğu tespit edilmiştir. Matematik maddelerine ilişkin DMF analiz sonuçlarına bakıldığında geometri, cebir ve olasılıkla ilgili konuları erkekler lehine çalıştığı istatistiki yorumlama, çok adımlı problem çözme ve matematiksel muhakeme ile ilgili konuların ise kadınlar lehine çalıştığı gözlemlenmiştir.

Yurtdışında yapılan çalışmalar incelendiğinde, yapılan çalışmaların bir kısmı DMF belirleme yöntemlerinin kıyaslanması, bir kısmı cinsiyet veya etnik kökenlere göre yanlılığın incelenmesi gibi konuları incelemektedir. Çalışmalarda geometrik şekil, mantıksal ilişki kurma gibi beceri gerektiren maddelerin erkekler lehine sözel ve cebirsel ifade gibi beceri gerektiren maddelerin ise bayanlar lehine sonuçlandığı tespit edilmiştir.

ÜÇÜNCÜ BÖLÜM MATERYAL VE YÖNTEM

Bu bölümde araştırmanın modeli, evren ve örnekleme, veriler, verilerin toplanması ve verilerin analizi ile ilgili bilgiler yer almaktadır.

3.1.ARAŞTIRMANIN MODELİ

Araştırmada; sekizinci sınıflara 2012 yılında uygulanan SBS Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilimler alt testlerinin yanlılık içerip içermediğinin belirlenmesi amaçlanmıştır. Bu araştırmada, var olan bir durum herhangi bir değişiklik yapılmadan ortaya konulması amaçlanmıştır. Var olan bir durumun var olduğu biçimiyle, herhangi bir değiştirme, etkileme olmaksızın betimlenmesi amaçlandığından bu araştırma, tarama modelinde betimsel bir araştırmadır (Karasar, 2008).

3.2.EVREN VE ÖRNEKLEM

Araştırmanın evrenini, 2012 yılında SBS'ye giren sekizinci sınıf öğrencilerinin maddelere verdikleri yanıtların tümü oluşturmaktadır. 2012 yılında SBS'ye giren öğrenci sayısının, SBS testlerini aldıkları kitapçık türüne ve cinsiyetlerine göre dağılımı Tablo-2'te verilmiştir.

Tablo-2: Kitapçık ve Cinsiyet Türüne Göre Öğrenci Evren Dağılımı:

	A Kitapçığı	B Kitapçığı	Toplam
Kadın	266907	261071	527978
Erkek	271289	276279	547568
Toplam	538196	537350	1075546

Tablo-2 incelendiğinde sınava katılan birey sayısının 1075546 olduğu ve bu katılımcıların cinsiyet ve kitapçık türüne göre birbirlerine yakın bir dağılım gösterdikleri anlaşılmaktadır.

Verilerin analizi için evrenin tamamına ilişkin veriler bulunmaktadır. Ancak analizde yaşanabilecek güçlüklerin önlenmesi için, her kitapçık ve cinsiyetin yaklaşık %20'si, tabakalı örneklem yönteminin orantılı seçim yöntemi ile örneklem oluşturulmuştur. Oluşturulan bu örneklem basit seçkisiz örnekleme yöntemi kullanılarak meydana getirilmiştir. Basit seçkisiz örnekleme yönteminde, evrendeki tüm elemanlar, birbirine göre eşit ve bağımsız seçilme şansına sahiptir (Karasar, 2008).

Tabakalı örnekleme yöntemi sosyal bilimlerdeki pek çok araştırma için, özellikle homojen olmayan evrenlerde, uygun bir örnekleme yoludur (Baykul, 1997). Bu çalışmada tabakalı örnekleme yönteminin kullanılması tercih edilmiştir. Tabakalı örnekleme yönteminde orantısız ve orantılı olmak üzere iki türlü seçim yolu vardır. Orantısız seçimde tabakalardaki birey sayısı göz önüne alınmadan her tabakadan eşit sayıda birey örnekleme alınır. Orantılı seçimde ise örnekleme alınacak bireyler tabakalardaki birey sayısına orantılı olarak seçilir. İşlemleri kolaylaştırdığı için bu çalışmada, tabakalı örnekleme yönteminin orantılı seçimi kullanılmıştır (Sümbüloğlu ve Sümbüloğlu, 1993:252).

Araştırmanın örneklemini oluşturan istatistiksel bilgiler kitapçık ve cinsiyet türüne göre Tablo-3'te verilmiştir.

Tablo-3: Kitapçık ve Cinsiyet Türüne Göre Öğrenci Örneklem Dağılımı:

	A Kitapçığı	B Kitapçığı	Toplam
Kadın	53785	52620	106405
Erkek	56776	53182	109958
Toplam	110561	105802	216363

Tablo-3'te çalışmanın örneklemini oluşturan bireylerin 216363 olduğu, kitapçık türü ve cinsiyete göre benzer dağılım gösterdikleri anlaşılmaktadır.

3.3.VERİLERİN ELDE EDİLMESİ

Araştırmada, sekizinci sınıflara 2012 yılında uygulanan SBS Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilimler alt testlerinin maddeleri dikkate alınmıştır. Bu alt testlerden Türkçe testi 23, Matematik, Fen ve Teknoloji ve Sosyal Bilimler testi ise 20'şer maddeden oluşmaktadır. Araştırma verilerini 2012 yılında SBS'ye giren ve bu alt testlere verilen yanıtlar oluşturmaktadır. Araştırmada ele

alınan problemin çözümü için kullanılan veriler, MEB Eğitim Teknolojileri Genel Müdürlüğü Ölçme ve Değerlendirme Daire Başkanlığı'ndan yazılı izinle sağlanmıştır.

3.4.VERİLERİN ÇÖZÜMLENMESİ

Araştırmada ele alınan soruların cevaplandırılabilmesi için, MEB Eğitim Teknolojileri Genel Müdürlüğü Ölçme ve Değerlendirme Daire Başkanlığı'ndan sağlanan, öğrencilerin Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilimler alt testlerine verdikleri yanıtlar öncelikle alt testlere göre ayrıştırılıp “Microsoft Office Excel” programına aktarılarak her alt testin A ve B grupları ayrıştırılmıştır. Gruplara göre ayrıştırılan yanıtlar gruplar bazında cinsiyetlere göre ayrıştırılmıştır. Excel programı ile veriler doğru cevaplar doğrultusunda “1-0” formatına çevrilmiştir. Her alt testin grup ve cinsiyetlere göre ayrıştırılmasından sonra veriler SPSS-20 paket programına aktarılıp her alt testin gruplar ve cinsiyet bazında tabakalı örnekleme orantılı seçim, seçkisiz yöntemle yaklaşık %20 si rastgele seçilmiştir. Yapılan örneklem oluşturma işleminden sonra uygun işlemlerle cinsiyetler ve gruplar birleştirilerek veriler Winsteps 3.8.1 programı yardımıyla bu programa uygun olacak şekilde Metin Belgesi Dosyasına aktarılmıştır. Aktarılan dosyada uygun syntaxlar yazılıp verilerin Winsteps paket programında analizi için hazır hale getirilmiştir.

Verilerin Rasch modeli sayıtlılarını karşıladığını tespit etmek için SPSS ve Amos paket programları kullanılmıştır. Verilerin normal dağılım gösterme durumlarını incelemek için SPSS paket programından faydalanılırken, tek boyutluluğuna ilişkin Doğrulayıcı Faktör Analizi (DFA) Amos paket programı yardımıyla yapılmıştır.

Rasch modellerinde DMF analiz tespiti için ConQuest (Wu, Adams, Wilson ve Haldane, 2007), Facets (Linacre, 2009) ve Winsteps (Linacre, 2010) paket programları kullanılmaktadır (Karami, 2011).

Winsteps paket programı Rasch modelinde DMF analizi için en çok kullanılan programlardan biridir (Wang ve Chen, 2005). KTK'ya göre yapı geçerliğini daha iyi belirlemek için nokta çift serili korelasyonların 0.3'ten büyük olması gerektiği ifade edilirken, Linacre'ye(2007) göre Winsteps programında nokta çift serili (veya nokta ölçüm) korelasyonlar pozitif olmalıdırlar. Her madde, yapıya önemli bir katkıda bulunmalıdır (Bold ve Fox, 2007). Uygun madde, artan uyum

istatistiğin ortalamalar karesi (MNSQ) ile ölçülür (Bond ve Fox, 2001). Uyum istatistiğın beklenen değeri 1.0 dır. Bu değeri 0'dan sonsuza kadar bir sayı alabilir. Sapmalar, maddeler ve model arasında uyum eksikliğıne işaret etmektedir. Beklenenden düşük değeri madde artıklığı veya binişikliğı olarak yorumlanabilir (Ariffin ve ark., 2010). Ayrıca Winsteps programı için uyum istatistiklerinden olan iç uyum istatistikleri, maddelere karşı cevapları etkileyen, beklenmeyen davranışları sezinlemeye duyarlıyken; dış uyum istatistikleri ise insanların maddeler üzerindeki beklenmeyen davranışlarına duyarlıdır (Linacre, 2002). Madde uyum istatistikleri MNSQ (ortalamalar karesi) beklenen değeri 1.0 dır. MNSQ >2.0 durumunda ölçüm sistemini bozmaktadır. Değeri $1.5 < MNSQ \leq 2.0$ olması ölçüm yapısı için verimsizken değeri $0.5 < MNSQ \leq 1.5$ aralıkta olması ölçüm için verimlidir. MNSQ değeri 0.5'tan küçük olası ölçüm için daha az verimlidir (Linacre, 2002; Smith, 2000). Bond ve Fox (2007) iç uyum ve dış uyum ölçek maddeleri için madde ortalamalar karesinin 0.6 ile 1.4 arası değışebileceğini söylemektedir. Ayrıca nokta çift serili (PTMEA) korelasyon değeri 0.20 değerinden büyük olması maddenin test yapısıyla uygun olduğunu göstermektedir (Streiner ve Norman, 1995).

Winsteps paket programında DMF tespiti için DMF Kontrast değerine bakılmaktadır. Lai ve Eton'a (2002) göre DMF Kontrast değeri 0,5 logit değeri olması likert tipi ölçekler için kritik değerdir. Ayrıca Pallant ve Tennant'a (2007) göre, cinsiyet bakımından DMF değeri 0,5 logit değeri altında olması maddenin DMF içermediğini göstermektedir. Negatif DMF Kontrast değeri maddenin odak gruba kolay geldiğini ve maddenin odak gruba avantaj sağladığını göstermektedir.

Değışen Madde Fonksiyonu (DMF) Winsteps paket programı aracılığıyla tespit edildikten sonra, yanlılık durumunun belirlenmesi için uzman görüşüne başvurulmuştur. Uzmanlar, ilgili alt test uzmanları, Psikolojik Danışma ve Rehberlik uzmanları ve Ölçme Değerlendirme uzmanından oluşmuştur. Toplam 10 uzman olmak üzere 2 Doç.Dr, 3 Yrd. Doç. Dr. ve 5 Arş. Gör.'den oluşmaktadır.

DÖRDÜNCÜ BÖLÜM BULGULAR VE TARTIŞMA

4.1.BULGULAR

Rasch modelinde madde ve yetenek parametrelerinin kestirilmesinde; ölçülen yeteneğin normal dağılım göstermesi, tek boyutlu, test maddelerinin yerel bağımsız olması ve maddelerin “a” (madde ayırıcılık) parametresinin 1 olması gibi sayıtlıları karşılaması gerekmektedir (Gelbal, 1994). Yapılan analizlerin Rasch modeline uygunluğu için *Türkçe, Matematik, Sosyal Bilimler ve Fen ve Teknoloji* alt testlerinin; normal dağılım, tek boyutluluk, test maddelerinin yerel bağımsızlık özellikleri incelenmiştir. Curry, Bashow ve Rentz (1978)’e göre Rasch modeli sayıtlılarından biri olan testteki maddelerin ayırıcılık değerlerinin bir değerinde olması gerekliliği koşulunu sağlayan maddelerin kolay elde edilememesinden dolayı bu sayıtlıyı sağlamayan maddelerle de ölçümün yapılabileceğini belirtmiştir (akt. Kaptan,1994). Aşağıda alt testlere ilişkin normallik, tek boyutluluk ve yerel bağımsızlık varsayımlarının karşılanıp karşılanmadığına yönelik çalışmalar yapılmıştır.

4.1.2.Normallik

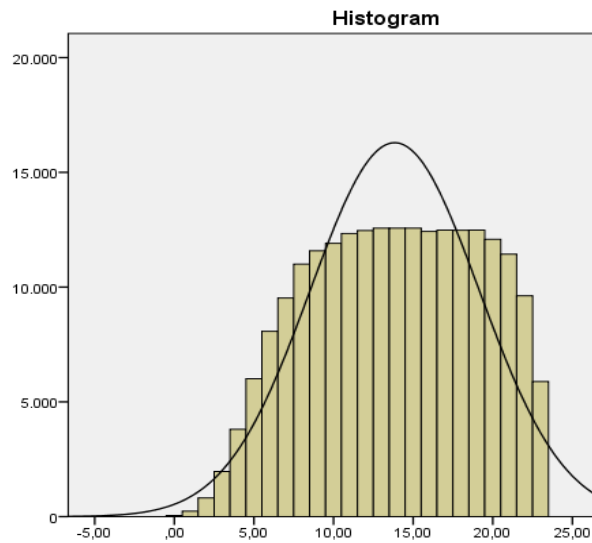
Rasch modeline ilişkin önemli sayıtlılardan biri verilerin normal bir dağılım göstermesidir. Bu bağlamda *Türkçe, Matematik, Sosyal Bilgiler ve Fen ve Teknoloji* alt testlerinin normal dağılım gösterip göstermediği incelenmek için normallik grafiği, çarpıklık katsayısı ve basıklık katsayılarına bakılmıştır..

Aşağıda *Türkçe, Matematik, Sosyal Bilgiler ve Fen ve Teknoloji* alt testlerinin betimsel istatistikleri verilmiştir.

Tablo 4: Testlere İlişkin Betimsel İstatistikler

Test	N	\bar{X}	Mod	Med.	S.S	Var.	Basıklık	Çarpıklık
Türkçe	216354	13.84	14	14	5.30	28.07	-0.99	-0.12
Matematik	216354	7.16	5	6	4.75	22.53	0.09	0.94
Fen ve Teknoloji	216354	9.16	6	8	4.73	22.41	-0.71	0.47
Sosyal Bilimler	216354	11.95	16	12	4.56	20.76	-0.88	-0.26

Şekil-4'te 2012 Seviye Belirleme Sınavına ilişkin 23 maddelik Türkçe alt Testine ilişkin verilerin dağılımı verilmiştir.



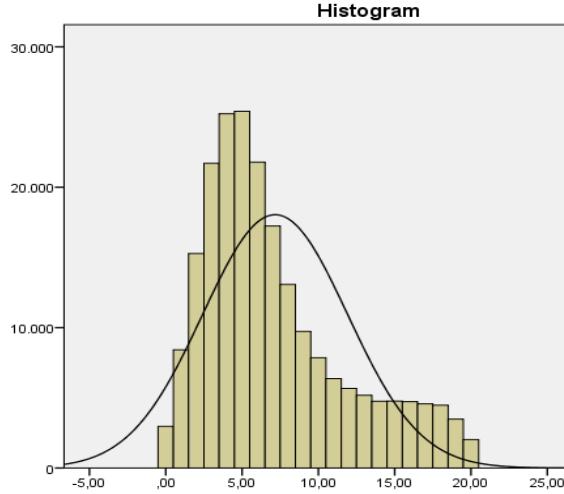
Şekil-4: Türkçe Alt Testine İlişkin Puan Dağılımı

23 maddelik Türkçe alt testine ilişkin veriler aşağıdaki gibidir:

Tablo-4 ve Şekil-4 incelendiğinde 23 soruluk Türkçe testi puan ortalaması 13.84 mod ve medyan değerleri 14, basıklık değeri -0.99 ve çarpıklık değeri -0.12 olarak tespit edilmiştir. Basıklık ve Çarpıklık değerlerinin sıfıra yakın olması beklenmektedir. Çarpıklık katsayısı -1 ve +1 arası alınan değerler, normal dağılım olarak kabul edilebilir (Büyüköztürk, 2002). Türkçe testine ilişkin değerler incelendiğinde testin normal dağılım gösterdiği varsayılabilir.

Bir diğer test olan matematik alt testine ilişkin puan dağılımı aşağıda verilmiştir.

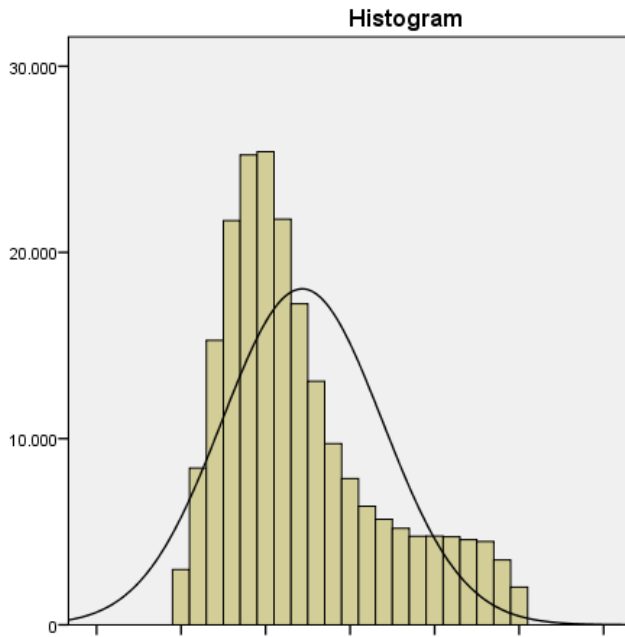
Şekil-5'te Matematik alt testine ilişkin puan dağılımı verilmiştir.



Şekil-5: Matematik Testi Puan Dağılımı

Tablo-4 ve Şekil-5 incelendiğinde 20 maddelik Matematik testi puan ortalaması 7.16, mod değeri 5, medyan değerleri 6, basıklık değeri 0.09 ve çarpıklık değeri 0.94 olarak tespit edilmiştir. Matematik testine ilişkin değerler incelendiğinde testin normal dağılım gösterdiği varsayılabilir.

SBS-2012 alt testlerinden bir diğeri olan Fen ve Teknoloji alt testine ilişkin puan dağılımları Şekil-6'da sunulmuştur.

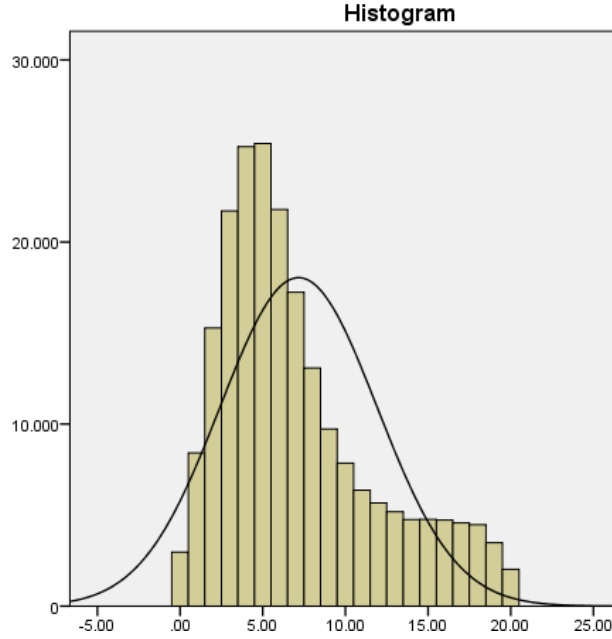


Şekil-6: Fen ve Teknoloji alt Testi Puan Dağılımı

20 maddelik Fen Bilgisi Testine ilişkin veriler aşağıdaki gibidir:

Tablo-4 ve Şekil 6 incelendiğinde 20 soruluk Fen Bilgisi testi puan ortalaması 9.16, mod değeri 6, medyan değerleri 8, basıklık değeri -0.71 ve çarpıklık değeri 0.47 olarak tespit edilmiştir. Fen Bilgisi testine ilişkin değerler incelendiğinde testin normal dağılım gösterdiği varsayılabilir.

Bir diğer test olan Sosyal Bilgiler alt testine ilişkin puan dağılımı Şekil-7'de verilmiştir.



Şekil-7: Sosyal Bilgiler Testi Puan Dağılımı

20 maddelik Sosyal Bilimler Testine ilişkin veriler aşağıdaki gibidir:

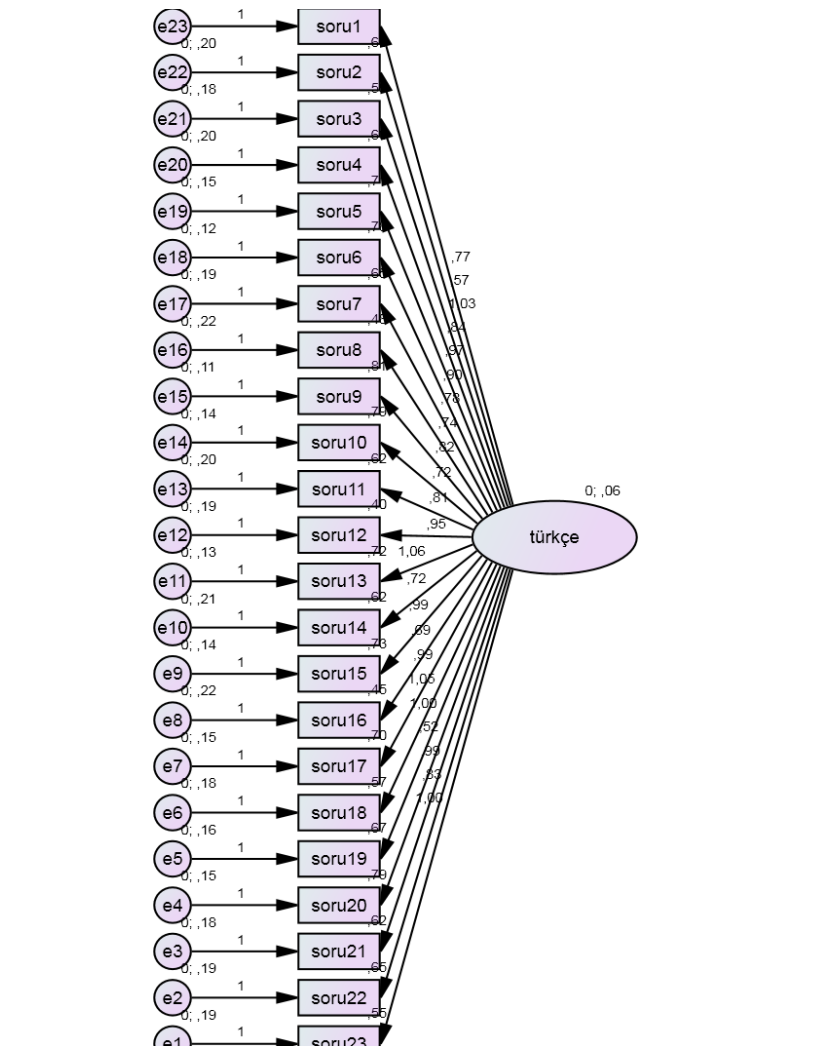
Tablo-4 ve Şekil-7 incelendiğinde 20 soruluk Sosyal Bilgiler testi puan ortalaması 11.95, mod değeri 16, medyan değerleri 12, basıklık değeri -0.88 ve çarpıklık değeri -0.26 olarak tespit edilmiştir. Sosyal Bilgiler testine ilişkin değerler incelendiğinde testin normal dağılım gösterdiği varsayılabilir.

4.1.2. Tek Boyutluluk

Rasch modelinin bir diğer önemli sayıtlısı ise ölçülen özelliğin tek boyutlu olmasıdır. Testlerin tek boyutluluğunu ölçmek için faktör analizi yapılmıştır. 2012 SBS Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilgiler testine ilişkin Doğrulayıcı faktör analiz sonuçları aşağıda verilmiştir.

DFA’da uyumlu olup olmadığı sınanan modelin yeterliğini ortaya koymak üzere pek çok uyum indeksi kullanılmaktadır. Bu çalışmada yapılan DFA için RMSEA (Root Mean Square Error of Approximation), CFI (Comparative Fit Index), NFI (Normed Fit Index), RFI (Relative Fit Index), IFI (Incremental Fit Index) ve TLI (Tucher-Lewis Index) uyum indeksleri incelenmiştir. CFI, NFI, RFI, IFI ve TLI indeksleri için kabul edilebilir uyum değeri 0.90 ve mükemmel uyum değeri 0.95 olarak kabul edilmektedir (Bentler ve Bonett, 1980; Bentler, 1980; Marsh, Hau, Artelt, Baumert ve Peschar, 2006). RMSEA için ise 0.08 kabul edilebilir uyum ve 0.05 mükemmel uyum değeri olarak kabul edilmiştir (Byrne ve Campbell, 1999; Browne ve Cudeck, 1993).

Şekil-8’de Türkçe alt testinin tek boyutluluk analizine ilişkin DFA sonuçları aşağıda sunulmuştur.

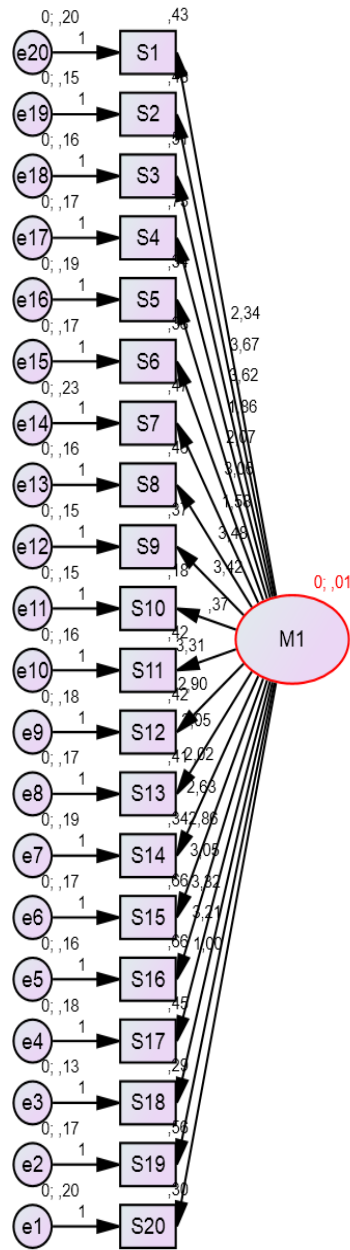


Şekil-8:Türkçe Alt Testinin Tek boyutluluğuna İlişkin Doğrulayıcı Faktör Analiz Sonuçları

Türkçe alt testine ilişkin RMSEA, CFI, IFI, NFI, RFI ve TLI değerlerine bakılarak maddelerin tek boyut altında gösterdikleri uyum incelenmiştir. Şekil-8'de Türkçe alt testine ilişkin RMSEA değeri 0.021, CFI değeri 0.97, IFI 0.97, NFI değeri 0.97, RFI değeri 0.96 ve TLI 0.96 olarak tespit edilmiştir. Doğrulayıcı faktör analizi verileri incelendiğinde, Türkçe alt testine ilişkin maddelerin bütün değerlerde mükemmel uyum gösterdikleri için tek boyuta uygun veriler olduğu anlaşılmaktadır.

Bir diğer alt test olan Matematik alt testine ilişkin yapılan doğrulayıcı faktör analizinde RMSEA, CFI, IFI, NFI,RFI ve TLI değerlerine bakılarak maddelerin tek boyut altında gösterdikleri uyum incelenmiştir.

Şekil-9’de Matematik alt testinin tek boyutluluk analizine ilişkin DFA sonuçları aşağıda sunulmuştur.



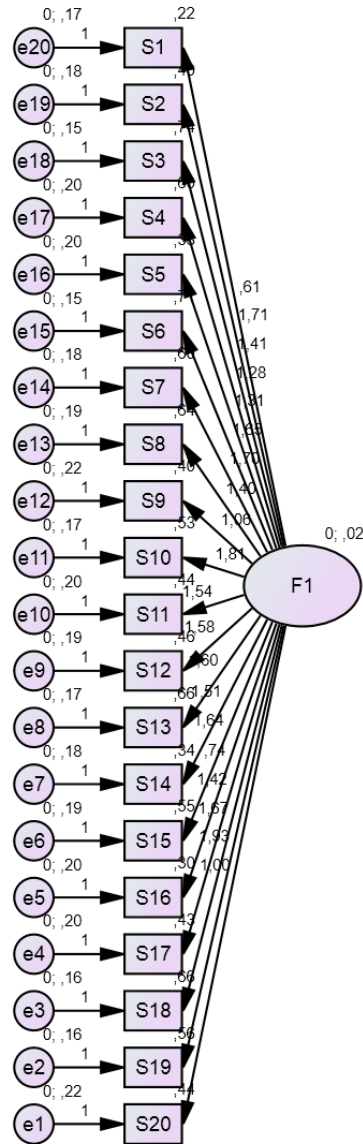
Şekil-9:Matematik Alt Testinin Tek boyutluluğuna İlişkin Doğrulayıcı Faktör Analiz Sonuçları

Şekil-9’da Matematik alt testine ilişkin RMSEA değeri 0.029, CFI değeri 0.95, IFI 0.95, NFI değeri 0.95, RFI değeri 0.94 ve TLI 0.94 olarak tespit edilmiştir. DFA verileri incelendiğinde Matematik testine ilişkin sonuçlardan RMSEA, IFI, NFI ve CFI değerlerinin mükemmel uyum, RFI ve TLI değerlerinin kabul edilebilir uyum

gösterdikleri anlaşılmaktadır. Yani Matematik testine ilişkin maddelerin tek boyuta uygun veriler olduğu anlaşılmaktadır.

2012-SBS alt testlerinden Fen ve Teknoloji alt testine ilişkin yapılan doğrulayıcı faktör analizinde RMSEA, CFI, IFI, NFI, RFI ve TLI değerlerine bakılarak maddelerin tek boyut altında gösterdikleri uyum incelenmiştir

Şekil-10'da Fen ve Teknoloji alt testinin tek boyutluluk analizine ilişkin DFA sonuçları aşağıda sunulmuştur.



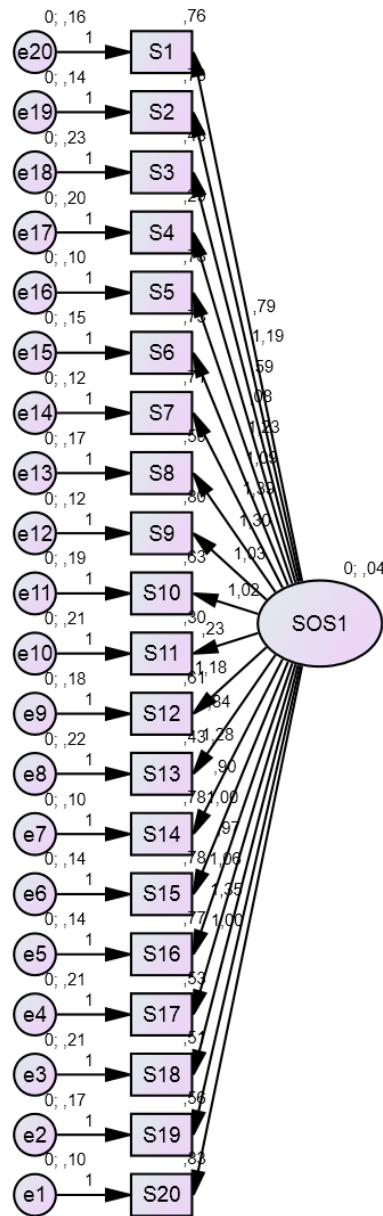
Şekil-10: Fen ve Teknoloji Alt Testinin Tek boyutluluğuna İlişkin Doğrulayıcı Faktör Analiz Sonuçları

Şekil-10'da Fen ve Teknoloji alt testine ilişkin RMSEA değeri 0.030, CFI değeri 0.94, IFI 0.94, NFI değeri 0.94, RFI değeri 0.93 ve TLI 0.93 olarak tespit

edilmiştir. DFA verileri incelendiğinde Matematik testine ilişkin sonuçlardan RMSEA değerlerinin mükemmel uyum, CFI, IFI, NFI ve TLI değerlerinin ise kabul edilebilir uyum gösterdikleri anlaşılmaktadır. Yani Fen ve Teknoloji testine ilişkin maddelerin tek boyuta uygun veriler olduğu anlaşılmaktadır.

Bir diğer test olan Sosyal Bilgiler alt testine ilişkin yapılan doğrulayıcı faktör analizinde RMSEA, CFI, IFI, NFI, RFI ve TLI değerlerine bakılarak maddelerin tek boyut altında gösterdikleri uyum incelenmiştir

Şekil-11’de Sosyal Bilgiler alt testinin tek boyutluluk analizine ilişkin DFA sonuçları aşağıda sunulmuştur.



Şekil-11: Sosyal Bilgiler Alt Testinin Tek boyutluluğuna İlişkin Doğrulayıcı Faktör Analiz Sonuçları

Şekil-11’de Sosyal Bilgiler testine ilişkin RMSEA değeri 0.026, CFI değeri 0.96, IFI 0.96, NFI değeri 0.96, RFI değeri 0.95 ve TLI 0.93 olarak tespit edilmiştir. DFA verileri incelendiğinde Sosyal Bilgiler testine ilişkin sonuçlardan RMSEA, CFI IFI, NFI ve RFI değerlerinin mükemmel uyum, TLI değerinin ise kabul edilebilir uyum gösterdikleri anlaşılmaktadır. Yani Sosyal Bilgiler testine ilişkin maddelerin tek boyuta uygun veriler olduğu anlaşılmaktadır.

SBS-2012 Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilimlere ilişkin veriler incelendiğinde her bir testin kendi içerisinde tek bir boyutlu olduğu sonucuna varılmıştır. Veriler Rasch modelinin tek boyutluluk sayıltısını karşılamaktadır.

4.1.3.Yerel Bağımsızlık

Yerel bağımsızlık, cevaplayıcının farklı maddelere verdiği cevapların istatistiksel olarak birbirinden bağımsız olması olarak tanımlanabilmektedir. Bunun bir sonucu olarak bireyin bir maddedeki performansı, diğer maddedeki performansını etkilememelidir. Yani yetenek düzeyi benzer bireylerin, benzer sorulara verdikleri cevapların benzer olması beklenir. Yerel bağımsızlık; madde setindeki ilişkilerin tek yetenek ile açıklanması olduğundan, tek boyutluluk varsayımının karşılanması, yerel bağımsızlık varsayımının da sağlandığı şekilde yorumlanabilir (Lord, 1980; Hambleton ve Swaminathan, 1985).

Sonuç olarak SBS-2012 Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilimlere ilişkin veriler incelendiğinde her bir testin kendi içerisinde tek bir boyutu olduğu tespit edildiğinden veriler yerel bağımsızlığı sağlamaktadır. Verilerin normallik, tek boyutluluk ve yerel bağımsızlık sayıtlılarını karşılamaları, verilerin Rasch modeli analizlerine uygun olduğunu göstermektedir.

Tablo-5'te SBS-2012 *Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilgiler* alt testlerine ilişkin birey ve madde güvenilirlik değerleri verilmiştir.

Tablo-5: Test Türlerinin Birey ve Madde Güvenirlik Değerleri:

	Madde (N)	Bireyler		Maddeler	
		Güvenirlik	Ayrışıklık	Güvenirlik	Ayrışıklık
Türkçe	23	0.78	1.84	0.86	120.3
Matematik	20	0.72	1.60	0.84	118.6
Fen ve Teknoloji	20	0.76	1.78	0.84	141.4
Sosyal Bilgiler	20	0.75	1.74	0.84	169.0

Tablo-5 incelendiğinde Türkçe alt testine ilişkin bireylerin güvenilirlik değeri 0.78 olarak saptanırken madde güvenilirlik değeri 0.86 olarak saptanmıştır. Matematik testine ilişkin birey güvenilirlik değeri 0.72 iken madde güvenilirlik değeri 0.84 olarak belirlenmiştir. Fen ve Teknoloji testi incelendiğinde, birey güvenirligi 0.76 değerini alırken madde güvenirligi 0.84 değerindedir. Sosyal Bilgiler testi incelendiğinde ise birey güvenilirlik değeri 0.75 olarak saptanırken madde güvenirligi 0.84 olarak saptanmıştır.

4.1.4. Türkçe Alt Testine İlişkin Bulgular

Tablo-6'da SBS-2012 Türkçe alt testine ilişkin madde güçlük düzeyi, madde iç ve dış uyum indeksleri ve nokta çift serili madde test korelasyonu verilmiştir.

Tablo-6: Türkçe Alt Testine İlişkin Madde Güçlük Düzeyi, Uyum İndeksi ve Nokta Çift Serili Korelasyon Değerleri:

Ham Puan	Toplam Puan	Madde güçlük Düzeyi	Model S.E	İç Uyum		Dış Uyum		Nokta Çift Serili Kor.(PTMEA)	Maddeler
				MNSQ	ZSTD	MNSQ	ZSTD		
80519	199137	1,32	.1	.98	-.99	1.04	9,0	.53	12
92639	212117	1,12	.1	1.10	.99	1.18	9,9	.45	1
89819	198855	1,07	.1	1.16	.99	1.26	9,9	.42	16
97378	203529	.90	.1	1.11	.99	1.19	9,9	.44	8
106553	193344	.53	.1	.97	-.99	0.95	-9,9	.52	23
108027	185769	.37	.1	.95	-.99	0.90	-9,9	.54	18
121959	203787	.27	.1	.95	-.99	0.93	-9,9	.52	3
120684	198783	.22	.1	1.05	.99	1.06	9,9	.46	4
121109	194550	.14	.1	1.07	.99	1.06	9,9	.45	11
129289	208307	.14	.1	1.11	.99	1.21	9,9	.40	14
126511	201826	.11	.1	.96	-.99	0.92	-9,9	.50	21
136912	210451	-.05	.1	1.06	-.99	1.06	9,9	.43	7
126405	193939	-.06	.1	1.03	-.99	1.03	7,7	.45	22
139380	209884	-.14	.1	1.16	-.99	1.23	9,9	.36	2

Tablo-6 Devam

Ham Puan	Toplam Puan	Madde güçlük Düzeyi	Model S.E	İç Uyum		Dış Uyum		Nokta Çift Serili Kor.(PTMEA)	Maddeler
				MNSQ	ZSTD	MNSQ	ZSTD		
129765	193327	-.16	.1	.93	-.99	.84	-9,9	.51	19
146578	208855	-.34	.1	.91	-.99	.85	-9,9	.50	17
150562	210764	-.42	.1	.91	-.99	.85	-9,9	.49	5
147424	201671	-.51	.1	.84	-.99	.73	-9,9	.53	13
152206	207734	-.56	.1	.88	-.99	.80	-9,9	.50	15
166943	212839	-.91	.1	.87	-.99	.81	-9,9	.48	6
162498	205138	-.92	.1	.95	-.99	1.11	9,9	.41	10
168503	212581	-.97	.1	1.06	.99	1.24	9,9	.34	20
172454	211182	-1.15	.1	.87	-.99	.69	-9,9	.46	9

Tablo-6’da 216363 kişinin SBS Türkçe alt testi kapsamında 23 maddenin güçlük ve uyumlarına ilişkin istatistiksel bilgileri verilmiştir. Tablo-6’da maddelerin güçlük düzeyi ve madde test uyumu incelenmiştir. Maddeler güçlük düzeyi yüksek olandan düşük olana doğru sıralanmıştır. Tablo-6’da; maddelere doğru cevap veren sayısı, maddeyi yanıtlayan toplam kişi sayısı, madde güçlük katsayısı, modelin standart hatası, iç uyum ortalamalar karesi, iç uyum ZSTD katsayısı, dış uyum ortalamalar karesi, dış uyum ZSTD katsayısı, nokta çift serili korelasyon katsayısı ve madde numaraları verilmiştir. Tablo-6’da sorulara cevap veren kişi sayısına bakıldığında 212839 kişi ile 6. madde en yüksek iken en düşük sayı 185769 kişi ile 18.madde olduğu tespit edilmiştir.

Tablo-6’da bireylerin maddeleri doğru cevaplamalarına ilişkin verilere bakıldığında en yüksek doğru cevaplanma oranı 172454 kişi ile 9.maddeyken, en düşük doğru cevaplanma oranı 80519 kişi ile 2. madde olarak gözlemlenmektedir. 23 maddelik Türkçe testine ilişkin veriler incelendiğinde madde güçlük düzeyi en yüksek olan 1,32 ile 12.madde iken güçlük düzeyi en düşük madde -1,15 ile 9.madde olarak saptanmıştır. Modelin standart hatası tüm maddeler için 0.1 olarak hesaplanmıştır.

Maddelere ilişkin test uyumu; iç uyum ve dış uyum ortalama kareler değerleri incelenmiştir. Madde fit istatistiklerinin MNSQ (ortalamalar karesi) beklenen değerleri 1.0 dır. MNSQ değerinin 2.0’den büyük ölçülmesi test sistemini bozmaktadır. Değerin $1.5 < MNSQ \leq 2.0$ arasında değer alması ölçüm yapısı için verimsizdir. Değer $0.5 < MNSQ \leq 1.5$ arasında ise ölçüm için verimlidir. MNSQ değerinin 0.5’ten küçük olması ölçüm için daha az verimlidir (Linacre, 2002; Smith,

2000). Tablo-6'daki Türkçe alt testine ilişkin 23 maddenin uyum indeksleri incelendiğinde iç uyum indekslerinin 0.84 ile 1.16 arasında deęiřtięi gözlemlenmektedir. 13.madde 0.84 iken, 16.ve 2.maddeler 1.16 deęerini almıřtır. Mevcut maddelerin dıř uyum ortalamalar karesi uyum indeksi incelendiğinde maddelerin 0.69 ile 1.24 arası deęer aldıęı gözlemlenmektedir. 9.madde 0.69 deęeri ile en düşük deęeri alırken 20.madde 1.24 ortalamalar karesi deęeri ile en yüksek deęeri almıřtır.

Tablo-6 incelendięinde, maddelerin nokta çift serili (PTMEA) korelasyon deęerleri 0.54 ile 0.34 deęerleri arasında deęiřmektedir. En yüksek deęeri 0.54 ile 18.madde alırken, en düşük deęeri 0.34 ile 20.maddenin aldıęı gözlemlenmektedir.

Soru-1 ve Soru-16'nın aynı güçlük düzeyinin sahip olduğu tespit edilmiştir. Soru-8'in bu soruların güçlük düzeyine göre daha az güçlük düzeyinde olduğu gözlemlenmiştir. Buna karşın madde güçlük düzeyi en düşük olan madde Soru-9'dur. Bu soruyu takiben Soru-20'nin en düşük 2. madde güçlük düzeyinde olduğu tespit edilmiştir. Soru-10 ve Soru-6 benzer güçlük düzeyinde bu soruları takip etmektedir. Diğer maddelerin orta güçlük düzeyinde olduğu tespit edilmiştir. Maddelerin güçlük düzeyi yükseldikçe maddenin zorluk derecesi artmaktadır.

2012-SBS Türkçe alt testine ilişkin Değişen Madde Fonksiyonu (DMF) kontrast değeri Tablo-7'de sunulmuştur.

Tablo-7: Türkçe Alt Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değerleri:

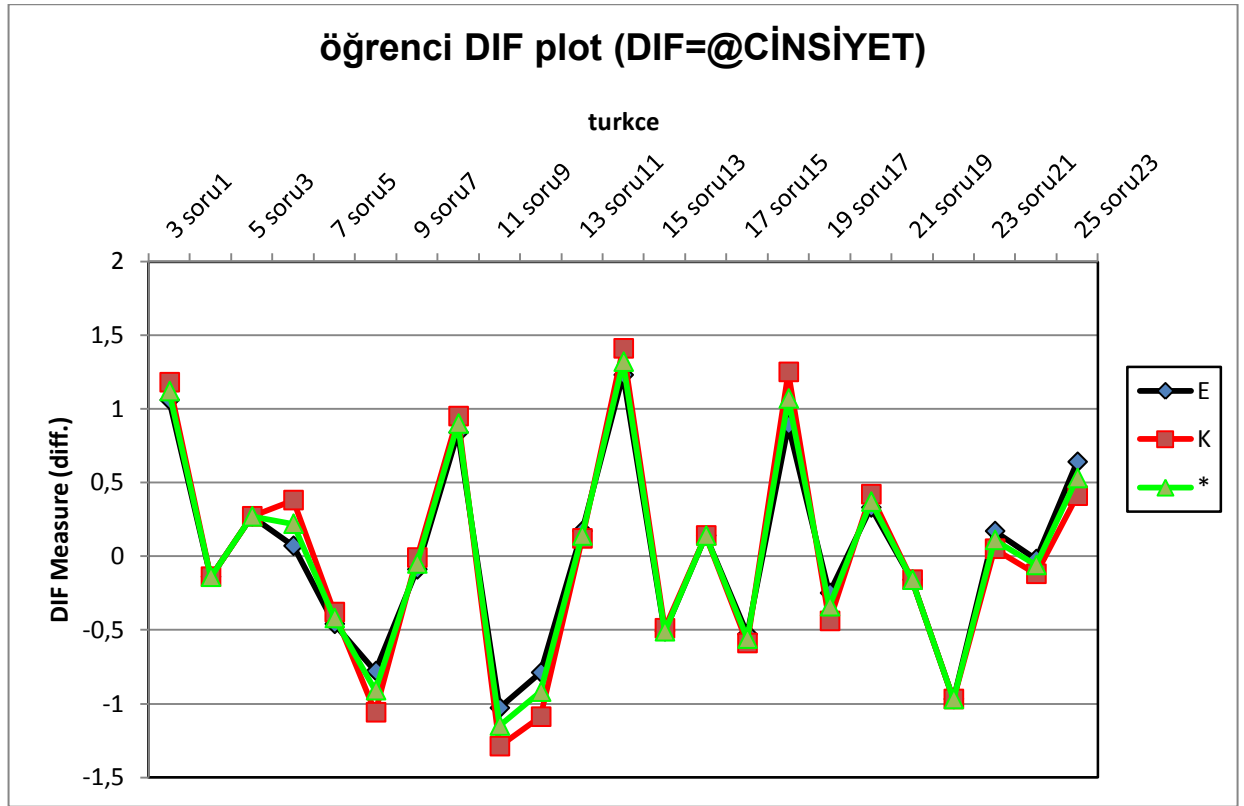
Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	T	Madde
Erkek	1.06	Kız	1.18	-.12	-11.4	1
Erkek	-.14	Kız	-.14	.00	.00	2
Erkek	.27	Kız	.27	.00	.00	3
Erkek	.07	Kız	.38	-.31	-29.1	4
Erkek	-.46	Kız	-.38	-.08	-7.78	5
Erkek	-.78	Kız	-1.06	.28	23.77	6
Erkek	-.09	Kız	-.01	-.07	-7.07	7
Erkek	.84	Kız	.95	-.11	-10.2	8
Erkek	-1.03	Kız	-1.29	.26	20.90	9
Erkek	-.79	Kız	-1.09	.31	25.06	10
Erkek	.17	Kız	.12	.05	4.87	11
Erkek	1.23	Kız	1.41	-.18	-16.5	12
Erkek	-.51	Kız	-.49	-.02	-1.90	13
Erkek	.14	Kız	.14	.00	.00	14
Erkek	-.53	Kız	-.59	.07	6.05	15
Erkek	.89	Kız	1.25	-.35	-33.2	16
Erkek	-.25	Kız	-.44	.18	16.73	17
Erkek	.33	Kız	.42	-.09	-8.44	18
Erkek	-.16	Kız	-.16	.00	.00	19
Erkek	-.97	Kız	-.97	.00	.00	20
Erkek	.17	Kız	.05	.12	11.24	21
Erkek	-.02	Kız	-.12	.10	8.86	22
Erkek	.64	Kız	.41	.23	21.58	23

Tablo-7'de katılımcıların cinsiyetlerine göre Türkçe alt testinin 23 maddesine vermiş oldukları cevaplara ilişkin DMF ölçüm değerleri verilmiştir. Tablo-7'de erkeklerin maddelere göre DMF ölçüm değerleri, kızların maddelere göre DMF değerleri, toplam DMF ölçüm değeri, t değerleri ve madde numaraları verilmiştir.

Tablo-7'ye bakıldığında erkek öğrencilerin Türkçe alt testi maddeleri DMF ölçüm değerleri -1.03 ile 1.23 logit değeri arasında değiştiği görülmektedir. Türkçe testinin erkek yanıtlayıcılarının en yüksek DMF değerini -1.03 logit değeri ile 9.madde alırken en yüksek DMF değerini 1.23 logit değeri ile 12.madde almıştır. Kız öğrencilerin Türkçe alt testi maddelerine ilişkin DMF ölçüm değerlerin ise -1.29 logit değeri ile 1.41 logit değeri arasında değiştiği gözlemlenmektedir. Türkçe testine dair kız öğrencilerin aldığı puanlara göre madde bazında en düşük DMF değeri -1.29 logit değeri ile 9.maddesi alırken, en yüksek DMF puanını 1.41 logit değeri ile 12.maddesinin aldığı görülmektedir.

Tablo-7'de Türkçe testinin cinsiyet bazında DMF'ye ilişkin değerleri incelendiğinde maddelerin DMF kontrast değerleri -0.35 logit ile 0.31 logit arasında değerler aldığı görülmektedir. Türkçe testinde en düşük değeri -0.35 logit değeri ile 16.madde alırken en yüksek değeri 0.31 logit değeri ile 10.madde almıştır. DMF kontrast değerlerinin negatif değer alması yanlılığın odak grup lehine(birinci) olması anlamına gelmektedir. Uygun DMF kontrast değeri, -0.5 ile 0.5 logit değeri arasındadır. Ayrıca Türkçe alt testine ilişkin maddelerin t değeri -33.2 ile 25.06 arasında değer aldığı görülmektedir. Bu t değerlerinin en düşüğünü -33.2 değeri ile 16.madde alırken en yüksek değerini, 25.06 ile 10.madde almıştır.

Aşağıda Türkçe alt testine ait maddeler bazında gruplara ilişkin Değişen Madde Fonksiyonu (DMF) kontrast değeri şekil halinde verilmiştir.



Şekil-13: Türkçe Maddelerinin Cinsiyete göre DMF Değişimi

Şekil-13'te Türkçe alt testinin cinsiyete göre DMF değişimi gösterilmiştir. Şekilde yeşil renk ortalama DMF değerini yansıtırken kırmızı renk olan ölçümler kız öğrencilerin Türkçe testi maddelerindeki DMF oranını temsil etmektedir. Siyah renkte olan ölçümler ise erkek öğrencilerin Türkçe testi maddelerindeki DMF oranını temsil etmektedir. Kırmızı ve siyah renkteki değişen DMF oranları farkı .5 üzerinde olan maddelerin anlamlı düzeyde DMF içerdiğini göstermektedir. Şekilde görüldüğü gibi bütün maddelerde .5'ten daha az bir oranda değişimin mevcut olduğu anlaşılmaktadır. Şekilde en yüksek oranda değişimin 4.madde, 10.madde ve 16. maddede olduğu görülmektedir.

4.1.5. Matematik Alt Testine İlişkin Bulgular

Tablo-8'da SBS-2012 Matematik alt testine ilişkin madde güçlük düzeyi, madde iç ve dış uyum indeksleri ve nokta çift serili madde test korelasyonu verilmiştir.

Tablo-8: Matematik Alt Testine İlişkin Madde Güçlük Düzeyi, Uyum İndeksi ve Nokta Çift Serili Korelasyon Değerleri:

Ham Puan	Toplam Puan	Madde Güçlük Düzeyi	Model S.E	İç uyum		Dış Uyum		Nokta Çift Serili Kor. (PTMEA)	Maddeler
				MNSQ	ZSTD	MNSQ	ZSTD		
31289	168731	1.86	.01	1.49	9.9	2.46	9.9	.22	10
33166	110524	.91	.01	1.55	9.9	1.77	9.9	.31	20
50629	163461	.90	.01	.84	-9.9	.84	-9.9	.63	18
39935	116875	.61	.01	1.24	9.9	1.35	9.9	.47	14
40980	120913	.56	.01	1.24	9.9	1.29	9.9	.46	5
75884	200511	.37	.01	.87	-9.9	.84	-9.9	.60	9
61031	154958	.31	.01	.96	-9.9	.95	-9.9	.58	6
54306	127774	.15	.01	.98	-7.0	.95	-9.2	.58	13
61757	142345	.13	.01	.91	-9.9	.87	-9.9	.61	11
73449	168098	.11	.01	.99	-5.5	.97	-6.6	.55	12
62840	146640	.02	.01	1.13	9.9	1.16	9.9	.48	1
84228	181380	-.05	.01	.95	-9.9	.92	-9.9	.55	17
84278	174730	-.17	.01	.87	-9.9	.84	-9.9	.59	8
95113	203354	-.18	.01	1.22	9.9	1.27	9.9	.38	7
95395	194761	-.27	.01	.83	-9.9	.76	-9.9	.61	2
99429	192563	-.42	.01	.83	-9.9	.75	-9.9	.61	3
107092	184560	-.71	.01	.87	-9.9	.81	-9.9	.56	19
122276	182267	-1.23	.01	.91	-9.9	.86	-9.9	.50	15
127142	187897	-1.25	.01	.86	-9.9	.74	-9.9	.53	16
138037	188292	-1.63	.01	1.00	-.3	1.02	3.1	.41	4

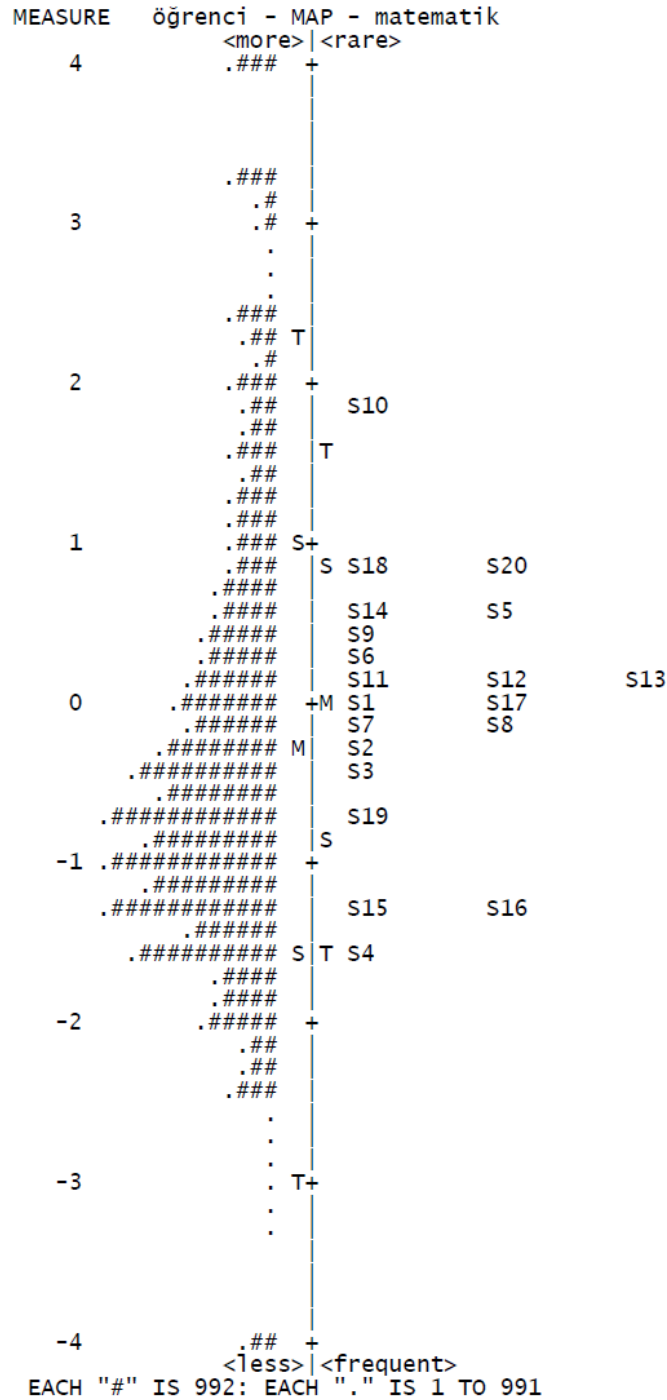
Tablo-8'de, 216363 kişinin SBS Matematik alt testine ait 20 maddesinin güçlük düzeyi ve uyumlarına ilişkin istatistiksel bilgileri verilmiştir. Tablo-8'de maddelerin güçlük düzeyi ve madde test uyumu incelenmiştir. Maddeler güçlük düzeyi yüksek olandan düşük olana doğru sıralanmıştır. Tabloda maddelere doğru cevap veren sayısı, maddeyi yanıtlayan toplam kişi sayısı, madde güçlük katsayısı, modelin standart hatası, iç uyum ortalamalar karesi (MNSQ), iç uyum (ZSTD) katsayısı, dış uyum ortalamalar karesi(MSQ), dış uyum (ZSTD) katsayısı, nokta serili madde test korelasyon katsayısı ve madde numaraları verilmiştir. Tablo-8'de bireylerin soruları cevaplama oranlarına bakıldığında, en yüksek cevaplama oranı 203354 kişi ile 7. madde iken en düşük oran 110524 kişi ile 20.maddenin olduğu tespit edilmiştir.

Tablo-8’de bireylerin maddeleri doğru cevaplamalarına ilişkin verilere bakıldığında en yüksek doğru cevaplanma oranı 138037 kişi tarafından doğru cevaplanan 4.maddeyken en düşük doğru cevaplanma oranı 31289 kişi ile 10. madde olarak gözlemlenmektedir. 20 maddelik Matematik alt testi incelendiğinde, madde güçlük düzeyi en yüksek olan madde 1,86 ile 10.madde iken güçlük düzeyi en düşük madde -1,63 ile 4.madde olarak saptanmıştır. Modelin standart hatası tüm maddeler için 0.1 olarak hesaplanmıştır.

Maddelerin test uyumunu incelemek için; iç uyum (MNSQ) ve dış uyum (MNSQ) değerleri incelenmektedir. Madde uyum (fit) istatistikleri MNSQ (ortalamalar karesi) beklenen değerleri 1.0 dır. MNSQ değerinin 2’den küçük olması ölçüm sistemini bozmaktadır. Değerin $1.5 < MNSQ \leq 2.0$ arasında olması, ölçüm yapısı için verimsizdir. Değerin $0.5 < MNSQ \leq 1.5$ arasında olması ölçüm için verimlidir. MNSQ < 0.5 olması ise ölçüm için daha az verimlidir (Linacre, 2002, Smith, 2000). Tablo-8’de Matematik alt testine ilişkin 20 maddenin uyum indeksleri incelendiğinde, iç uyum indekslerinin 0.83 ile 1.55 arasında değiştiği gözlemlenmektedir. 2.madde ve 3.madde 0.83 değerini alırken 20.madde 1.55 değerini almıştır. Mevcut maddelerin dış uyum ortalamar karesi indeksi incelendiğinde maddelerin 0.74 ile 2.46 arası değer aldığı gözlemlenmektedir. 16.madde 0.74 değeri ile en düşük değeri alırken 10.madde 2.46 ortalamar karesi değeri ile en yüksek değeri almıştır.

Tablo-8 incelendiğinde Matematik testine alt ilişkin maddelerin nokta çift serili (PTMEA) korelasyon değerleri 0.22 ile 0.63 değerleri arasında değişmektedir. En düşük PTMEA korelasyon değerini 0.22 ile 10.madde alırken, en yüksek değeri 0.63 ile 18.maddenin aldığı gözlemlenmektedir.

INPUT: 216363 öğrenci 23 matematik REPORTED: 214722 öğrenci 20 matematik 2
 CATS WINSTEPS 3.80.1



Şekil-14: Matematik Alt Testi Madde Güçlük ve Öğrenci Dağılımı:

Şekil-14'te 216363 kişiye ait SBS 2012 Matematik alt testine ilişkin 20 maddenin, güçlük ve bireyler bazında dağılımı verilmiştir. Şekil-14'te her “#” işareti 992 kişiyi temsil etmektedir. Yapılan analizler sonucunda maddelerden güçlük düzeyi en yüksek olan madde Soru-10 olarak tespit edilmiştir. Bu soruyu takiben

Soru-18 ve Soru-20 yaklaşık olarak aynı güçlük düzeyinde olduğu tespit edilmiştir. Soru-14 ve Soru-5 maddeleri, bu maddelerin güçlük düzeyinden daha az güçlük düzeyinde olduğu tespit edilmiştir. Buna karşın madde güçlük düzeyi en düşük olan madde Soru-4 ve bu maddeyi takiben Soru-15 ve Soru-16 en düşük madde güçlük düzeyinde olduğu tespit edilmiştir. Soru-19'un madde güçlük değeri bu soruları takip etmektedir. Diğer maddelerin orta güçlük düzeyinde olduğu tespit edilmiştir. Maddelerin güçlük düzeyi yükseldikçe maddenin zorluk derecesi artmaktadır.

2012-SBS Matematik alt testine ilişkin Değişen Madde Fonksiyonu (DMF) kontrast değeri Tablo-9'da sunulmuştur.

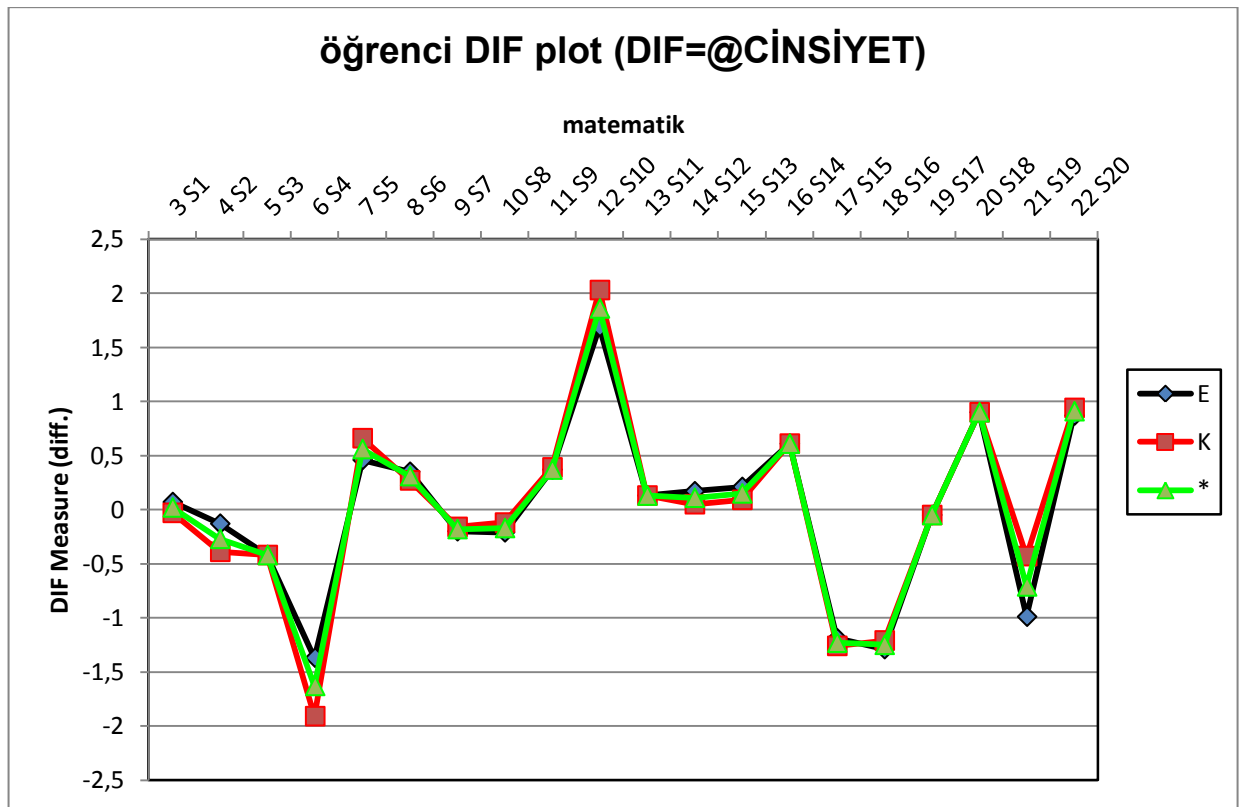
Tablo-9: Matematik Alt Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değerleri:

Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	T	Madde
Erkek	.07	Kız	-.03	.10	8.24	1
Erkek	-.13	Kız	-.39	.26	24.36	2
Erkek	-.42	Kız	-.42	.00	.00	3
Erkek	-1.37	Kız	-1.91	.54	45.94	4
Erkek	.46	Kız	.66	-.20	-13.1	5
Erkek	.35	Kız	.27	.07	5.90	6
Erkek	-.20	Kız	-.16	-.05	-4.43	7
Erkek	-.21	Kız	-.12	-.09	-8.02	8
Erkek	.37	Kız	.39	-.02	-1.83	9
Erkek	1.69	Kız	2.03	-.34	-22.6	10
Erkek	.13	Kız	.13	.00	.00	11
Erkek	.17	Kız	.05	.11	9.73	12
Erkek	.21	Kız	.09	.12	8.97	13
Erkek	.61	Kız	.61	.00	.00	14
Erkek	-1.19	Kız	-1.26	.07	6.09	15
Erkek	-1.29	Kız	-1.21	-.08	-7.31	16
Erkek	-.05	Kız	-.05	.00	.00	17
Erkek	.90	Kız	.90	.00	.00	18
Erkek	-.99	Kız	-.43	-.56	-50.8	19
Erkek	.87	Kız	.94	-.07	-4.43	20

Tablo-9'da katılımcıların cinsiyetlerine göre Matematik alt testinin 20 maddesine vermiş oldukları cevaplara göre DMF ölçüm değerleri verilmiştir. Tablo-9'da erkek katılımcıların maddelere göre DMF ölçüm değerleri, kızların maddelere göre DMF değerleri, toplam Dmf ölçüm değeri, t değerleri ve madde numaraları verilmiştir.

Tablo-9'a göre erkek öğrencilerin Matematik alt testi maddelerinin Dmf ölçüm değerleri -1.37 ile 1.69 logit arasında değiştiği görülmektedir. Matematik alt testine ilişkin erkek yanıtlayıcılarının en düşük DMF değerini -1.37 logit ile 4.madde alırken en yüksek DMF değerini 1.69 logit ile 10.madde almıştır. Kız öğrencilerin Matematik alt testine ilişkin madde DMF ölçüm değerleri ise -1.91 logit ile 2.03 logit arasında değerler aldığı görülmektedir. Matematik alt testine dair kız öğrencilerin aldığı puanlara göre madde bazında en düşük DMF değerini -1.29 logit ile 4.madde alırken en yüksek DMF puanını ise 1.41 logit değeri ile 10.madde almıştır.

Tablo-9'da Matematik alt testinin cinsiyet bazında DMF'ye ilişkin değerleri incelendiğinde, maddelerin DMF kontrast değerlerinin -0.56 logit ile 0.54 logit arasında değerler aldığı görülmektedir. Matematik alt testinde en düşük değeri -0.56 logit ile 19.madde alırken en yüksek değeri 0.54 logit değeri ile 4.madde almıştır. DMF kontrast değerlerinin negatif değer alması yanlılığın odak (birinci) grup lehine olması anlamına gelmektedir. DMF'nin uygun değeri 0.5 ile -0.5 değer aralığıdır. Ayrıca Matematik testine ilişkin maddelerin t değeri -50.8 ile 45.94 arasında değer aldığı görülmektedir. Bu t değerlerinin en düşüğü -50.8 değeri ile 19.madde alırken en yüksek değeri 45.94 ile 4.madde almıştır.



Şekil-15: Matematik Maddelerinin Cinsiyete göre DMF Değişimi

Şekil-15'te Matematik alt testinin cinsiyete göre DMF değişimi gösterilmiştir. Şekilde yeşil renkte olan DMF değeri, ortalama DMF değerini yansıtırken kırmızı renkte olan ölçümler kız öğrencilerin Matematik alt testi maddelerindeki DMF oranını temsil etmektedir. Siyah renkte olan ölçümler ise erkek öğrencilerin Matematik alt testi maddelerindeki DMF oranını temsil etmektedir. Kırmızı ve siyah renkler arasındaki DMF değişim oranlarının 0.5 üzerinde olması, maddelerin anlamlı düzeyde DMF içerdiğini göstermektedir. Şekil-15'te görüldüğü gibi 19. ve 4.maddelerin 0.5'ten fazla bir oranda DMF kontrast değeri içerdiği görülmektedir. Diğer bütün maddelerde 0.5'ten daha az bir oranda değişimin mevcut olduğu anlaşılmaktadır. Şekilde en yüksek oranda değişimin 19. ve 4. maddelerde olduğu görülmektedir.

4.1.6.Fen ve Teknoloji Alt Testine İlişkin Bulgular

Aşağıdaki tabloda Fen ve Teknoloji alt testine ilişkin madde güçlük düzeyi, uyum indeksleri ve nokta çift serili (PTMEA) korelasyon değerleri verilmiştir.

Tablo-10: Fen ve Teknoloji Alt Testine İlişkin Madde Güçlük Düzeyi, Uyum İndeksi ve Nokta Çift Serili Korelasyon Değerleri:

Ham Puan	Toplam Puan	Madde Güçlük Düzeyi	Model S.E	İç uyum		Dış Uyum		Nokta Çift Serili Kor. (PTMEA)	Maddeler
				Mnsq	ZSTD	Mnsq	ZSTD		
44768	201123	1.70	.01	1.23	9.9	1.63	9.9	.33	1
57855	190763	1.18	.01	1.24	9.9	1.46	9.9	.35	16
65721	188173	.89	.01	.97	-9.9	1.03	7.6	.52	14
77622	205663	.68	.01	1.05	9.9	1.10	9.9	.47	5
74348	184025	.58	.01	1.18	9.9	1.28	9.9	.40	9
76889	188460	.55	.01	.92	-9.9	.94	-9.9	.55	2
76178	173656	.42	.01	1.04	9.9	1.08	9.9	.49	17
83682	188667	.36	.01	1.20	9.9	1.31	9.9	.37	20
89955	202939	.32	.01	.99	-2.8	1.02	4.3	.50	11
79129	170646	.25	.01	1.00	-1.5	.98	-4.4	.52	12
92089	169990	-.18	.01	.90	-9.9	.83	-9.9	.56	10
114738	204703	-.30	.01	.95	-9.9	.89	-9.9	.51	15
109638	189985	-.34	.01	.86	-9.9	.78	-9.9	.57	19
126553	208464	-.56	.01	.91	-9.9	.87	-9.9	.52	7
110070	180610	-.57	.01	1.04	9.9	1.10	9.9	.44	4
125353	194314	-.76	.01	.98	-9.9	1.05	9.4	.46	8
134331	202431	-.85	.01	.88	-9.9	.81	-9.9	.52	18
136128	203537	-.89	.01	.90	-9.9	.83	-9.9	.50	13
147351	207222	-1.14	.01	.85	-9.9	.72	-9.9	.52	6
157224	211845	-1.35	.01	.88	-9.9	.85	-9.9	.48	3

Tablo-10'da 216363 kişinin SBS Fen ve Teknoloji alt testi 20 maddesinin güçlük ve uyumlarına ilişkin istatistiksel bilgileri verilmiştir. Tablo-10'da maddelerin güçlük düzeyi ve madde test uyumu incelenmiştir. Maddeler güçlük düzeyi yüksek olandan düşük olana doğru sıralanmıştır. Tabloda maddelere doğru cevap veren kişi sayısı, maddeyi yanıtlayan toplam kişi sayısı, madde güçlük katsayısı, modelin standart hatası, iç uyum ortalamalar karesi (MNSQ), iç uyum (ZSTD) katsayısı, dış uyum ortalamalar karesi (MSQ), dış uyum (ZSTD) katsayısı, nokta serili test korelasyon katsayısı ve madde numaraları verilmiştir. Tablo-10'da bireylerin soruları cevaplama oranlarına bakıldığında en yüksek cevaplama oranı 211845 kişi ile 3. Madde ait iken en düşük oran 170646 kişi ile 12.maddeye ait olduğu tespit edilmiştir.

Tablo-10'da bireylerin maddeleri doğru cevaplamalarına ilişkin verilere bakıldığında en yüksek doğru cevaplanma oranı 157224 kişi tarafından doğru cevaplanan 3.maddeyken en düşük doğru cevaplanma oranı 44768 kişi ile 1. madde olarak gözlemlenmektedir. 20 maddelik Fen ve Teknoloji alt testi incelendiğinde madde güçlük düzeyi en yüksek olan madde 1,70 ile 1.madde iken güçlük düzeyi en düşük madde -1,35 ile 3.madde olarak saptanmıştır. Modelin standart hatası tüm maddeler için 0.1 olarak hesaplanmıştır.

Maddelerin test uyumu; iç uyum(MNSQ) ve dış uyum (ÖNSQ) değerleri açısından incelenmiştir. Madde uyum (fit) istatistikleri için beklenen MNSQ (ortalamalar karesi) değerleri 1.0 dır. MNSQ değerinin 2'dan büyük olması, sistemini bozmaktadır. Değerin $1.5 < \text{MNSQ} \leq 2.0$ arasında olması ölçüm yapısı için verimsizdir. Değerin $0.5 < \text{MNSQ} \leq 1.5$ arasında olması ölçüm için verimlidir. $\text{MNSQ} < 0.5$ olması ise ölçüm için daha az verimlidir (Linacre, 2002, Smith, 2000). Tablo-10 incelendiğinde Fen ve Teknoloji alt testine ilişkin 20 maddenin uyum indeksleri incelendiğinde iç (infit) uyum indekslerinin 0.85 ile 1.25 arasında değiştiği gözlemlenmektedir. 6.madde 0.85 değeri ile en küçük değeri alırken 16.madde ise 1.25 değeri ile en büyük değeri almıştır. Mevcut maddelerin dış uyum (outfit) ortalamalar karesi indeksi incelendiğinde, maddelerin 0.72 ile 1.63 arasında değerler aldığı gözlemlenmektedir. 6.madde 0.72 değeri ile en düşük değer alırken 1.madde 1.63 ortalamalar karesi değeri ile en yüksek değeri almıştır.

Şekil-16’da 216363 kişiye ait SBS 2012 Fen ve Teknoloji alt testine ilişkin 20 maddenin güçlük ve bireyler bazında dağılımı verilmiştir. Şekil-16’da her “#” işareti 977 kişiyi temsil etmektedir. Yapılan analizler sonucunda maddelerden güçlük düzeyi en yüksek olan madde Soru-1 olarak tespit edilmiştir. Bu maddeyi takiben Soru-16’nın güçlük düzeyi yüksek madde olarak geldiği Şekil-16’dan anlaşılmaktadır. Soru-14’ün bu soruların güçlük düzeyine göre daha az güçlük düzeyinde olduğu tespit edilmiştir. Buna karşın madde güçlük düzeyi en düşük olan madde Soru-3 olduğu ve bu maddeyi takiben Soru-6’nın düşük madde güçlük düzeyinde olduğu tespit edilmiştir. Soru-13 ve Soru-18 benzer güçlük düzeyinde bu soruları takip etmektedir. Diğer maddelerin orta güçlük düzeyinde olduğu tespit edilmiştir. Maddelerin güçlük düzeyi yükseldikçe maddenin zorluk derecesi artmaktadır.

Aşağıdaki tabloda Fen ve Teknoloji alt testine ilişkin Değişen Madde Fonksiyonu (DMF) değerleri verilmiştir.

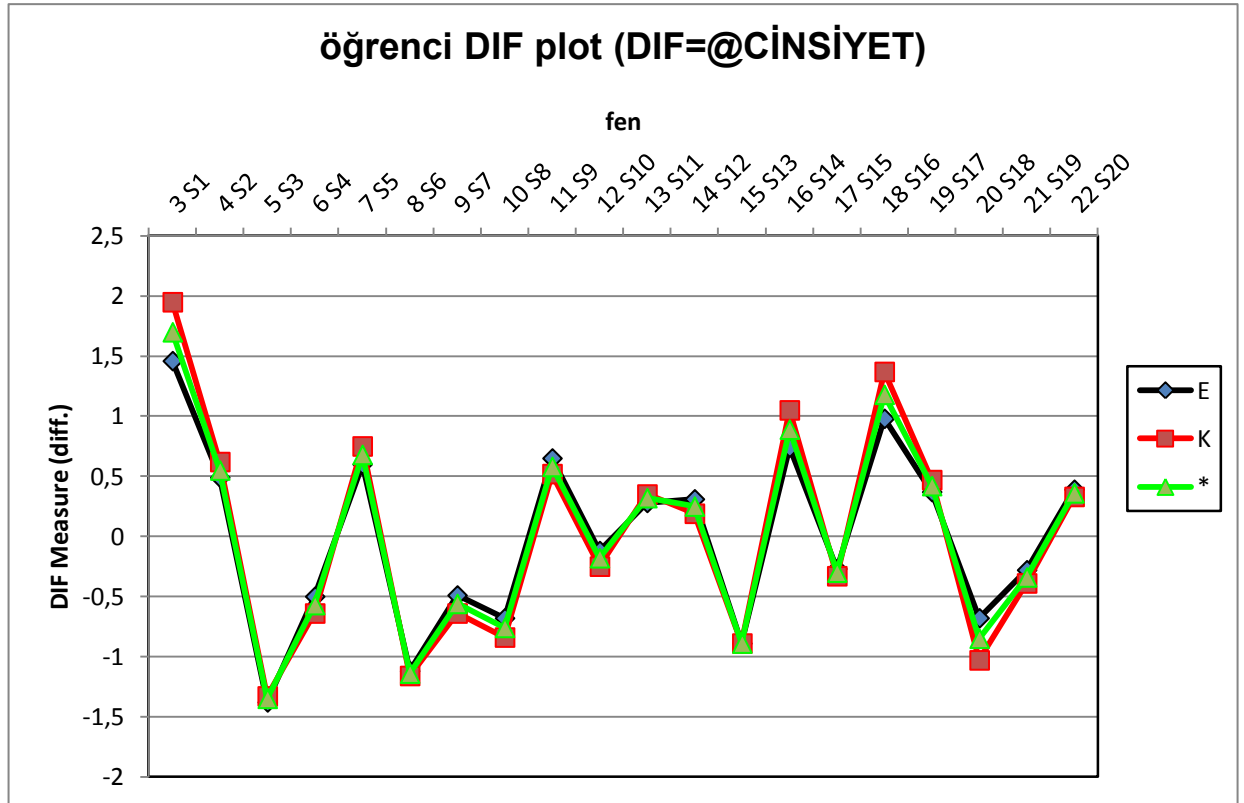
Tablo-11: Fen ve Teknoloji Alt Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değerleri:

Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	T	Madde
Erkek	1.46	Kız	1.95	-.49	-38.8	1
Erkek	.49	Kız	.62	-.13	-11.6	2
Erkek	-1.38	Kız	-1.33	-.05	-4.96	3
Erkek	-.50	Kız	-.64	.14	12.49	4
Erkek	.60	Kız	.75	-.15	-14.1	5
Erkek	-1.12	Kız	-1.16	.05	4.18	6
Erkek	-.49	Kız	-.64	.15	14.85	7
Erkek	-.68	Kız	-.84	.17	15.51	8
Erkek	.65	Kız	.52	.13	11.55	9
Erkek	-.12	Kız	-.25	.13	11.07	10
Erkek	.28	Kız	.35	-.07	-6.27	11
Erkek	.31	Kız	.19	.12	10.62	12
Erkek	-.89	Kız	-.89	.00	.00	13
Erkek	.74	Kız	1.05	-.31	-27.6	14
Erkek	-.27	Kız	-.33	.06	5.56	15
Erkek	.98	Kız	1.37	-.39	-33.1	16
Erkek	.37	Kız	.47	-.11	-9.29	17
Erkek	-.68	Kız	-1.03	.35	32.51	18
Erkek	-.28	Kız	-.39	.10	9.39	19
Erkek	.39	Kız	.33	.06	5.08	20

Tablo-11’de katılımcıların cinsiyetleri bakımından Fen ve Teknoloji alt testinin 20 maddesine vermiş oldukları cevaplara göre DMF ölçüm değerleri verilmiştir. Tablo-11’de erkeklerin maddelere göre DMF ölçüm değerleri, kızların maddelere göre DMF değerleri, toplam DMF ölçüm değeri, t değerleri ve madde numaraları verilmiştir.

Tablo-11’e göre erkek öğrencilerin Fen ve Teknoloji alt testi maddeleri DMF ölçüm değerlerinin -1.38 ile 1.46 logit arasında değiştiği görülmektedir. Fen ve Teknoloji alt testi erkek yanıtlayıcılarının en düşük DMF değeri -1.38 logit ile 3.maddeyken en yüksek DMF değeri 1.46 logit ile 1.madde olmuştur. Kız öğrencilerin Fen ve Teknoloji alt testi maddelerinin DMF ölçüm değerleri ise -1.33 logit ile 1.95 logit arasında değerler aldığı görülmektedir. Fen ve Teknoloji alt testine dair kız öğrencilerin aldığı puanlara göre madde bazında en düşük DMF değerini -1.33 logit değeri ile 3.madde alırken en yüksek DMF puanını 1.95 logit değeri ile 11.maddenin aldığı görülmektedir.

Tablo-11’de Fen ve Teknoloji alt testinin cinsiyet bazında, DMF’ye ilişkin değerleri incelendiğinde, maddelerin DMF kontrast değerleri -0.49 logit ile 0.35 logit değeri arasında değerler aldığı görülmektedir. Fen ve Teknoloji alt testinde en düşük değeri -0.49 logit değeri ile 1.madde alırken en yüksek değeri 0.35 logit değeri ile 18.madde almıştır. DMF kontrast değerlerinin negatif değer alması maddenin odak grup lehine(birinci) çalıştığını yani odak gruba avantaj sağladığını göstermektedir. DMF değerinin mutlak değer olarak 0.5 logit arasında değer alması maddenin DMF içermediğini göstermektedir. Ayrıca Fen ve Teknoloji alt testine ilişkin maddelerin t değeri -38.8 ile 32.51 arasında değer aldığı görülmektedir. Bu t değerlerinin en düşüğünü -38.8 değeri ile 1.madde alırken en yüksek değeri 32.51 ile 18.madde almıştır.



Şekil-17: Fen ve Teknoloji Maddelerinin Cinsiyete göre DMF Değişimi

Şekil-17’de Fen ve Teknoloji alt testinin cinsiyete göre DMF değişimi gösterilmiştir. Şekilde yeşil renkte olan ölçümler, DMF değerinin ortalama değerini yansıtırken kırmızı renkte olan ölçümler kız öğrencilerin Fen ve Teknoloji alt testi maddelerindeki DMF oranını temsil etmektedir. Siyah renkte olan ölçümler ise erkek öğrencilerin Fen ve Teknoloji alt testi maddelerindeki DMF oranını temsil etmektedir. Kırmızı ve siyah renkler arasındaki DMF ölçüm değeri 0.5 üzerindeyse, bu maddelerin anlamlı düzeyde DMF içerdiğini göstermektedir. Şekilde görüldüğü gibi bütün maddelerde 0.5’ten daha az bir oranda değişimin mevcut olduğu anlaşılmaktadır. Şekilde en yüksek oranda değişimin 1.madde, 16.madde ve 18. maddede olduğu görülmektedir.

4.1.7.Sosyal Bilgiler Alt Testine İlişkin Bulgular

Aşağıdaki tabloda Sosyal Bilgiler alt testine ilişkin madde günlük düzeyi ve uyum indeksleri verilmiştir.

Tablo-12: Sosyal Bilgiler Alt Testine İlişkin Madde Güçlük Düzeyi, Uyum İndeksi ve Nokta Çift Serili Korelasyon Değerleri:

Ham Puan	Toplam Puan	Madde Güçlük Düzeyi	Model S.E	İç uyum		Dış Uyum		Nokta Çift Serili Kor. (PTMEA)	Maddeler
				Mnsq	ZSTD	Mnsq	ZSTD		
57570	200336	2.04	.01	1.46	9.9	2.12	9.9	.19	4
58286	192294	1.94	.01	1.36	9.9	1.94	9.9	.26	11
83393	191302	1.20	.01	1.09	9.9	1.22	9.9	.45	13
84160	196450	1.16	.01	1.21	9.9	1.40	9.9	.38	3
107459	207751	.70	.01	1.01	3.3	1.02	6.5	.50	18
109890	203626	.59	.01	1.07	9.9	1.08	9.9	.46	17
115814	203643	.42	.01	.88	-9.9	.82	-9.9	.58	19
120404	201446	.26	.01	.91	-9.9	.84	-9.9	.56	8
122899	202045	.17	.01	.96	-9.9	.88	-9.9	.53	12
131031	206203	.03	.01	1.02	9.9	.97	-6.6	.48	10
147639	204530	-.48	.01	.82	-9.9	.66	-9.9	.58	7
151636	207823	-.55	.01	.94	-9.9	.82	-9.9	.50	6
145747	198864	-.57	.01	.89	-9.9	.75	-9.9	.53	2
162017	214002	-.75	.01	1.05	9.9	1.04	7.0	.41	1
162484	210146	-.85	.01	.94	-9.9	.86	-9.9	.47	16
157185	198628	-.95	.01	.81	-9.9	.61	-9.9	.55	5
164198	208466	-.95	.01	.97	-9.9	.93	-9.4	.44	15
156863	197685	-.96	.01	.79	-9.9	.58	-9.9	.56	14
168690	209805	-1.08	.01	.88	-9.9	.79	-9.9	.48	9
176429	210207	-1.37	.01	.84	-9.9	.64	-9.9	.49	20

Tablo-12’de 216363 kişinin SBS Sosyal Bilgiler alt testine ait 20 maddenin güçlük ve uyumlarına ilişkin istatistiksel bilgileri verilmiştir. Tablo-12’de maddelerin güçlük düzeyi ve madde test uyumu incelenmiştir. Maddeler güçlük düzeyi yüksek olandan düşük olana doğru sıralanmıştır. Tabloda maddelere doğru cevap veren katılımcı sayısı, maddeyi yanıtlayan toplam katılımcı sayısı, madde güçlük katsayısı, modelin standart hatası, iç uyum ortalamalar karesi (MNSQ), iç uyum (ZSTD) katsayısı, dış uyum ortalamalar karesi (MNSQ), dış uyum (ZSTD) katsayısı, nokta serili korelasyon katsayısı ve madde numaraları verilmiştir. Tablo-12’de bireylerin soruları cevaplama oranlarına bakıldığında en yüksek değer 214002 kişi ile 1. madde iken en düşük oranın ise 191302 kişi ile 13.madde olduğu tespit edilmiştir.

Tablo-12’de bireylerin maddeleri doğru cevaplamalarına ilişkin verilere bakıldığında en yüksek doğru cevaplanma oranı 176429 kişi ile 20.maddeyken, en düşük doğru cevaplanma oranı 57570 kişi ile 4. madde olarak gözlemlenmektedir. 20 maddelik Sosyal Bilgiler alt testi incelendiğinde madde güçlük düzeyi en yüksek ölçüm 2,04 ile 4.madde iken güçlük düzeyi en düşük madde -1,37 ile 20.madde

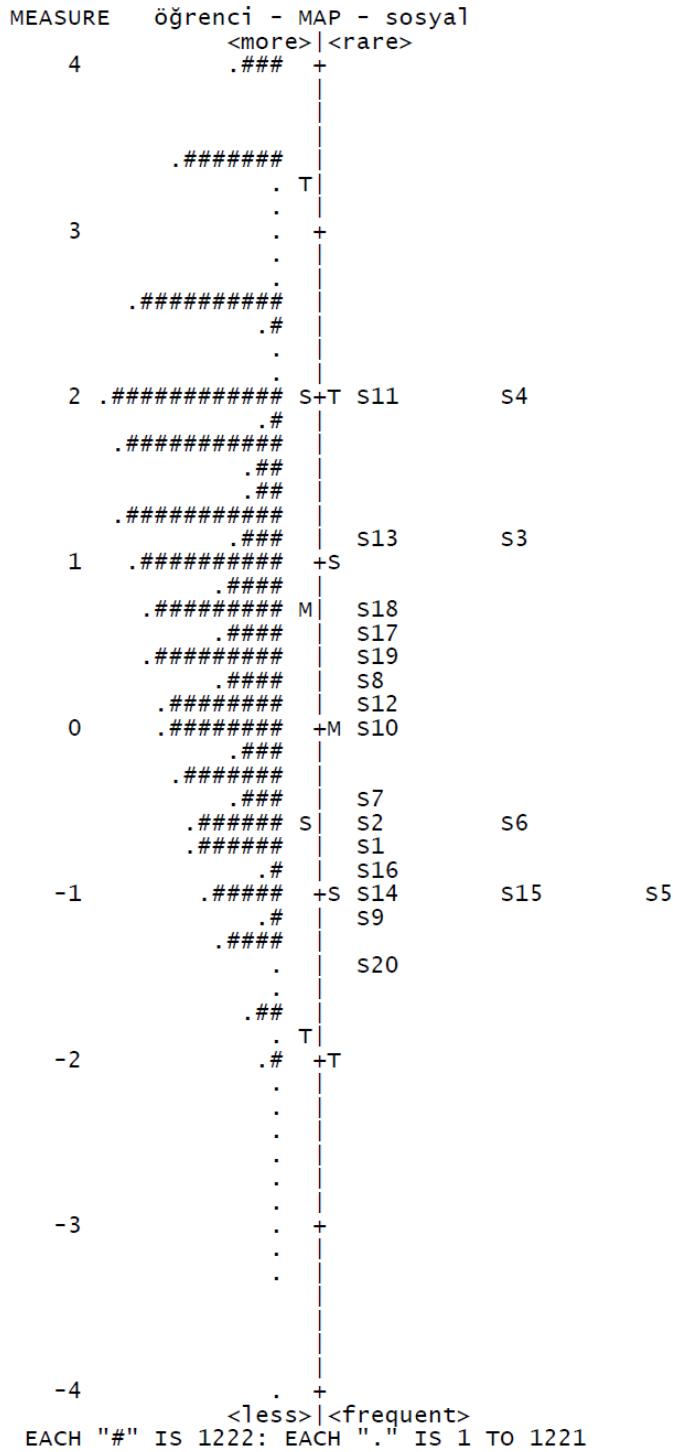
olarak saptanmıştır. Modelin standart hatası tüm maddeler için 0.1 olarak hesaplanmıştır.

Maddelerin test uyumu; iç uyum(MNSQ) ve dış uyum(MNSQ) değerlerine bakılarak incelenmektedir. Madde uyum(fit) istatistikleri MNSQ (ortalamalar karesi) beklenen değerleri 1.0 dır. $MNSQ > 2.0$ olması ölçüm sistemini bozmaktadır. Değerin $1.5 < MNSQ \leq 2.0$ arasında olması ölçüm yapısı için verimsizdir. Değer $0.5 < MNSQ \leq 1.5$ arasında olursa ölçüm için verimlidir. MNSQ değerinin 0.5'ten küçük olması ölçüm için daha az verimlidir (Linacre, 2002, Smith, 2000). Tablo-12 incelendiğinde Sosyal Bilgiler alt testine ilişkin 20 maddenin uyum indeksleri incelendiğinde iç uyum indekslerinin 0.79 ile 1.46 arasında değiştiği gözlemlenmektedir. 14.madde 0.79 iken 4. madde 1.46 değerini almıştır. Mevcut maddelerin dış uyum ortalamalar karesi indeksi incelendiğinde maddelerin 0.58 ile 2.12 arası değer aldığı gözlemlenmektedir. 14.madde 0.58 değeri ile en düşük değer alırken 4.madde 2.52 ortalamalar karesi değeri ile en yüksek değeri almıştır.

Tablo-12 incelendiğinde maddelerin nokta çift serili (PTMEA) korelasyon değerleri 0.19 ile 0.58 değerleri arasında değişmektedir. En düşük değeri 0.19 ile 4.madde alırken en yüksek değeri 0.58 ile 7. ve 19.maddelerin aldığı gözlemlenmektedir.

Aşağıdaki şekilde maddelere ilişkin güçlük düzeyini bireyler bazında dağılımı verilmiştir.

INPUT: 216363 öğrenci 23 sosyal REPORTED: 216206 öğrenci 20 sosyal 2 CATS
WINSTEPS 3.80.1



Şekil-18: Sosyal Bilgiler Alt Testi Madde Güçlük ve Öğrenci Dağılımı:

Şekil-18’de 216363 kişiye ait SBS 2012 Sosyal Bilgiler alt testine ilişkin 20 sorunun güçlük ve bireyler bazında dağılımı verilmiştir. Şekil-18’de her “#” işareti

1221 kişiyi temsil etmektedir. Yapılan analizler sonucunda maddelerden Soru-4 ile Soru-11'in benzer güçlük düzeyi ile en yüksek olan maddeler olduğu tespit edilmiştir. Bu maddeleri takiben Soru-3 ve Soru-13 benzer güçlük düzeyinde olduğu tespit edilmiştir. Soru-18'in bu maddelerin güçlük düzeyinden daha az güçlük düzeyinde olduğu tespit edilmiştir. Buna karşın madde güçlük düzeyi en düşük olan madde Soru-2 ve bu maddeyi takiben Soru-9'un düşük madde güçlük düzeyinde olduğu tespit edilmiştir. Soru-5, Soru-14 ve soru-15 benzer güçlük düzeyinde bu maddeleri takip etmektedir. Diğer maddelerin orta güçlük düzeyinde olduğu tespit edilmiştir. Maddelerin güçlük düzeyi yükseldikçe maddenin zorluk derecesi artmaktadır.

Aşağıdaki tabloda Sosyal Bilgiler alt testine ilişkin Değişen Madde Fonksiyonu (DMF) değerleri verilmiştir.

Tablo-13: Sosyal Bilgiler Alt Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değerleri:

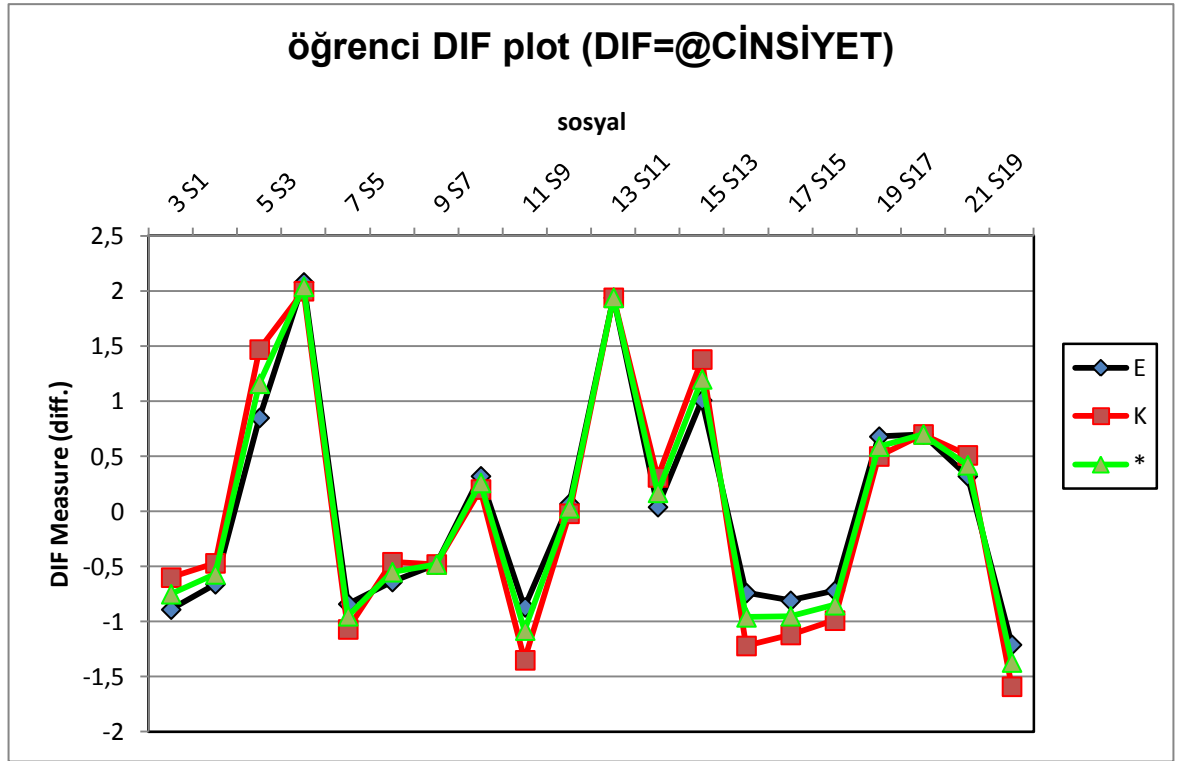
Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	T	Madde
Erkek	-.89	Kız	-.60	-.29	-25.4	1
Erkek	-.66	Kız	-.47	-.19	-16.1	2
Erkek	.85	Kız	1.47	-.62	-57.5	3
Erkek	2.08	Kız	2.00	.08	7.13	4
Erkek	-.84	Kız	-1.07	.23	18.01	5
Erkek	-.64	Kız	-.46	-.19	-16.6	6
Erkek	-.48	Kız	-.48	.00	.00	7
Erkek	.32	Kız	.20	.12	11.12	8
Erkek	-.87	Kız	-1.35	.48	38.38	9
Erkek	.07	Kız	-.02	.10	8.96	10
Erkek	1.94	Kız	1.94	.00	.00	11
Erkek	.04	Kız	.31	-.27	-25.6	12
Erkek	1.01	Kız	1.38	-.37	-34.2	13
Erkek	-.74	Kız	-1.22	.47	36.91	14
Erkek	-.81	Kız	-1.12	.31	25.51	15
Erkek	-.72	Kız	-.99	.27	22.63	16
Erkek	.68	Kız	.50	.18	17.25	17
Erkek	.70	Kız	.70	.00	.00	18
Erkek	.32	Kız	.51	-.20	-19.1	19
Erkek	-1.21	Kız	-1.59	.37	27.84	20

Tablo-13'te katılımcıların cinsiyetlerine göre Sosyal Bilgiler alt testinin 20 maddesine vermiş oldukları cevaplara göre DMF ölçüm değerleri verilmiştir. Tablo-13'te erkeklerin maddelere göre DMF ölçüm değerleri, kızların maddelere göre DMF değerleri, DMF kontrast ölçüm değeri, t değerleri ve madde numaraları verilmiştir.

Tablo-13'e göre erkek öğrencilerin Sosyal Bilgiler alt testi maddelerine ilişkin DMF ölçüm değerlerinin -1.21 ile 2.08 logit değeri arasında değiştiği görülmektedir. Sosyal Bilgiler alt testine ilişkin erkek yanıtlayıcılarının en düşük DMF değerini -1.21 logit değeri ile 20.madde alırken en yüksek DMF değerini 2.08 logit değeri ile 4.madde almıştır. Kız öğrencilerin Sosyal Bilgiler alt testi madde DMF ölçüm değerleri ise -1.59 logit ile 2.00 logit arası değerler aldığı görülmektedir. Sosyal Bilgiler alt testine dair kız öğrencilerin aldığı puanlara göre madde bazında en düşük DMF değerini -1.59 logit değeri ile 20.maddesi alırken en yüksek DMF puanını 2.00 logit değeri ile 4.maddenin aldığı görülmektedir.

Tablo-13'te Sosyal Bilgiler alt testinin cinsiyet bazında DMF'ye ilişkin değerleri incelendiğinde, maddelerin DMF kontrast değerleri -0.62 logit ile 0.48 logit arasında değerler aldığı görülmektedir. Sosyal Bilgiler alt testinde en düşük değeri -0.62 logit değeri ile 3.madde alırken en yüksek değeri 0.48 logit değeri ile 9.madde almıştır. DMF kontrast değerlerinin negatif değer alması yanlılığın odak grup(birinci) lehine olması anlamına gelmektedir. DMF ölçüm değerinin mutlak değer olarak 0.5 logit aralında bir değer alması beklenen bir değerdir. Ayrıca Sosyal Bilgiler alt testine ilişkin maddelerin t değerlerinin -33.2 ile 25.06 arasında değer aldığı görülmektedir. Bu t değerlerinin en düşüğü -57.5 değeri ile 3.madde alırken en yüksek değeri 38.38 ile 9.madde almıştır.

Aşağıdaki şekilde Sosyal Bilgiler alt testine ilişkin cinsiyet bazında Değişen Madde Fonksiyonu (DMF) ölçüm dağılımı verilmiştir.



Şekil-19: Sosyal Bilgiler Maddelerinin Cinsiyete göre DMF Değişimi

Şekil-19'da Sosyal Bilgiler alt testinin cinsiyete göre DMF değişimi gösterilmiştir. Şekilde yeşil renk DMF değerinin ortalama değerini yansıtırken kırmızı renkte olan ölçümler kız öğrencilerin Sosyal Bilgiler alt testi maddelerindeki DMF oranını temsil etmektedir. Siyah renkte olan ölçümler ise erkek öğrencilerin Sosyal Bilgiler alt testi maddelerindeki DMF oranını temsil etmektedir. Kırmızı ve siyah renklerdeki DMF oranları arasındaki fark 0.5 üzerinde olan maddeler için anlamlı düzeyde DMF içerdiğini göstermektedir. Şekilde 3. ,9. ve 14. maddelerin yaklaşık olarak 0.5'ten büyük değer aldıkları görülmektedir. Şekilde görüldüğü gibi diğer bütün maddeler arasındaki değişimin 0.5'ten daha az bir oranda değişim gösterdiği anlaşılmaktadır. Şekilde en yüksek oranda değişimin 3.madde, 9.madde ve 14. maddede olduğu görülmektedir.

4.1.8. Değişen Madde Fonksiyonu (DMF) İçeren Maddelere İlişkin Uzman Görüşlerine ait Bulgular

Yapılan istatistiksel işlemler sonucunda iki Matematik alt testi maddesi ve bir Sosyal Bilgiler alt testi maddesi olmak üzere toplam üç maddenin Değişen Madde Fonksiyonu (DMF) içerdiği tespit edilmiştir. Elde edilen DMF'li maddelerin madde

etkisinden mi, yanlılıktan mı kaynaklandığını tespit etmek için uzman görüşlerine başvurulmuştur. Matematik alt testi 4. maddesine ilişkin uzman görüşlerinden iki katılımcı maddenin erkekler lehine bir uzman bayanlar lehine diğer uzmanlar ise madde etkisinden kaynaklandığını dile getirmiştir. Maddenin bayanlar lehine olmasından dolayı uzman görüşleri doğrultusunda maddenin yanlılık içermediği anlaşılmıştır. Benzer şekilde matematik alt testi 19. maddesine ilişkin uzman görüşleri incelendiğinde sadece bir katılımcının maddenin erkekler lehine olabileceği kanısı oluşmuştur. Diğer katılımcılar maddenin *madde etkisinden* kaynaklanabileceği dile getirmişlerdir. Sosyal Bilgiler alt testi üçüncü maddesi incelendiğinde ilgili uzmanlardan ikisi maddenin *madde yanlılığından* kaynaklanabileceği görüşünü öne sürmüştür.

4.2.TARTIŞMA

Bir testte bulunması gereken özelliklerin başında geçerlik ve güvenilirlik gelmektedir. Geçerlik ve güvenilirliği kapsama yansıtacak maddeler seçilerek, sınav için uygun ortam hazırlanarak, maddelerin test ile tutarlı olmalarını sağlamak gibi önlemler alınarak sağlanabilir. Testin geçerlik ve güvenilirliğini sağlamanın bir yolu da test ile ilgili yanlılık çalışmaları yapmaktır. Ulusal düzeyde yapılan tüm sınavlar gibi SBS'nin de herhangi bir gruba avantaj sağlamaması gerekmektedir. Testlerde bulunabilecek bu olumsuzluğun önüne geçebilmek için yanlılık çalışmalarının yapılması kaçınılmazdır. Yanlılıkla ilgili çalışmalar yapılırken madde parametrelerinin gruptan bağımsız, grup yeteneğinin de teste bağlı olmaksızın kestirilmesi büyük önem taşımaktadır. Bu çalışmada, SBS alt testlerinin cinsiyet değişkenine göre, madde parametrelerinin gruptan bağımsız, grup yeteneğinin de teste bağlı olmaksızın Rasch modeli ile madde yanlılıkları incelenmiştir. Yapılan analizlerin Rasch modeline uygunluğu için testlerin; normal dağılım, tek boyutluluk, test maddelerinin yerel bağımsızlık özellikleri incelenmiştir. Yapılan analizler sonucunda verilerin Rasch modeli analizlerine uygun olduğu anlaşılmıştır.

23 maddelik Türkçe alt testi incelendiğinde madde güçlük düzeyi yüksek ve düşük olan maddelerin mevcut olduğu ancak testin genel olarak orta güçlükte olduğu anlaşılmaktadır. Madde test uyumu kapsamında, teste ilişkin iç ve dış uyum indeksleri ve nokta çift serili korelasyon (PTMEA) değerleri incelenmiştir. Analiz sonuçlarına göre teste ilişkin bütün madde sayısal verilerinin istenilen değer aralığına uygun olduğu ve testin genel yapısını bozmadığı anlaşılmaktadır.

2012 SBS Türkçe alt testine ilişkin DMF kontrast değerleri incelenmiştir. DMF kontrast değerlerinin negatif değer alması yanlılığın odak grup(birinci grup) lehine olduğunu göstermektedir. Beklenen DMF değeri sıfır iken mutlak değerce 0.5 değeri kabul edilebilir değerdir. Türkçe alt testine ilişkin analiz sonuçları incelendiğinde kimi maddelerin negatif kontrast değer kimilerin ise pozitif kontrast değer aldığı gözlemlenmesine karşın maddelerin tümünün beklenen aralıkta değer aldığı anlaşılmıştır

Sonuç olarak Tablo-7 ve Şekil-13'te Türkçe alt testine ilişkin cinsiyet bakımından DMF değerleri incelendiğinde, bütün maddelerin -0,50 ile 0,50 arasında değer aldığı tespit edilmiştir. Yani Türkçe alt testine ilişkin cinsiyet açısından bir yanlılığın olmadığı tespit edilmiştir. Diğer bir ifade ile 23 maddelik Türkçe alt testinin hiçbir sorusu cinsiyet açısından bir gruba avantaj yada dezavantaj sağlamamıştır. Bu sonuç Kalaycıoğlu ve Kelecioğlu'nun (2011) yapmış oldukları çalışmalarında, 2005 yılı Öğrenci Seçme Sınavında (ÖSS) Türkçe alt testinde maddelerin cinsiyete göre bir yanlılık içermediği sonucu ile benzerlik göstermektedir. Benzer şekilde Yurdugül ve Aşkar'ın Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı verileri ile yapmış oldukları çalışmada, Türkçe alt testi maddelerinin cinsiyete göre DMF içermediği yada ihmal edilebilir düzeyde DMF içerdiği sonucu ile örtüşmektedir (Yurdugül ve Aşkar, 2004).

2012 SBS'ye alt testlerinden bir diğeri olan Matematik alt testine ilişkin 20 madde incelendiğinde madde güçlük düzeyi yüksek ve düşük olan maddelerin mevcut olduğu ancak testin genel olarak orta güçlükte olduğu anlaşılmaktadır. Madde test uyumu kapsamında, teste ilişkin iç ve dış uyum indeksleri ve nokta çift serili korelasyon (PTMEA) değerleri incelenmiştir. Analiz sonuçlarına göre Matematik alt testine ilişkin bir maddenin iç uyum indeksinin, iki maddenin dış uyum indeksinin ve bir maddenin ise PTMEA değerinin beklenen değerden yüksek olduğu anlaşılmıştır. Bu maddelere ilişkin sayısal değerler incelendiğinde sadece bir maddenin düzeltilerek tekrar teste dahil edilmesi gerektiği anlaşılmıştır.

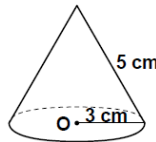
2012 SBS Matematik alt testine ilişkin DMF kontrast değerleri incelenmiştir. DMF kontrast değerlerinin negatif değer alması yanlılığın odak grup(birinci grup) lehine olduğunu göstermektedir. Beklenen DMF değeri sıfır iken mutlak değerce 0.5 değeri kabul edilebilir değerdir. Matematik alt testine ilişkin analiz sonuçları incelendiğinde kimi maddelerin negatif kontrast değer kimilerin ise

pozitif kontrast değer aldığı gözlemlenmektedir. Teste ilişkin analiz sonuçlarına bakıldığında iki maddenin beklenen değer dışında sonuçlandığı anlaşılmıştır. Bu maddelerden birinin negatif değer olarak erkekler lehine çalıştığı diğerinin ise pozitif değer olarak bayanlar lehine çalıştığı anlaşılmıştır.

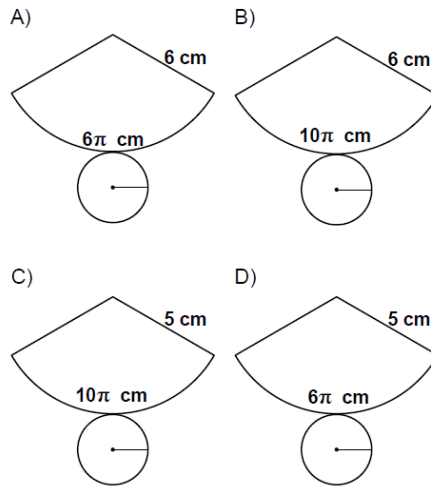
Sonuç olarak Tablo-9 ve Şekil-15 incelendiğinde Matematik alt testine ilişkin maddelerden 4. ve 19. maddelerin 0.54 logit ve -0.56 logit değerlerini alarak, yanlılık gösterebilecekleri; diğer maddelerin ise yanlılık göstermediği tespit edilmiştir.

Aşağıdaki şekillerde Matematik alt testinde DMF içeren maddeler verilmiştir.

4. Şekildeki O noktası, verilen dik dairesel koninin taban merkezidir.



Şekil üzerindeki verilere göre bu koninin açınımı aşağıdakilerden hangisi olabilir?



Şekil-20: Matematik Alt Testi 4.soru

Şekilde-20'de görülen Matematik alt testi 4.sorusunun yapılan istatistiksel işlemler sonucunda 0.54 logit DMF kontrast değeri içerdiği tespit edilmiştir. DMF kontrast değerinin pozitif olması referans grup lehine çalıştığı anlamına gelmektedir. Yani maddenin kız öğrencilere avantaj sağladığını göstermektedir. Yapılan bir çalışmada ilköğretim düzeyinde kız öğrencilerin hem matematik hem de sözel dil

becerilerine dayalı derslerde erkek öğrencilere oranla daha başarılı olduklarını ancak ortaöğretimden itibaren geometri alanında erkeklerin daha başarılı oldukları açıklanmaktadır (Amrein ve Berliner, 2002). Bu bulgu, Zenisky, Hambleton ve Robin'nin (2003) yapmış oldukları çalışma ile çelişmektedir. Zenisky, Hambleton ve Robin'e (2003) göre tablo, şekil, grafik gibi görzel-uzamsal zeka gerektiren maddelerin erkekler lehine çalışmaktadır.

Şekil-20'deki sorunun DMF içerdiği tespit edildikten sonra maddenin yanlıktan mı, madde etkisi(item impact) mi kaynaklandığını tespit etmek için uzman görüşlerine başvurulmuştur. İlgili uzman görüşlerinden sonra maddenin yanlılık içermediği madde etkisinden kaynaklandığı anlaşılmıştır.

19. Bir yarışma programında, verilen her doğru cevaba + 3 puan, her yanlış cevaba – 2 puan verilmektedir.

Bu yarışmaya katılan Aysun, sorulan 5 sorunun tümünü cevaplamıştır. Yarışma sonunda 10 puan aldığına göre, Aysun kaç soruyu doğru cevaplamıştır?

A) 2 B) 3 C) 4 D) 5

Şekil-21: Matematik Alt Testi 4.soru

Şekilde-21'de görülen Matematik alt testi 19.sorusunun yapılan istatistiksel işlemler sonucunda -0.56 logit DMF kontrast değeri içerdiği tespit edilmiştir. DMF kontrast değerinin negatif olması odak grup lehine çalıştığı anlamına gelmektedir. Yani maddenin erkek öğrencilere avantaj sağladığını göstermektedir. Yapılan bir çalışmada ilköğretim düzeyinde kız öğrencilerin hem matematik hem de sözel dil becerilerine dayalı derslerde erkek öğrencilere oranla daha başarılı olduklarını ancak ortaöğretimden itibaren geometri alanında erkeklerin daha başarılı oldukları açıklanmaktadır (Amrein ve Berliner, 2002). Elde edilen bu bulgu, Yurdugül ve Aşkar'ın yapmış olduğu çalışma ile örtüşmektedir. Yurdugül ve Aşkar yapmış oldukları çalışmada matematik sorularının sözel olarak ifade edilen sorularının erkekler lehine sonuçlandığı saptanmıştır(Yurdugül ve Aşkar, 2004). Willingham ve

Cole (1997), matematik alanında yaptıkları araştırmada, özellikle matematiksel işlemler ve hesaplamalar alanında kız öğrencilerin daha başarılı olduklarını, problem çözmeye dayalı performanslarda ise erkek öğrencilerin daha başarılı olduklarını ifade etmişlerdir. Şekil-21'deki maddenin problem çözüme becerisi ile ilgili bir madde olmasından dolayı bu bulgu Willingham ve Cole'un bulgusu ile benzerlik göstermektedir.

Şekil-21'deki sorunun DMF içerdiği tespit edildikten sonra maddenin yanlılıktan mı, madde etkisi(item impact) mi kaynaklandığını tespit etmek için uzman görüşlerine başvurulmuştur. İlgili uzman görüşlerinden sonra maddenin yanlılık içermediği madde etkisinden kaynaklandığı anlaşılmıştır.

20 maddelik Fen ve Teknoloji alt testi incelendiğinde madde güçlük düzeyi yüksek ve düşük olan maddelerin mevcut olduğu ancak testin genel olarak orta güçlükte olduğu anlaşılmaktadır. Madde test uyumu kapsamında, teste ilişkin iç ve dış uyum indeksleri ve nokta çift serili korelasyon (PTMEA) değerleri incelenmiştir. Analiz sonuçlarına göre teste ilişkin bütün madde sayısal verilerinin istenilen değer aralığına uygun olduğu ancak bir maddenin dış uyum indeksinin istenilen değerden yüksek olduğu anlaşılmıştır. Ancak bu maddeye ilişkin diğer sonuçlar incelendiğinde testin yapısını bozmadığı ve teste ilişkin bütün maddelerin testin genel yapısına uygun olduğu anlaşılmaktadır.

2012 SBS Fen ve Teknoloji alt testine ilişkin DMF kontrast değerleri incelenmiştir. DMF kontrast değerlerinin negatif değer alması yanlılığın odak grup(birinci grup) lehine olduğunu göstermektedir. Beklenen DMF değeri sıfır iken mutlak değerce 0.5 değeri kabul edilebilir değerdir. Fen ve Teknoloji alt testine ilişkin analiz sonuçları incelendiğinde kimi maddelerin negatif kontrast değer kimilerinin ise pozitif kontrast değer aldığı gözlemlenmesine karşın maddelerin tümünün beklenen aralıkta değer aldığı anlaşılmıştır.

Sonuç olarak, Tablo-11 ve Şekil-17'de Fen ve Teknoloji alt testine ilişkin cinsiyet bakımından DMF değerleri incelendiğinde, bütün maddelerin -0,50 ile 0,50 arasında değer aldığı tespit edilmiştir. Yani Fen Bilgisi testine ilişkin cinsiyet açısından bir yanlılığın olmadığı ortaya çıkmıştır. Elde edilen bu bulgu, Gök ve arkadaşlarının 2010'da değişen madde fonksiyonunu belirlemede kullanılan Mantel-Haenszel (MH) ve lojistik regresyon (LR) tekniklerini karşılaştırmak amacıyla

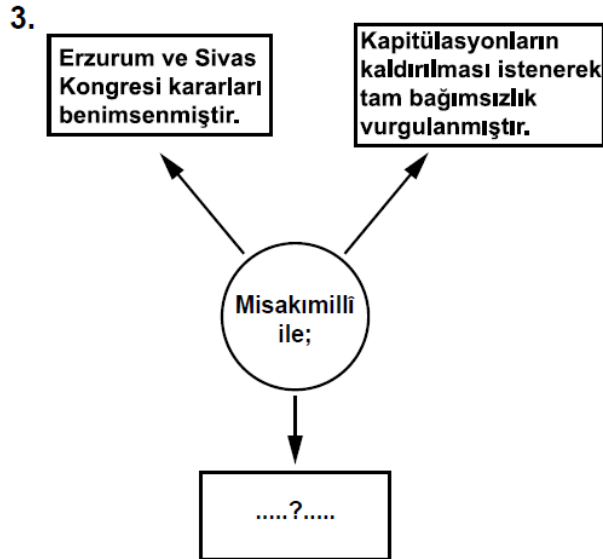
Ortaöğretim Kurumları Sınavı (OKS) verileri ile yaptıkları çalışmada Fen Bilgisi alt testinde cinsiyete göre DMF veren maddelere ilişkin MH tekniğine ait sonuçlar ile örtüşmektedir. Gök ve arkadaşları 2005 OKS sınavı Fen bilgisi alt testinde cinsiyete göre maddelerin tamamı ihmal edilebilir düzeyde yada DMF içermediğini tespit etmişlerdir(Gök ve Ark., 2010). Ayrıca Doğan ve Öğretmen değişen madde fonksiyonlarını tespit etmede kullanılan Ki-kare, Mantel-Haenszel ve Lojistik Regresyon tekniklerini karşılaştırarak uygulamada ortaya çıkan benzerlik ve farklılıkları belirlemek amacıyla 2003 yılı Ortaöğretim Kurumları Seçme ve Yerleştirme (OKÖSYS) sınav verilerini inceleyerek yapmış oldukları çalışmada da Fen Bilgisi alt testinde DMF içeren maddeye rastlamamışlardır (Doğan ve Öğretmen, 2008).

20 maddelik Sosyal Bilgiler alt testi incelendiğinde madde güçlük düzeyi yüksek ve düşük olan maddelerin mevcut olduğu ancak testin genel olarak orta güçlükte olduğu anlaşılmaktadır. Madde test uyumu kapsamında, teste ilişkin iç ve dış uyum indeksleri ve nokta çift serili korelasyon (PTMEA) değerleri incelenmiştir. Analiz sonuçlarına göre teste ilişkin bütün madde sayısal verilerinin istenilen değer aralığına uygun olduğu ancak iki maddeye ilişkin dış uyum indeksi ve PTMEA değerlerinin istenilen değer aralığında olmadığı anlaşılmıştır. Bu maddelerin gözden geçirilip teste dahil edilmesi anlaşılmıştır.

2012 SBS Sosyal Bilgiler alt testine ilişkin DMF kontrast değerleri incelenmiştir. DMF kontrast değerlerinin negatif değer alması yanlılığın odak grup(birinci grup) lehine olduğunu göstermektedir. Beklenen DMF değeri sıfır iken mutlak değerce 0.5 değeri kabul edilebilir değerdir. Sosyal Bilgiler alt testine ilişkin analiz sonuçları incelendiğinde kimi maddelerin negatif kontrast değer kimilerin ise pozitif kontrast değer aldığı gözlemlenmiştir. Analiz sonuçlarına göre Sosyal Bilgiler alt testine ilişkin bir maddenin DMF kontrast değeri beklenen değer dışındadır. Elde edilen bulguya göre bu maddenin erkeklere yanlılık sağlayabileceği anlaşılmıştır.

Sonuç olarak, Tablo-13 ve Şekil-19'da incelendiğinde Sosyal Bilgiler testine ilişkin maddelerden 3. madde -0.62 logit değeri aldığından yanlılık göstermiş olabildiği tespit edilirken diğer maddelerin yanlılık göstermediği tespit edilmiştir.

Aşağıdaki şekilde Sosyal Bilgiler alt testinde DMF içeren madde verilmiştir.



Misakımillî'nin önemi düşünüldüğünde şemada "?" ile gösterilen yere aşağıdakilerden hangisi getirilebilir?

- A) Türk vatanının sınırları çizilmiştir.
- B) Güçler birliği esası benimsenmiştir.
- C) Millî ekonominin ilkeleri saptanmıştır.
- D) İlk kez millî egemenlikten söz edilmiştir.

Şekil-22: Sosyal Bilgiler Alt Testi 3.soru

Şekilde-22'de görülen Sosyal Bilgiler alt testi 3.sorusunun yapılan istatistiksel işlemler sonucunda -0.62 logit DMF kontrast değeri içerdiği tespit edilmiştir. Maddenin negatif DMF kontrast değer alması erkekler lehine çalıştığı anlamına gelmektedir. Sorunun savaş ile ilgili bir konu olmasından dolayı erkekler lehine çalıştığı düşünülmektedir. Bu sonuç Kalaycıoğlu ve Kelecioğlu'nun 2005-ÖSS'de yer alan maddelerin, DMF içerip içermediği ile ilgili yaptıkları çalışmada, Sosyal Bilgiler testinde tarihle ilgili bir maddenin DMF içeren bulgusu ile örtüşmektedir. Araştırmacılar Tarih testinin siyaset ve savaş ile ilgili sorularında erkekler lehine çıkabileceğini vurgulamıştır (Kalaycıoğlu ve Kelecioğlu, 2011). Benzer olarak Zwick ve Ercikan'ın yapmış oldukları bir çalışmada, Tarih testi bazı maddelerinin erkekler lehine DMF içerdiği sonucuna varmışlardır(Zwick ve Ercikan, 1989). Le'ni (1999) yapmış olduğu çalışmada tarih sorularının erkek lehine bir sonuç gösterdiğini tespit etmiştir. Elde edilen bu sonuç Le'nin bulgusu ile örtüşmektedir.

Şekil-22'deki sorunun DMF içerdiği tespit edildikten sonra maddenin yanlıktan mı, madde etkisinden (item impact) mi kaynaklandığını tespit etmek için uzman görüşlerine başvurulmuştur. İlgili uzman görüşlerinden sonra maddenin yanlılık içermediği madde etkisinden kaynaklandığı anlaşılmıştır.

BEŞİNCİ BÖLÜM SONUÇ VE ÖNERİLER

5.1.SONUÇ

Sonuç olarak bu çalışmada, 2012-Seviye Belirleme Sınavı alt testlerinden Türkçe, Matematik, Fen ve Teknoloji ve Sosyal Bilgiler testlerinin maddeleri cinsiyet yanlılığı bakımından incelenmiştir. Yapılan istatistiksel işlemler sonucunda iki Matematik alt testi maddesi ve bir Sosyal Bilgiler alt testi maddesi olmak üzere toplam üç maddenin Değişen Madde Fonksiyonu (DMF) içerdiği tespit edilmiştir. DMF tespit edilen maddelerin yanlılıktan mı madde etkisinden mi kaynaklandığını tespit etmek için uzman görüşüne başvurulmuştur. Uzman görüş doğrultusunda maddelerin yanlılıktan kaynaklanmadığı, madde etkisinden kaynaklandığı anlaşılmıştır. Bu bağlamda 2012 Seviye Belirleme Sınavı verilerinin cinsiyet yanlılığı bakımından geçerli ve güvenilir bir test olduğu tespit edilmiştir.

5.2.ÖNERİLER

- Bu çalışmada 2012 Seviye Belirleme Sınavı Türkçe, Matematik, Fen Bilgisi ve Sosyal Bilgiler alt testinin cinsiyet açısından yanlılığı incelenmiştir. Ulusal sınavların, bölgesel yanlılıkları incelenmesi faydalı görünmektedir.
- Yüksek öğrenim giriş sınavlarında öğrencilerin okul türü yada mezuniyet durumlarına göre yanlılık çalışmalarının yapılması önemli bir katkı sağlayabilecektir.
- İlgili alan yazın incelendiğinde DMF belirlemede gerek Klasik Tepki Kuramına göre, gerek Madde Tepki Kuramına göre birçok yöntem kullanılmıştır. Alan yazında DMF belirlemede birçok yöntem karşılaştırılması da yapılmıştır. Bu çalışmada Rasch modeli ile DMF içeren maddeler tespit edilmiştir. Rasch modeli ile diğer DMF belirleme yöntemlerinin karşılaştırılması yararlı olacaktır.
- Ulusal düzeyde yapılan ortaöğretime veya yüksek öğrenime giriş sınavlarına dair geçmişte yapılan yanlılık çalışmalarının karşılaştırılıp sınavlardaki yanlılık oranının farklılaşma oranının araştırılması ilgili alana katkı sağlayabilir.

- Farklı örneklem büyüklükleri ile yanlılık belirlenip, örneklem büyüğünün yöntem üzerindeki etkisi test edilebilir.

KAYNAKLAR

- Agresti, A. (1984). *Analysis of ordinal categorical data*. New York: John Wiley & Sons.
- Amrein, A. L. ve Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10, 18. 05.04.2013 tarihinde, <http://epaa.asu.edu/epaa/v10n18/> adresinden alınmıştır.
- Ariffin, S. R., Idris, R. ve Ishak, N. M. (2010). Differential item functioning in Malaysian generic skills instrument (MyGSI). *Jurnal Pendidikan Malaysia*, 35(1), 1-10.
- Baykul, Y. (1997). *İstatistik Metotlar ve Uygulamalar*. Ankara: Anı Yayıncılık, 2.Baskı.
- Baykul, Y. (2010). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*.2.Baskı Pegem.
- Bekçi, B. (2007). Ortaöğretim Kurumları Öğrenci Seçme Sınavının Değişen Madde Fonksiyonlarının Cinsiyete ve Okul Türüne Göre İncelenmesi. *Yayınlanmamış Yüksek lisans Tezi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 31, 419-456.
- Bentler, P. M., ve Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bond, T.G. ve Fox, C.M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. New Jersey: Lawrence Erlbaum Associates Publishers London.
- Bond, T.G. ve Fox, C.M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Second Edition*. New Jersey: Lawrence Erlbaum Associates Publishers London.
- Browne, M. W., & Cudeck, R. (1993). Alternativeways of assessing model fit. In K. A. Bollen & J. S.Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park. CA: Sage.
- Büyüköztürk, Ş. (2002). *Sosyal bilimler için veri analizi el kitabı*. Ankara: PegemA Yayıncılık.

- Byrne, B. M., ve Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-Cultural Psychology*, 30, 555.
- Camilli, G. ve Shepard L. A. (1994). *Methods for identifying biased test items*. London: Sage Publications
- Cleary, T. ve Hilton, T. L. (1968). An investigation of item bias. *Educational and Psychological Measurement*.
- Cole, N. S. ve Zieky, M. J. (2001). The new faces of fairness. *Journal of Educational Measurement*, 38(4), 369-382.
- Crooks, T. J.(1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 438-481
- Demircioğlu, G. (2009). Geçerlik ve Güvenirlik. *Ölçme ve Değerlendirme*. Karip. E. (Ed.) Pegem Yayıncılık, Ankara.
- Deveci (Ateşok), N. (2008). Üniversitelerarası Kurul Yabancı Dil Sınavının Madde Yanlılığı Bakımından İncelenmesi. *Yayınlanmamış Doktora Tezi*. Eğitim Bilimleri Enstitüsü, Eğitimde Psikolojik Hizmetler Anabilim Dalı, Ankara Üniversitesi.
- Doğan, N., ve Öğretmen, T. (2010). Değişen Madde Fonksiyonunu Belirlemede Mantel-Haenszel, Ki-Kare ve Lojistik Regresyon Tekniklerinin Karşılaştırılması. *Eğitim ve Bilim*, 33(148), 100-112.
- Gamer, M. ve Engelhard Jr, G. (1999). Gender differences in performance on multiple-choice and constructed response mathematics items. *Applied Measurement in Education*, 12(1), 29-51.
- Gelbal, S. (1994). P madde güçlük indeksi ile Rasch modelinin b parametresi ve bunlara dayalı yetenek ölçüleri üzerine bir karşılaştırma. *Yayınlanmamış Doktora Tezi*, Eğitim Bilimleri Bölümü, Hacettepe Üniversitesi, Ankara.
- Gierl, M., Khaliq, S. N. ve Boughton, K. (1999). Gender differential item functioning in mathematics and science: Prevalence and policy implications. In *Annual Meeting of the Canadian Society for the Study of Education*, Sherbrooke.
- Gök, B., Kelecioğlu, H., ve Doğan, N. (2010). Değişen madde fonksiyonunu belirlemede Mantel-Haenszel ve Lojistik Regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 35(156), 3-16.

- Güler, N. (2008). *Klasik Test Kuramı, Genellenebilirlik Kuramı ve Rasch Modeli Üzerine Bir Araştırma*. Yayınlanmamış doktora tezi. Hacettepe Üniversitesi, Ankara.
- Hambleton, R. K., Swaminathan, H. ve Rogers H.J. (1991). *Fundamentals of Item Response Theory*. Sega.
- Hambleton, R. K., ve Rodgers, J. (1995). *Item bias review*. ERIC Clearinghouse on Assessment and Evaluation, the Catholic University of America, Department of Education.
- Hambleton, R.K. ve Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer Academic Publishers.
- Hanna, G. (1986). "Sex Differences in The Mathematics Achievement Of 8th Graders in Ontario." *Journal for Research in Mathematic Education* Vol. 17, p. 231-237
- Kalaycıoğlu, D. B. ve Kelecioğlu, H. (2011). Öğrenci Seçme Sınavı'nın Madde Yanlılığı Açısından İncelenmesi. *Eğitim ve Bilim*, 36(161), 3-13.
- Kan, Adnan, Hakan Atılğan ve Nuri Doğan (2006). *Eğitimde Ölçme ve Değerlendirme*, 3. baskı, Anı yayıncılık, Ankara.
- Kaptan, F.(1994). Rasch Modeli Madde Parametrelerini Kullanarak En Yüksek Olabilirlik Yöntemiyle Yetenegin Kestirilmesi, *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 10, 95-97
- Karami, H. (2011). An Introduction to Differential Item Functioning. *The International Journal of Educational and Psychological Assessment* September 2012, Vol. 11(2)
- Karasar, N. (2008). Bilimsel Araştırma Yöntemleri. *Ankara, Nobel Yayıncılık*.
- Kristjansson, E., Aylesworth, R., Mcdowell, I., ve Zumbo, B. D. (2005). A comparison of four methods for detecting differential item functioning in ordered response items. *Educational and Psychological Measurement*, 65(6), 935-953.
- Kurnaz, B. (2006). Peabody Resim Kelime Testinin Madde Yanlılığı Açısından İncelenmesi. *Yayınlanmamış Yüksek Lisans Tezi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Lai, J.S. ve Eton, D.T. (2002). *Clinically Meaningful Gaps*. *Rasch Measurement Transactions* 15(4): 850.

- Le, V. N. (1999). *Identifying Differential Item Functioning on the NELS: 88 History Achievement Test*. Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California, Los Angeles.
- Linacre, J. M. (2009). FACETS Rasch-model computer program (Version 3.66.0) [Computer software]. Chicago, IL: Winsteps.com.
- Linacre, J. M. (2010) Winsteps®(Version 3.70.0) [Computer Software]. Beaverton, Oregon:Winsteps.com.
- Linacre, J. M. ve Wright, B. D. (1987). *Item bias: Mantel-Haenszel and the Rasch model* (Memorandum No. 39). Chicago: MESA Psychometric Laboratory.
- Linacre, J.M. (2002). *Differential item and test functioning*. <http://www.rasch.org/rmt/rmt163g.htm>. adresinden 17.07.2013 tarihinde alınmıştır.
- Linacre, J.M. (2007). *A User's Guide to WINSTEPS Rasch-model Computer Programs*. Chicago: MESA Press.
- Lord, F.M. (1980). *Application of item response theory to practical testing problem*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Marsh, H. W., Hau, K.T., Artelt, C., Baumert, J., ve Peschar, J. L. (2006). OECD' s brief self-report measure of educational psychology' s most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, 6(4), 311-360.
- MEB. (2011). ““Türk Milli Eğitim Sisteminin Örgütlenmesi-2011” (Türkçe) (Ağ sayfası, metin, .html). <http://sgb.meb.gov.tr/eurydice/index.htm>” 11.10.2013 tarihinde adresinden alınmıştır. Milli Eğitim Bakanlığı Strateji Geliştirme Başkanlığı. Türkiye.
- MEB. (2012). SBS Başvuru Klavuzu, <http://oges.meb.gov.tr/> adresinden 10.06.2013 tarihinde alınmıştır.
- Mellenbergh, G. J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics*, 7, 105–118.
- Micheels, W. J. ve Karnes, M. R. (1950). *Measuring educational achievement*. McGraw-Hill.

- Nandakumar, R. (1993). A fortran 77 program for detecting differential item functioning through the mantel-haenszel statistic. *Educational and Psychological Measurement*, 53, 679–684.
- Osterlind, S. J. (Ed.). (1983). *Test item bias* (Vol. 30). Sage
- Özbek, Ö.Y. (2008). Ölçme Araçlarında Bulunması İstenen Nitelikler içinde S. Tekindal (Ed.). *Eğitimde Ölçme ve Değerlendirme*, 49-90.
- Özer (Özkan), Y. (2012). Öğrenci Başarılarının Belirlenmesi Sınavından (ÖBBS) Klasik Test Kuramı, Tek Boyutlu Ve Çok Boyutlu Madde Tepki Kuramı Modelleri İle Kestirilen Başarı Puanlarının Karşılaştırılması. *Yayınlanmamış Doktora Tezi*. Eğitim Bilimleri Enstitüsü Ölçme Ve Değerlendirme Anabilim Dalı, Ankara Üniversitesi.
- Pallant, J.F. ve Tennant, A. (2007). *An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS)*. *Br J Clin Psychol* 46: 1-18.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability* (Vol. 4, pp. 321-333). Berkeley, CA: University of California Press.
- Roeber, C. (2005). "That's not fair!" *Fairness, bias, and differential item functioning in language testing*. 18.05.2013 tarihinde Hawai'i Üniversitesi Sisteminden alınmıştır. Web site: <http://www2.hawaii.edu/~roever/brownbag.pdf>
- Schumacker, R. E. (2005). Test Bias and Differential İtem Functioning. *Applied Measurement Associates*, 1-2.
- Semerci, Ç. (2009). Eğitimde Ölçme ve Değerlendirme. *Ölçme ve Değerlendirme*. Karip. E. (Ed.) Pegem Yayıncılık, Ankara.
- Smith, R.M. (2000). Fit analysis in latent trait measurement models. *Journal of Applied Measurement* 1(2): 199-218.
- Streiner, D. ve Norman, G. (1995). *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford University Press, Oxford.
- Sümbüloğlu, K. ve Sümbüloğlu, V. (1993). *Biyoistatistik*. Ankara: Özdemir Yayıncılık, 4. Baskı.

- Swaminathan, H. ve Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational measurement*, 27(4), 361-370.
- Taylor, C. S. ve Lee, Y. (2012): Gender DIF in Reading and Mathematics Tests With Mixed Item Formats. *Applied Measurement in Education*, 25:3, 246-280
- Tekin, H. (2000). *Eğitimde Ölçme ve Değerlendirme*(16. Baskı). Ankara:Yargı Yayınevi.
- Turgut, M.F. (1995). *Eğitimde Ölçme Değerlendirme Metotları*. Yargıcı Yayıncılık, 10.Baskı, Ankara
- Turgut, M.F. ve Baykul, Y. (2010). *Eğitimde Ölçme Değerlendirme*. Pegem Yayıncılık, Ankara
- Wang, W. C. ve Chen, C. T. (2005). Item parameter recovery, standard error estimates, and fit statistics of the WINSTEPS program for the family of Rasch models. *Educational and Psychological Measurement*, 65(3), 376-404.
- Willingham, W. W. ve Cole, N. S. (Eds.). (1997). *Gender and fair assessment*. Psychology Press.
- Wu, M. L., Adams, R. J., Wilson, M. R. ve Haldane, S. A. (2007). ACER ConQuest Version 2: Generalized item response modeling software [computer program]. Camberwell: *Australian Council for Educational Research*.
- Yılmaz, A. (2009). Ölçme-Değerlendirmede Testler. *Ölçme ve Değerlendirme*. Karip. E. (Ed.) Pegem Yayıncılık, Ankara.
- Yılmaz, H. (1998). *Eğitimde ölçme ve değerlendirme*. Mikro Yayınevi, Konya.
- Yurdugül, H. (2003). *Ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavının madde yanlılığı açısından incelenmesi*, Yayınlanmamış Doktora tezi, Hacettepe Üniversitesi, Ankara
- Yurdugül, H. ve Aşkar, P. (2004). Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı'nın cinsiyete göre madde yanlılığı açısından incelenmesi.*Eğitim Bilimleri ve Uygulama*, 3(5), 3-20.
- Zenisky, A. L., Hambleton, R. K., ve Robin, F. (2003). Detection of differential item functioning in large-scale state assessments: A study evaluating a two-stage approach. *Educational and Psychological Measurement*, 63(1), 51-64.
- Zenisky, A. L., Hambleton, R. K., ve Robin, F. (2004). DIF detection and interpretation in large-scale science assessments: Informing item writing practices. *Educational Assessment*, 9(1-2), 61-78.

- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland, and H. Wainer (Eds.), *Differential item functioning* (pp. 337-347). Hillsdale, NJ: Lawrence Erlbaum.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- Zwick, R. ve Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of Educational Measurement*, 26(1), 55-66.

EKLER**Ek-1:Verilerin Milli Eğitim Bakanlıđından Alınmasına Dair Dilekçe**

Sayı : 62927161-300
 Konu : Y.Lisans Tez Çalışması Verileri

21/5/2013

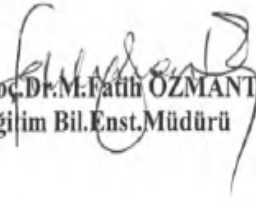
T.C.
MİLLİ EĐİTİM BAKANLIĐI
 Eğitim Teknolojisi Genel Müdürlüğü


ANKARA

Gaziantep Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitim Bilimleri Anabilim Dalı Yüksek Lisans öğrencisi Azmi TÜRKAN'a "2012 Seviye Belirleme Sınavı Sorularının Test Yakınlığı Bakımından İncelenmesi" başlıklı tezinde kullanmak üzere, 2012 yılında yapılan 8.Sınıf Seviye Belirleme Sınavı test verileri gerekmektedir.

Araştırmayı gerçekleştirebilmesi için 8.sınıf düzeyindeki öğrencilerin cinsiyetleri, sınava girdikleri şehirleri, SBS test kitapçıklarına ait kitapçık türü, kitapçıklardaki madde formatı, madde formatlarının başladığı ve bittiği madde numaraları, kitapçık türüne göre maddelerin doğru cevap anahtarları ile yanıtlayıcıların sorulara verdikleri cevaplara ihtiyaç duyulmaktadır.

Tez çalışmasında kullanılacak verilerin Eğitim bilimleri Enstitüsüne gönderilebilmesi hususunda gereğini rica ederim.


 Doç.Dr.M.Fatih ÖZMANTAR
 Eğitim Bil.Enst.Müdürü


 Prof.Dr.M.Yavuz COŞKUN
 Gaziantep Üniversitesi Rektörü

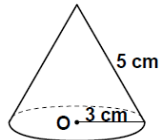
Ek-2: Yanlılığa İlişkin Uzman Görüşü

Uzman Görüş Formu

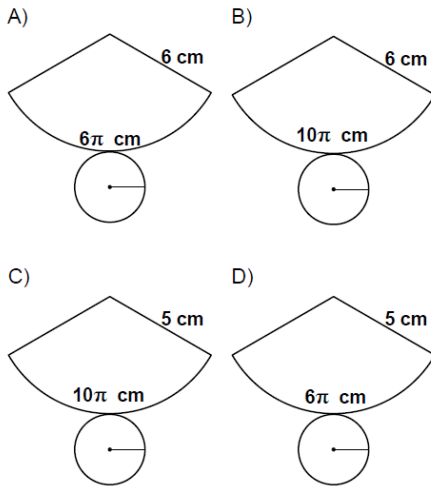
Yanlılık, farklı alt gruplardan gelen aynı yetenek düzeyine sahip bireylerin testin özelliklerinden ya da testin temel amacına uygun olmayan koşullardan dolayı herhangi bir grup lehine ya da aleyhine farklılaşmasıdır. Yapılan istatistiksel işlemler sonucunda aynı yetenek düzeyine sahip katılımcıların aşağıdaki 2 Matematik ve 1 Sosyal Bilgiler sorusunun cinsiyete açısından bir grup lehine sonuçlandığı tespit edilmiştir. Elde edilen bu bulguların yanlılık mı yoksa gerçek farktan kaynaklandığını tespit etmek amacıyla görüşünüze başvurulmuştur.

Azmi TÜRKAN

4. Şekildeki O noktası, verilen dik dairesel koninin taban merkezidir.



Şekil üzerindeki verilere göre bu koninin açınımı aşağıdakilerden hangisi olabilir?



Sizce madde hangi grup lehine çalışmıştır?

() Erkek () Bayan

Maddenin tercih ettiğiniz grup lehine çalışma nedeni aşağıdakilerden hangisi olabilir? Neden?

A) Maddede kullanılan şekilden. (.....)

- B) Maddede kullanılan sözel ifadeden. (.....)
- C) Maddenin bir cinsiyet grubuna aşına olmasından (.....)
- D) Maddenin cinsiyete bağlı olmaksızın, sadece madde etkisinden (gerçek farktan) kaynaklandığını düşünüyorum. (.....)

19. Bir yarışma programında, verilen her doğru cevaba + 3 puan, her yanlış cevaba – 2 puan verilmektedir.

Bu yarışmaya katılan Aysun, sorulan 5 sorunun tümünü cevaplamıştır. Yarışma sonunda 10 puan aldığına göre, Aysun kaç soruyu doğru cevaplamıştır?

- A) 2 B) 3 C) 4 D) 5

Sizce madde hangi grup lehine çalışmıştır?

- () Erkek () Bayan

Maddenin tercih ettiğiniz grup lehine çalışma nedeni aşağıdakilerden hangisi olabilir? Neden?

- A) Maddede kullanılan sözel ifadeden. (.....)
- B) Maddenin bir cinsiyet grubuna aşına olmasından (.....)
- C) Maddenin cinsiyete bağlı olmaksızın, sadece madde etkisinden (gerçek farktan) kaynaklandığını düşünüyorum. (.....)

3.



Misakımillî'nin önemi düşünüldüğünde şemada “?” ile gösterilen yere aşağıdakilerden hangisi getirilebilir?

- A) Türk vatanının sınırları çizilmiştir.
- B) Güçler birliği esası benimsenmiştir.
- C) Millî ekonominin ilkeleri saptanmıştır.
- D) İlk kez millî egemenlikten söz edilmiştir.

Sizce madde hangi grup lehine çalışmıştır?

() Erkek () Bayan

Maddenin tercih ettiğiniz grup lehine çalışma nedeni aşağıdakilerden hangisi olabilir? Neden?

- A) Maddede kullanılan şekilden. (.....)
- B) Maddede kullanılan sözel ifadeden. (.....)
- C) Maddenin bir cinsiyet grubuna aşına olmasından (.....)
- D) Maddenin cinsiyete bağlı olmaksızın, sadece madde etkisinden (gerçek farktan) kaynaklandığını düşünüyorum. (.....)

ÖZGEÇMİŞ

Azmi TÜRKAN 1984 yılında Diyarbakır'ın Dicle ilçesinde doğdu. İlköğrenimini Dicle'de bitiren araştırmacı ortaöğrenimini Diyarbakır Yunus Emre Lisesinde tamamlamıştır. Lisans eğitimini Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi İlköğretim Matematik Öğretmenliğinde 2011 yılında tamamlayan araştırmacı aynı yıl içersisinde Siirt Üniversitesinde Araştırma Görevlisi olarak başlamıştır. Araştırmacı Siirt Üniversitesinde hala Araştırma Görevlisi olarak görevine devam etmektedir.

VITAE

Azmi TÜRKAN was born in Diyarbakır in 1984. The author finished primary school in Dicle and attended Yunus Emre High School in Diyarbakır. He graduated from Dicle University, Ziya Gökalp Faculty of Education Department of Primary Mathematics Teaching in 2011. He started to work as a research assistant at Siirt University in the same year. The author is still holding the same position at Siirt University.