

## Temel Eğitimden Ortaöğretime Geçiş Sınavının (2014-Kasım) Cinsiyet Açısından Madde Yanlılığının İncelenmesi

### Investigation of Item Bias in 2014- November Transition Exam from Primary to Secondary Education In Terms Of Gender\*

Duriye Kübra SATICI\*\*, Yeşim ÖZER ÖZKAN\*\*\*

**Öz:** Eğitimde kullanılan testlerin ve maddelerinin, aynı yetenek düzeyinde olan fakat farklı gruplardan gelen bireylerin cevapları açısından farklı özellikler göstermemesi beklenmektedir. Bu gereklilikten hareketle araştırmada; TEOG sınavının Türkçe, matematik, fen ve teknoloji ve T.C. İnkılâp Tarihi ve Atatürkçülük testlerinin cinsiyete göre Değişen Madde Fonksiyonu (DMF) içerip içermediğinin belirlenmesi amaçlanmıştır. Var olan bir durumun var olduğu biçimiyle, herhangi bir değişiklik yapmaksızın betimlenmesi amaçlandığından araştırma, tarama modelinde betimsel türde desenlenmiştir. Araştırmanın evreni, 2014 yılı Kasım ayında uygulanan TEOG sınavına giren bütün öğrencilerdir. Araştırmanın örneklemini ise sınava giren öğrenciler arasından basit seçkisiz örnekleme yöntemi ile seçilmiş 40.000 öğrenci oluşturmaktadır. Örneklemdaki öğrencilerin %52,14'ü erkek ve % 47,86'sı kızdır. Maddelerin DMF içerip içermediği Winsteps paket programı aracılığıyla Rasch modeli ile tespit edilmiştir. Maddelerin DMF tespitine geçilmeden önce yapılan analiz işlemlerinin Rasch modelinin sayıltılarını karşılayabilmesi için Türkçe, matematik, fen ve teknoloji ve T.C. İnkılâp Tarihi ve Atatürkçülük testlerinin tek boyutluluk ve test maddelerinin yerel bağımsızlık özelliklerini sağlayıp sağlamadığı incelenmiştir. Tek boyutlu olduğu görülen Türkçe alt testinde 12, matematik alt testinde 20, fen ve teknoloji alt testinde 18 ve T.C. İnkılâp Tarihi ve Atatürkçülük alt testinde 14 madde olmak üzere toplam 64 madde için DMF analizleri yapılmıştır. Testler için yapılan analizlerde analize dahil edilen hiçbir alt test maddesinde DMF için uygun kontrast değer bulunmamıştır. Diğer bir anlatımla testlerde cinsiyet bakımından farklı işleyen herhangi bir maddeye rastlanmamıştır. Bu sonuç sınavın analize dahil edilen maddeler çerçevesinde cinsiyet açısından farklı grup öğrencilerin lehine ya da aleyhine çalışan maddenin olmadığına göstergesidir.

**Anahtar Kelimeler:** TEOG, Rasch analizi, madde yanlılığı, değişen madde fonksiyonu

**Abstract:** Tests and items are expected not to show different characteristics in terms of the answers of individuals who are at the same ability level but from different groups. The objective of this study was to indicate whether Turkish, math, science and technology and Turkish Republic History and Atatürk's courses put forth any difference due to the gender of the individuals or not in TEOG examination. The study was based on a descriptive model, as the objective was to describe a situation as it exists. The research includes all students who took part in the TEOG examination in November in 2014. The sample of the research consists of 40000 students that were chosen via random sampling method: 52.14% percentage of which consists of male and 47.86 of female. It was determined whether the items contains Differential Item Function (DIF) or not using the Rasch model and via the Winstep package software. Prior to determining DIF identification, it was examined whether Turkish, math, Revolution History and Kemalism and science and technology courses provide unidimensionality and local independency or not in order that the analysis processes can fulfill the requirements of the Rasch model. Following the acceptance of unidimensionality 12 items in Turkish subtests, 20 in math subtests, 18 in science and technology subtests and 14 in Revolution History and Kemalism subtest were subject to DIF analysis making up a total of 64 items. In the analysis for the subtests, no contrast value, that is suitable for the DIF, could be found. In other words, there couldn't be found any item that shows any different functions

\*Bu çalışma ilk yazarın yüksek lisans tezinden üretilmiştir

\*\*Öğretmen, MEB, Kahramanmaraş-Türkiye, e-posta: [kubra.satıcı@gmail.com](mailto:kubra.satıcı@gmail.com)

\*\*\*Yrd. Doç. Dr., Gaziantep Üniversitesi, Gaziantep Eğitim Fakültesi, Gaziantep-Türkiye, e-posta: [yozer80@gmail.com](mailto:yozer80@gmail.com)

in terms of gender in subtests.. This result indicates that there is no item that works for the behalf of or to the detriment of students from different gender groups.

*Keywords:* TEOG, Rasch analysis, item bias, differential item function

## Giriş

Türkiye’de ilköğretimden ortaöğretime, ortaöğretimden yükseköğretime geçişlerde çeşitli sınavlar uygulanmaktadır. Sınavların uygulanma amacına bağlı olarak alınan puanlar bireyler hakkında bazı kararlar almaya dayanak olur (Baykul, 2000). Bu sınavlar sonucunda öğrenciler uygun olan bir üst kuruma yerleşebilmekte veya yerleşebilmek umuduyla tekrar şansını denemektedir. Testlerin ortaöğretim kurumlarına ve yükseköğretim programlarına girişte kullanılması, sonuçları itibari ile gelecekte iş ve meslek yaşamını belirleyecek olması, test sonuçlarına göre verilecek kararların önemini daha da arttırmıştır. Bu derece önemli olan sınavlarda kullanılan ölçme araçlarının ölçme kriterlerine göre iyi analiz edilip değerlendirilmesi gerekmektedir. Küçük hataların büyük etkiler doğurabileceği sınavlar üzerinde çalışmaların yoğunlaşma nedeni de bu olsa gerekir. Günümüzde ilköğretimden sonra Anadolu, Fen ve Meslek Liseleri ile Özel liselerde okuyacak öğrencileri belirleme ve yerleştirme işlemi, Milli Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü (MEB- ODSGM) tarafından gerçekleştirilmektedir.

İlköğretimden ortaöğretime öğrenci seçme ve yerleştirme süreci tarihsel olarak incelendiğinde 1998-2008 yılları arasında, Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı (OKÖSYS) adıyla uygulandığı görülmektedir (Karakaya ve Kutlu, 2012). 2008 yılında MEB, ilköğretim programlarındaki değişikliği gerekçe göstererek, sınavı Seviye Belirleme Sınavı (SBS) adı altında uygulamaya başlamıştır (MEB, 2011). SBS, 2008 yılında 6. ve 7. sınıflarda, 2009 yılından 2011 yılına kadar 6., 7. ve 8. sınıflarda uygulanmıştır. Daha sonra 6.sınıftaki uygulamalardan vazgeçilmiş ve 2011 yılında 7. ve 8. sınıf; 2012 yılında ise sadece 8. sınıf öğrencilerine uygulanmıştır (MEB, 2011; MEB, 2012).

Gerek daha önceki uygulamaların beklenen sonuçları vermemesi, gerek 2010 yılından itibaren genel liselerin Anadolu liselerine dönüştürülmesi ve 2013-2014 eğitim öğretim yılında 4+4+4 eğitim sistemi düzenlemesiyle zorunlu eğitimin 12 yıla çıkarılması gibi adımlar, ortaöğretime geçiş sistemi tartışmalarını artırmıştır (Görmez ve Coşkun, 2015). MEB Eylül 2013’te yaptığı açıklamayla, 2013-2014 öğretim yılından itibaren uygulanacak yeni ortaöğretim geçiş sistemini, kamuoyuyla paylaşmıştır. Temel Eğitimden Ortaöğretime geçişteki bu sisteme göre altı temel ders için (Türkçe, matematik, fen ve teknoloji, T.C. İnkılap Tarihi ve Atatürkçülük, yabancı dil, din kültürü ve ahlak bilgisi) sekizinci sınıfta öğretmen tarafından dönemsel olarak yapılan sınavlardan bir tanesi ortak olarak gerçekleştirilmektedir. Ortak sınavlar, her dönem iki yazılısı olan derslerden birincisi, üç yazılısı olan derslerden ise ikincisi olmak üzere, akademik takvime uygun işleyen müfredatı kapsayacak şekilde, her dönem iki okul gününde yapılmaktadır. Sorular çoktan seçmeli olarak uygulanmakta ve diğer yıllardaki sınavlardan farklı olarak düzeltme formülü kullanılmamaktadır. Bu ise öğrencilere şans başarısını getirmekte ve seçkisiz hata kaynağı oluşturmaktadır.

Öğrenciler sınavlarda olağanüstü bir hal olmadığı sürece kendi okullarında sınava girmekte; öğretmenler ise başka okullarda görevlendirilmektedir. Ayrıca öğretmen kendi branşının sınavının yapılacağı gün sınav görevlisi olamamaktadır. Geçerli bir sebepten sınava giremeyen öğrencilere ise mazeret sınavı uygulanmaktadır. Bu uygulama önceki yıllardaki sınavlardan farklılık taşımaktadır. Böylesine önem ve emek verilen bir sınavın çıktıları ise öğrenciler için oldukça kritik sonuçlar doğurmaktadır. Sınav sonuçlarına göre bireyleri bir üst eğitim kurumuna yerleştirme, okul başarılarını izleme, bir işe yerleştirme gibi yaşamını etkileyecek sonuçlar elde edilmektedir. Bu nedenle verilecek kararların isabetliliği için ölçme araçlarının istenilen özelliği hatasız ölçmesi beklenmektedir.

Bireyler hakkında ölçme araçlarından elde edilen sonuçlara dayanarak verilecek olan yanlış karar veya kararların bireysel ve toplumsal sonuçları kuşkusuz önemlidir. Bu önem doğrultusunda ölçme araçlarının güvenilir ve geçerli olması büyük önem arz etmektedir. Micheels ve Karnes’e (1950) göre bir testte bulunması gereken özellikler: kapsamlı, ayırt edici, objektif,

kullanışlı, geçerli ve güvenilir olmasıdır. Tekin (2000), Yılmaz (1998) ve Turgut'a (1995) göre ise testlerde bulunması gereken özellikler; geçerlilik ve güvenilirlik ve kullanışlılıktır. Kullanışlılık, bir testin uygulama ve puanlama kolaylığı, mali açıdan ekonomikliği olarak tanımlanmaktadır. Geçerlilik ölçülmek istenilen özelliğe başka etkenler katılmadan testin amacına hizmet etme derecesidir (Yılmaz,1998). Güvenilirlik ise ölçülmek istenen özelliğin seçkisiz (random) hatalardan arınık olma derecesidir (Turgut ve Baykul, 2010; Tekin, 2000 ve Demircioğlu, 2009).

Ölçmedeki hatalar testlerin güvenilirliğini doğrudan etkileyebilmektedir. Dolayısıyla ölçme sürecinin olabildiğince hatadan arınık olması gerekmektedir. Olabildiğince denetimdeki ölçmelerin hatasız gerçekleşmesinin mümkün olmamasını ifade etmektedir. Az hata ile maksimum doğruluk sağlamak, gerçek sonuçlara yaklaşmak hedeflenmektedir. Hatalar istatistiksel olarak sabit, sistematik ve seçkisiz olmak üzere üç başlıkta tanımlanmaktadır. Bir ölçmeden diğerine miktarı değişmeyen hatalara sabit hata denir (Turgut ve Baykul, 2010). Sınav günü öğrencinin hastalığı, sınav koşullarının elverişsizliği, ya da bütün öğrenciler için eşit olmayışı, şans başarısı, öğretmenlerin cevapları puanlarken ve verdiği puanları toplarken dikkatsiz olması gibi etkenler seçkisiz hataların ortaya çıkmasına neden olur. Elde edilen ölçümlerde bu tür hata ne denli büyükse o ölçümlerin elde edildiği aracın güvenilirliği de o denli düşük olur. Diğer bir ifadeyle; bir ölçüm sonucu, içindeki seçkisiz hataların azlığı oranında güvenilir olur (Tekin, 2000).

Sistematik hata ise ölçülen büyüklüğe, ölçmeye veya ölçme koşullarına bağlı olarak sistematik olarak değişen hata olarak tanımlanmaktadır (Turgut ve Baykul, 2010). Bir öğretmen yazılı kâğıtları puanlanırken, yazısı güzel, düzgün ve okunaklı olanlara fazla puan verilmesi sistematik hatadır (Tekin, 2000). Bu türden sistematik hatalara istatistikte yanlılık denir. Diğer bir anlatımla yanlılık, farklı kategorilere ayrılmış bireylerin test puanlarının ayrıldıkları kategorilere bağlı olarak sistematik hata içermesidir (Camilli ve Shepard, 1994; Zumbo 1999).

Madde ve test yanlılığını belirlemenin ilk adımı Değişen Madde Fonksiyonunu (DMF) belirlemektir. DMF ölçme aracı ile ölçülen yetenek düzeyi bakımından benzer gruplarda ancak cinsiyet, okul türü, sosyo-ekonomik düzey gibi özellikler bakımından farklı gruplarda yer alan bireylerin maddeyi doğru yanıtlama olasılıklarının farklılaşması olarak tanımlanmaktadır (Hambleton, Swaminathan ve Rogers, 1991; Embretson ve Reise, 2000). DMF'nin ortaya çıktığı her koşulda yanlılık vardır demek doğru değildir. Zumbo'ya (1999) göre gruplar arasındaki bu farklılığın nedeni madde etkisinden veya DMF'den kaynaklanabilmektedir. Madde etkisi; bir maddeye doğru şekilde cevap veren ya da onaylayan farklı gruplardaki yanıtlayıcıların bir maddeyi doğru yanıtlama olasılıklarındaki gerçek farklılıkları ifade etmektedir. Maddenin ölçtüğü beceri, gruplar arasında önemli farklılıklar oluşturur (Zumbo,1999). Bu farklılık uygulayıcıların herhangi bir özelliğine göre ayırım yapıldığını değil, uygulayıcılar arasındaki tecrübe ya da önceden kazanılmış bilgi birikiminden doğan farklılıktan kaynaklanmaktadır. Dolayısıyla bir maddede DMF'ye rastlamak o maddenin yanlı olduğu anlamını içermemektedir. Eğer durum böyle olsaydı yanlılık belirlenirken uzman görüşü almaya gerek kalmaz ve DMF içeren bütün maddeler yanlıdır şeklinde ifade edilirdi. Bu sebeple DMF içeren madde ile yanlı madde arasında uzman görüşü farkı bulunmaktadır. Diğer bir anlatımla DMF'nin yanlılık çalışmalarının istatistiksel basamağı olduğu söylenebilir.

Eğitimde kullanılan testlerin ve maddelerinin, aynı yetenek düzeyinde olan fakat farklı gruplardan gelen bireylerin cevapları açısından farklı özellikler göstermemesi beklenmektedir. Diğer bir anlatımla test maddelerinin farklı cinsiyet, istatistiksel bölge veya sosyo-ekonomik düzey gruplarında bulunan aynı yetenek düzeyindeki bireylerden herhangi birinin lehine ya da aleyhine işlemiyor olması gerekmektedir. Test ya da maddelerin bu tür bir yanlılık taşıması ölçme araçlarının geçerliliğiyle ilgili önemli ve giderilmesi gereken bir sorundur. Özellikle öğrenciler hakkında önemli kararların alındığı seçme ve yerleştirme amacıyla kullanılan testlerin yanlılık içermemesi önemlidir.

Türkiye'de ulusal ve uluslararası ölçekte durum belirleme, seçme ve yerleştirme gibi amaçlarla farklı sınıf düzeyi ve ders alanlarında çeşitli sınavlar uygulanmaktadır. Ulusal değerlendirmelerden biri de Milli Eğitim Bakanlığı (MEB) tarafından uygulanan Temel

Öğretimden Orta Öğretime Geçiş (TEOG) sınavıdır. Öğrencileri bir üst eğitim kurumuna seçme amacıyla yapılan bu sınavın herhangi bir grup lehine yanlılık taşımaması gerekmektedir. Dolayısıyla öğrenciler hakkında önemli kararların verildiği bu sınavın yanlılık içermediğinin belirlenmesi önemli görülmektedir.

Türkiye’de madde yanlılığı ile ilgili yapılan ilk çalışmalar 1990’lı yılların ortalarına doğru başladığı görülmektedir. Son 10 yılda ise yanlılık çalışmalarında artış yaşanmıştır. Yapılan araştırmaların genel olarak geniş ölçekli sınavlar üzerinde yapıldığı görülmektedir. Alan yazın incelendiğinde çeşitli yıllarda uygulanan geniş ölçekli testlerin cinsiyet (Berberoğlu, 1995; Öğretmen ve Doğan, 2004; Yurdugül ve Aşkar, 2004a; Bekçi, 2007; Bakan Kalaycıoğlu, 2008; Deveci Ateşok, 2008; Bakan Kalaycıoğlu ve Kelecioğlu, 2011; Karakaya, 2012; Ulutaş, 2012; Karakaya ve Kutlu, 2012; Kan, Sünbül ve Ömür, 2013; Suna, 2012; Akalın, 2014; Türkan, 2014; Yalçın ve Tavşancıl, 2015; Zwick ve Erçikan, 1989; Harris ve Carlton, 1993; Gamer ve Engelhard, 1999; Le, 2009; Zenisky, Hambleton ve Robin, 2004; Lyons-Thomas, Sandilands ve Erçikan, 2014), okul türü (Bekçi, 2007; Bakan Kalaycıoğlu, 2008; Karakaya ve Kutlu, 2012) veya bölgesel (Yurdugül ve Aşkar, 2004b; Ulutaş, 2012; Özmen, 2014; Gök, Atalay Kabasakal ve Kelecioğlu, 2014; Zwick ve Erçikan, 1989; Erçikan ve Kim, 2009) kültürel yanlılıklarının incelendiği görülmektedir. Araştırmalarda çoğunlukla KTK'ya bağlı yöntemler tercih edilmiş olmakla beraber MTK'ya bağlı yöntemlerle yapılan araştırmalar oldukça az sayıdadır. Bu gereklilikten hareketle araştırmada aşağıdaki sorulara yanıt aranmıştır:

1. 2014 Temel Eğitimden Ortaöğretime Geçiş Sınavının Türkçe testi maddeleri Rasch Modeli yöntemiyle yapılan analizlerde cinsiyete göre DMF içermekte midir?
2. 2014 Temel Eğitimden Ortaöğretime Geçiş Sınavının matematik testi maddeleri Rasch Modeli yöntemiyle yapılan analizlerde cinsiyete göre DMF içermekte midir?
3. 2014 Temel Eğitimden Ortaöğretime Geçiş Sınavının fen ve teknoloji testi maddeleri Rasch Modeli yöntemiyle yapılan analizlerde cinsiyete göre DMF içermekte midir?
4. 2014 Temel Eğitimden Ortaöğretime Geçiş Sınavının T.C. İnkılap Tarihi ve Atatürkçülük testi maddeleri Rasch Modeli yöntemiyle yapılan analizlerde cinsiyete göre DMF içermekte midir?

## Yöntem

### *Araştırma Modeli*

Var olan bir durumun var olduğu biçimiyle, herhangi bir değiştirme, etkileme olmaksızın betimlenmesi amaçlandığından bu araştırma, tarama modelinde betimsel bir araştırmadır (Karasar, 2008).

### *Evren ve Örneklem*

Araştırmanın evrenini, 2014 Kasım TEOG sınavına giren bütün öğrenciler oluşturmaktadır. Araştırmanın örneklemini ise sınava giren bütün öğrenciler arasından seçkisiz atanmış 40000 öğrenci oluşturmaktadır. Örneklemdaki öğrencilerin %52,14’ü erkek ve % 47,86’sı kızdır. 2014 Kasım TEOG sınavına giren 40000 öğrencinin TEOG sınavında almış oldukları kitapçık türüne ve cinsiyetlerine göre dağılımı Tablo 1’de verilmiştir.

**Tablo 1. Kitapçık ve Cinsiyet Türüne Göre Öğrenci Örneklem Dağılımı**

	A Kitapçığı	B Kitapçığı	C Kitapçığı	D Kitapçığı	Toplam
Kadın	4814	4849	4702	4779	19144
Erkek	5186	5151	5298	5221	20856
Toplam	10000	10000	10000	10000	40000

Tablo 1 incelendiğinde her kitapçık türünün eşit miktarda bulunduğu, öğrencilerin cinsiyete göre dağılımının birbirine yakın olduğu görülmektedir.

### **Veri Toplama**

Araştırmada, sekizinci sınıfa giden öğrencilere 2014 yılı Kasım ayında uygulanan TEOG sınavının Türkçe, matematik, fen ve teknoloji ve T.C. İnkılap Tarihi ve Atatürkçülük testlerine ait maddeler dikkate alınmıştır. Yabancı dil ve din kültürü ahlak bilgisi testleri araştırmaya dâhil edilmemiştir. Yabancı dil testi İngilizce, Almanca, İtalyanca ve Fransızca alt testleri olarak uygulanmakta; dini bilgiler testi ise din kültürü ve ahlak bilgisi ve Musevilik alt testleri olarak uygulanmaktadır. Örnekleme sınırlandırmamak için bu testler analize dahil edilmemiştir. Testlerin her biri 20 şer maddeden oluşmaktadır. Veriler MEB Ölçme Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğünden yazılı izinle alınmıştır. Verilerin MEB'den alınmasına ilişkin dilekçe Ek 1'de sunulmuştur.

### **Verilerin Analizi**

Araştırmadan elde edilen veriler MEB Ölçme Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü'nden Excel formatında öğrencilerin ad, soyadı yerine sıra numarası, sınava katıldığı il, cinsiyeti, dersin adı, kitapçık türü ve öğrenci cevapları şeklinde gönderilmiştir. Araştırmaya katılan öğrencilerin maddelere vermiş oldukları cevapların doğru ya da yanlış olduğunu belirleyebilmek için verilerde cevap anahtarı gönderilmemiştir. Bu nedenle MEB ODSGM'den her dersin 4 farklı kitapçık türüne göre cevap anahtarı bulunmuştur. Cevap anahtarları Excel dosyasına aktarılmıştır ve öğrencilerin "A,B,C,D" formatında olan cevapları verilen cevaplar doğrultusunda "1-0" formatına çevrilmiştir. Yapılan "1-0" kodlamasından sonra cinsiyetlerle gruplar birleştirilerek Winsteps 3.8.1 programı yardımıyla bu programa uygun olacak şekilde Metin Belgesi Dosyasına aktarılmıştır. Aktarılan dosyada uygun syntaxlar yazılmış ve verilerin Winsteps paket programında kullanılacak hale getirilmiştir.

Verilerin Rasch modeli sayıtlarını sağlayabildiğini tespit etmek için SPSS ve Lisrel paket programları kullanılmıştır. Verilerin normal dağılım gösterip göstermediği SPSS; tek boyutlu olup olmadığı ise Lisrel paket programı ile analiz edilmiştir.

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 programı Rasch modelinde DMF tespiti için en çok tercih edilen uygulamalardan biridir. Winsteps programından elde edilen uyum istatistikleri olan iç uyum istatistikleri, maddelere verilen cevapları etkileyen, beklenmeyen davranışları sezme duyarlıyken; dış uyum istatistikleri ise insanların maddeler üzerindeki beklenmeyen davranışlarına duyarlıdır (Linacre, 2002). Madde uyum istatistikleri MNSQ (ortalamalar karesi) beklenen değerleri 1.0 dır.  $MNSQ > 2.0$  durumunda ölçüm sistemini olumsuz etkilemektedir. Değerin  $1.5 < MNSQ \leq 2.0$  olması ölçüm yapısı için yetersizken değer  $0.5 < MNSQ \leq 1.5$  aralıkta olması ölçüm için verimli ve yeterlidir. MNSQ değerinin 0,5'ten küçük olması ölçüm için daha az verimlidir (Linacre, 2002). Bond ve Fox (2007) iç uyum ve dış uyum ölçek maddeleri için madde ortalamalar karesinin 0,6 ile 1,4 arası değişebileceğini söylemektedir. Ayrıca nokta çift serili (PTMEA) korelasyon değerinin 0,20 değerinden büyük olması maddenin test yapısıyla uygun olduğunu göstermektedir.

Winsteps paket programında DMF'nin bulunup bulunmadığını anlayabilmek için analiz sonuçlarından DMF kontrast değerine bakılmaktadır. Lai ve Eton'a (2002) göre DMF kontrast değerinin 0,5 logit değerinde olması likert tipi ölçekler için kritik değerdir. Cinsiyet bakımından DMF değerinin 0,5 logit değerinin altında olması maddenin DMF içermediğini göstermektedir. Negatif DMF kontrast değeri maddenin odak gruba kolay geldiğini ve maddenin odak gruba avantaj sağladığını göstermektedir. Yapılan analizlerde maddelerin DMF içerip içermediğini tespit etmek için DMF kontrast değerleri kontrol edilmiştir. Logit değerlerinin istenilen aralıklarda olması göz önüne alınarak değerlendirmeler incelenmiştir

### **Bulgular**

Yapılan analizlerde Rasch modeline göre maddelerin DMF içerip içermediği incelenmiştir. Rasch modelinde madde ve yetenek parametrelerinin kestirilmesinde; ölçülen yeteneğin normal dağılım göstermesi, tek boyutlu olması, test maddelerinin yerel bağımsız olması ve maddelerin

“a” (madde ayırıcılık) parametresinin 1 olması gibi sayıtları karşılaması gerekmektedir (Gelbal, 1994). Bu nedenle yapılan istatistiklerin Rasch modelinin sayıtlarını karşılayıp karşılamadığını test etmek amacıyla Türkçe, matematik, fen ve teknoloji ve T.C. İnkılap Tarihi ve Atatürkçülük testlerinin; normal dağılım, tek boyutluluk ve maddelerin yerel bağımsızlığı analiz edilmiştir. Curry, Bashow ve Rentz (1978)’e göre Rasch modeli sayıtlarından biri olan testteki maddelerin ayırıcılık değerlerinin bir değerinde olması gerekliliği koşulunu sağlayan maddelerin kolay elde edilememesinden dolayı bu sayıltıyı sağlamayan maddelerle de ölçümün yapılabileceğini belirtmiştir (akt. Kaptan,1994). Aşağıda sırasıyla bütün testlere ait verilere normallik, tek boyutluluk, yerel bağımsızlık analizleri yapılmış ve testler bazında her biri değerlendirilmiştir.

Türkçe, matematik, fen ve teknoloji ve T.C. İnkılap Tarihi ve Atatürkçülük testlerinin normal dağılım gösterip göstermediğini belirlemek amacı ile normallik grafiği ile çarpıklık ve basıklık katsayıları hesaplanmıştır. Tablo 2’de Türkçe testine ilişkin betimsel istatistikler verilmiştir.

**Tablo 2. Türkçe Testine İlişkin Betimsel İstatistikler**

Test	N	$\bar{x}$	Mod	Medyan	s	s <sup>2</sup>	Basıklık	Çarpıklık
Türkçe	40000	5.54	3.00	5.00	2.54	6.43	-0.97	0.18

Tablo 2’de 20 soruluk Türkçe testi puan ortalaması 5.54, mod değeri 3.00, medyan değeri 5.00, standart sapması 2.54, varyansı 6.43, basıklık değeri -0.97 ve çarpıklık değeri 0.18 olarak tespit edilmiştir. Geniş örnekleme sahip bir testin normal dağılım gösterebilmesi için basıklık ve çarpıklık değerinin sıfıra yakın olması gerekmektedir. Çarpıklık katsayısı -1 ve +1 arası alınan değerler, normal dağılım olarak kabul edilebilir (Büyüköztürk, 2002). Dolayısıyla elde edilen değer Türkçe testi verilerinin normalden sapma göstermediğini ortaya koymaktadır. Araştırma kapsamında incelenen bir diğer test olan matematik testine ilişkin betimsel istatistikler Tablo 3’te sunulmuştur.

**Tablo 3. Matematik Testine İlişkin Betimsel İstatistikler**

Test	N	$\bar{x}$	Mod	Medyan	s	s <sup>2</sup>	Basıklık	Çarpıklık
Matematik	40000	4.62	4.00	5.00	1.78	3.19	0.26	0.36

Tablo 3’te matematik testi puan ortalaması 4.62, mod değeri 4.00, medyan değeri 5.00, standart sapması 1.78, varyansı 3.19, basıklık değeri 0.26 ve çarpıklık değeri 0.36 olarak tespit edilmiştir. Çarpıklık katsayısının -1 ile +1 arasında olmasından dolayı matematik testine ilişkin verilerin normal dağılım gösterdiği şeklinde yorumlanabilir. Araştırma kapsamında incelenen üçüncü test olan fen ve teknoloji testine ilişkin betimsel istatistikler Tablo 4’te sunulmuştur.

**Tablo 4. Fen ve Teknoloji Testine İlişkin Betimsel İstatistikler**

Test	N	$\bar{x}$	Mod	Medyan	S	S <sup>2</sup>	Basıklık	Çarpıklık
Fen Teknoloji	40000	4.97	5.00	5.00	1.76	3.09	0.06	0.15

Tablo 4 incelendiğinde fen ve teknoloji testi puan ortalaması 4.97, mod değeri 5.00, medyan değeri 5.00, standart sapması 1.76, varyansı 3.09, basıklık değeri 0.06 ve çarpıklık değeri 0.15 olarak tespit edilmiştir. Çarpıklık katsayısının -1 ile +1 arasında olmasından dolayı fen ve teknoloji testine ilişkin verilerin normal dağılım gösterdiği varsayılabilir. Son test olan T.C. İnkılap Tarihi ve Atatürkçülük testine ilişkin betimsel istatistikler ise Tablo 5’te verilmiştir.

**Tablo 5. T.C. İnkılap Tarihi ve Atatürkçülük Testine İlişkin Betimsel İstatistikler**

Test	N	$\bar{x}$	Mod	Med.	S	S <sup>2</sup>	Basıklık	Çarpıklık
T.C. İnkılap Tarihi ve Atatürkçülük	40000	4.90	5.00	5.00	2.07	4.30	-0.43	0.30

Tablo 5 incelendiğinde T.C. İnkılap Tarihi ve Atatürkçülük testi puan ortalaması 4.90, mod değeri 5.00, medyan değeri 5.00, standart sapması 2.07, varyansı 4.30, basıklık değeri -0.43 ve çarpıklık değeri 0.30 olarak tespit edilmiştir. Çarpıklık katsayısının -1 ile +1 arasında olmasından dolayı T.C. İnkılap Tarihi ve Atatürkçülük testine ilişkin verilerin normal dağılım gösterdiği varsayılabilir. Bu analizler ile testteki bütün maddelerin normal dağıldığı gözlenmiştir.

Normallik sayılıştısının ardından Rasch analizi için gerekli koşullardan biri olan verinin tek boyutlu yapıya sahip olup olmadığının incelenmesi aşamasına geçilmiştir. Tek boyutluluk, testteki bir set madde ile sadece tek bir yetenek ölçülmesidir (Hambleton ve diğ., 1991). Bireyin test performansı için gerekli olan tek bir baskın yetenek olduğunda kullanılan modellere tek boyutlu modeller denir (Embertson ve Reise, 2000). Güvenilir sonuçlar elde edebilmek için, verilerin tek boyutluluğunu gerektiren psikometrik yöntemlerin uygulanmasından önce tek boyutluluk varsayımının doğrulanması gereklidir (Gao, 1997, s.6). Bu amaçla testlere Doğrulayıcı Faktör Analizi (DFA) uygulanmış ve elde edilen sonuçlar aşağıda testler bazında sunulmuştur. Model 0.05 anlamlılık düzeyinde, Maximum Likelihood yöntemi kullanılarak test edilmiştir. İlk aşamada model için elde edilen uyum indeksleri incelenmiştir. Elde edilen uyum indekslerine ilişkin kabul edilebilir ve iyi uyum değerleri (Schermelleh-Engel ve Moosbrugger, 2003) Tablo 6'da sunulmuştur.

**Tablo 6.** Uyum İndekslerine İlişkin Kabul Edilebilir ve İyi Uyum Değerleri

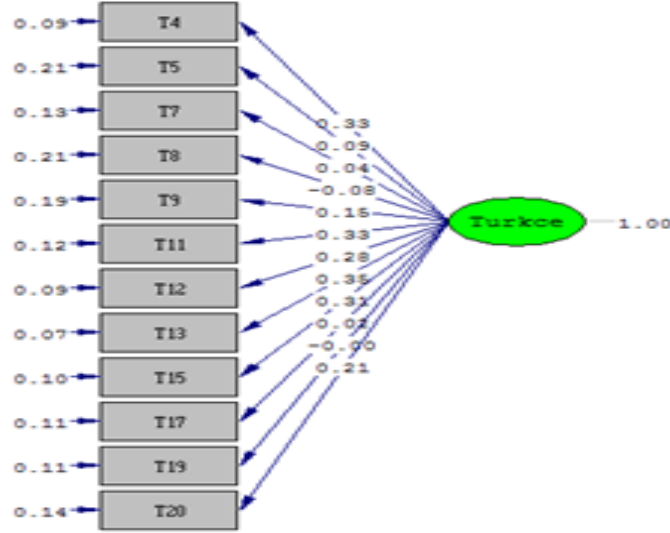
Uyum İndeksleri	İyi Uyum	Kabul Edilebilir Uyum
p değeri	$0.05 < p \leq 1.00$	$0.01 \leq p \leq 0.05$
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 \leq RMSEA \leq 0.10$
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 \leq SRMR \leq 0.10$
NNFI	$0.97 \leq NNFI \leq 1.00$	$0.95 \leq NNFI \leq 0.97$
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI \leq 0.95$
GFI	$0.95 \leq GFI \leq 1.00$	$0.90 \leq GFI \leq 0.95$

Matematik testine ilişkin uyum değerleri incelendiğinde RMSEA değerinin 0.04, GFI değerinin 0.97, AGFI 0.96 ve SRMR 0,03 olması iyi bir uyumu göstermektedir. Uyum indekslerinden sadece NFI ve CFI değerleri iyi uyum göstermemiştir. Matematik testine ilişkin DFA sonucunda elde edilen modele ilişkin faktör yükleri, Şekil 1'de yer almaktadır.



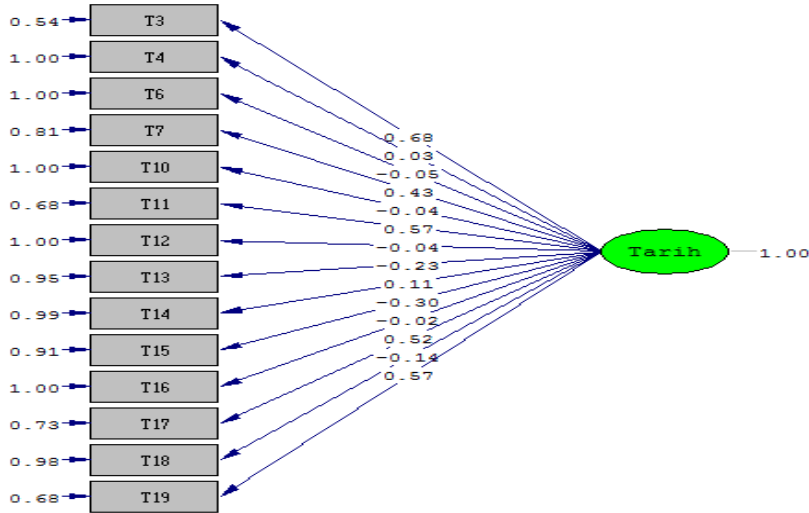
**Şekil 1.** Matematik Testinin Tek Boyutluluğuna İlişkin DFA Sonuçları

Araştırma kapsamında incelenen ikinci alt test olan Türkçe'de t değeri anlamlı olmayan sekiz madde (1., 2., 3., 6., 10., 14., 16. ve 18.) testten çıkarılmıştır. Kalan maddelerden elde edilen modele ilişkin uyum değerleri incelendiğinde RMSEA değerinin 0.05, GFI 0.97, AGFI 0.96, NFI 0.96, CFI 0.96 ve SRMR 0.05 olması iyi bir uyumu göstermektedir. Türkçe testine ilişkin DFA sonucunda elde edilen modele ilişkin faktör yükleri, Şekil 2'de yer almaktadır.



Şekil 2. Türkçe Testinin Tek Boyutluluğuna İlişkin DFA Sonuçları

Üçüncü test olan T.C. İnkılap Tarihi ve Atatürkçülük testinde t değeri anlamlı olmayan altı madde (1., 2., 5., 8., 9. ve 20.) testten çıkarılmıştır. Kalan maddelerden elde edilen modele ilişkin uyum değerleri incelendiğinde RMSEA değerinin 0.07, GFI 0.94, AGFI 0.92 ve SRMR 0.06 olması kabul edilebilir bir uyumu göstermektedir. NFI ve CFI değerleri iyi uyum göstermemiştir. TC. İnkılap Tarihi ve Atatürkçülük testine ilişkin DFA sonucunda elde edilen modele ilişkin faktör yükleri, Şekil 3'te yer almaktadır.

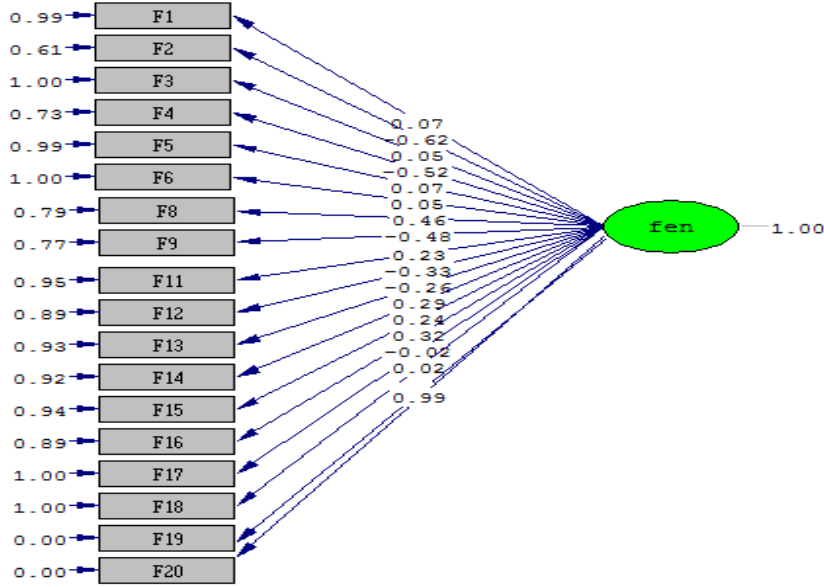


Şekil 3. T.C. İnkılap Tarihi ve Atatürkçülük Testinin Tek Boyutluluğuna İlişkin DFA Sonuçları

Araştırma kapsamında incelenen bir diğer test olan fen ve teknoloji testinde t değeri anlamlı olmayan iki madde (7 ve 10) testten çıkarılmıştır. Kalan maddelerden elde edilen modele ilişkin uyum değerleri incelendiğinde RMSEA değerinin 0.06, GFI 0.94, AGFI 0.93 ve SRMR 0.05 olması kabul edilebilir bir uyumu göstermektedir. NFI ve CFI değerleri iyi uyum



göstermemiştir. Fen ve teknoloji testine ilişkin DFA sonucunda elde edilen modele ilişkin faktör yükleri, Şekil 4'te yer almaktadır.



Şekil 4. Fen ve Teknoloji Testinin Tek Boyutluluğuna İlişkin DFA Sonuçları

Yapılan analizler sonucunda testlerden bazı maddeler atılarak tek boyutluluk için yeterli uyum indeksleri sağlanmıştır. Yerel bağımsızlık cevaplayıcının bir maddeye vermiş olduğu cevabın diğer maddelerden bağımsız olmasını gerektirmektedir. Bireyin bir maddede göstermiş olduğu başarı diğer maddeyi etkilememelidir. Cevaplayıcının her maddeye vermiş olduğu cevapların istatistiksel olarak birbirinden bağımsız olması gerekmektedir. Diğer bir anlatımla performans düzeyleri benzer olan öğrencilerin aynı soruya aynı cevabı vermeleri beklenmektedir. Yerel bağımsızlık; madde setindeki ilişkilerin tek yetenek ile açıklanması olduğundan, tek boyutluluk varsayımının karşılanması, yerel bağımsızlık varsayımının da sağlandığı şekilde yorumlanabilir (Emberson ve Reise, 2000; Hambleton ve Swaminathan, 1985).

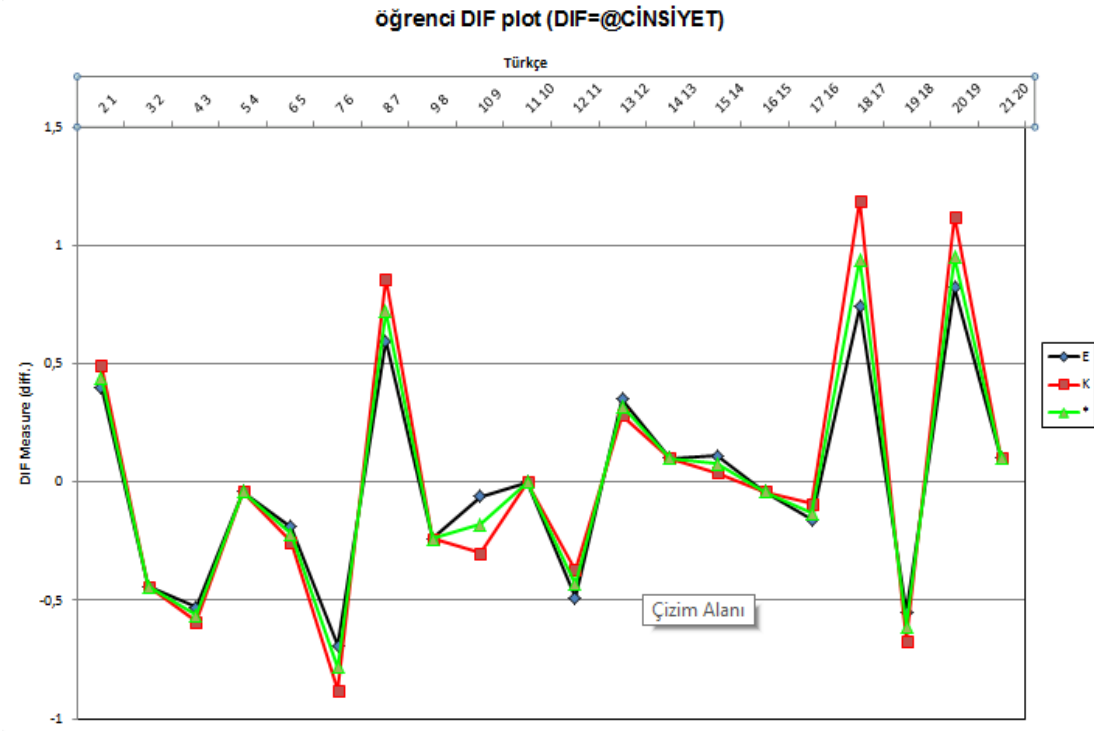
TEOG-Kasım 2014 verileri incelendiğinde Türkçe, matematik, fen ve teknoloji ve T.C. İnkılâp Tarihi ve Atatürkçülük testlerinin normallik ve tek boyutluluk varsayımını sağladığı yapılan analizlerle gösterilmiştir. Her bir testin kendi içerisinde tek boyutlu olduğu sağlandığı için yerel bağımsızlık sağlanmaktadır. Yerel bağımsızlık için farklı analizler yapmaya gerek yoktur. Verilerin normallik, tek boyutluluk ve yerel bağımsızlık şartlarını sağlamış olması analizlerin Rasch modeli ile yapılabileceğinin göstergesidir. Bu uygunluk çerçevesinde Türkçe, matematik, fen ve teknoloji ve T.C. İnkılâp Tarihi ve Atatürkçülük testlerine ilişkin DMF kontrast değerlerine ilişkin bulgular dersler bazında aşağıda verilmiştir.

Tablo 7'de öğrencilerin maddelere vermiş oldukları cevaplardan faydalanılarak oluşturulan Türkçe testinin DMF ölçüm değerleri verilmiştir. Tablo 7'de erkeklerin DMF ölçüm istatistikleri, kızların DMF ölçüm istatistikleri, toplam grubun ölçüm DMF değerleri, t değerleri ve maddelerin sıra numarası sunulmuştur. Tablo 7 incelendiğinde Türkçe testine verilen cevaplara göre erkek katılımcıların DMF ölçüm değerleri -0.74 ile 0.76 logit değeri aralığında değişmektedir. Türkçe testinde erkek katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -0.74 logit ölçümü ile 11. maddedir. En yüksek DMF değeri 0.76 logit ölçümü ile 19. maddedir. Türkçe testine verilen cevaplara göre kız katılımcıların DMF ölçüm değerleri -0.68 ile 1.21 logit değeri aralığında değişmektedir. Türkçe testinde kız katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -0.68 logit ölçümü ile 11. maddedir. En yüksek DMF değeri 1.21 logit ölçümü ile 17. maddedir.

**Tablo 7. Türkçe Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değeri**

Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	T	Madde
Erkek	-0.25	Kız	-0.25	.00	.00	4
Erkek	-0.39	Kız	-0.53	.14	5.96	5
Erkek	.52	Kız	.83	-.31	-10.1	7
Erkek	-0.44	Kız	-0.52	.07	2.95	8
Erkek	-0.24	Kız	-0.59	.34	13.6	9
Erkek	-0.74	Kız	-0.68	-.06	-2.61	11
Erkek	.23	Kız	.13	.10	3.54	12
Erkek	-0.05	Kız	-0.10	.06	2.20	13
Erkek	-0.24	Kız	-0.24	.00	.00	15
Erkek	.67	Kız	1.21	-.45	-16.5	17
Erkek	.76	Kız	1.12	-.36	-11.2	19
Erkek	-0.04	Kız	-0.11	.06	2.42	20

Tablo 7 incelendiğinde maddelerin DMF kontrast değerlerinin -0.45 ile 0.34 logit değerleri arasında değerler aldığı görülmektedir. Türkçe testinin en az logit değerini alan maddesi -0.45 ile 17. madde olurken, en yüksek logit değerini alan maddesi 0.34 ile 9. maddedir. DMF kontrast değerinin negatif yönlü olması yanlılığın odak grup (erkek) lehine olduğunu göstermektedir. Dolayısıyla Türkçe testindeki 7., 11., 17. ve 19. madde erkek öğrenciler lehine çıktığını göstermektedir. Uygun DMF kontrast değerleri -0.5 ile 0.5 logit değerleri arasındadır. Türkçe testine ilişkin t ölçümleri -16.5 ile 13.6 değerleri arasında değişmektedir. En düşük t değerini -16.5 ile 17. madde alırken; en yüksek t değerini 13.6 ile 9. madde almaktadır. Türkçe testinin cinsiyete göre DMF değişimi Şekil 5’de verilmiştir.



**Şekil 5. Türkçe Testinin Cinsiyete Göre DMF Değişimi**

Şekil 5’te Türkçe testine ait olan bütün maddelerin DMF ölçüm değerleri cinsiyetlere göre farklı kategorilerde gösterilmiştir. Şekil 5 incelendiğinde yeşil renk ile ifade edilen

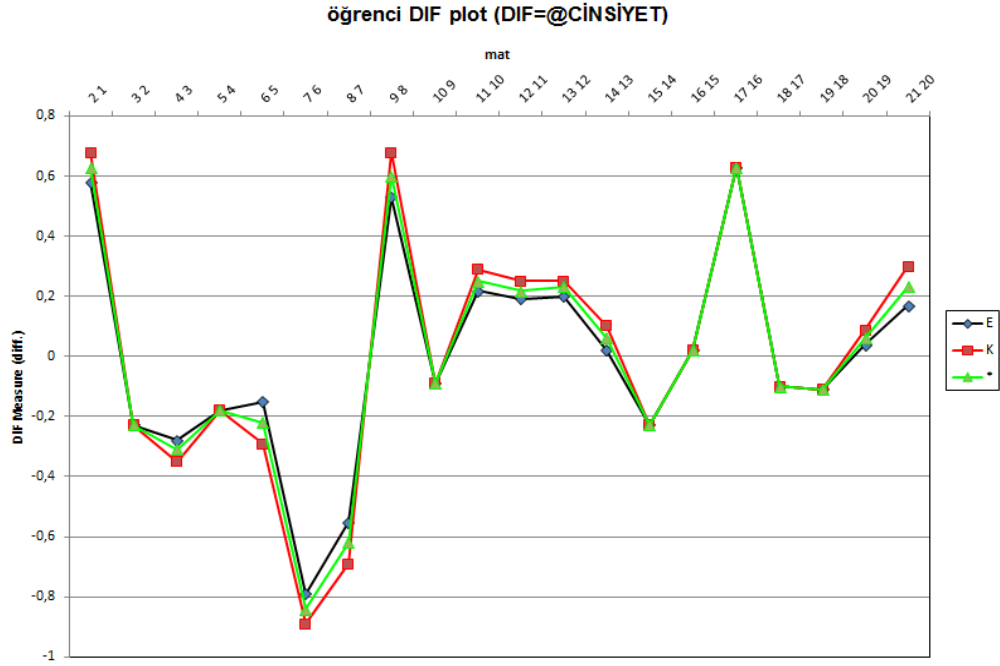
ölçümler ortalama olması gereken DMF değerini göstermekte iken, kırmızı renk ile gösterilen ölçümler kız öğrencilerin Türkçe testine ait ortalama DMF değerlerini göstermekte ve siyah renkli ölçümler ise erkek öğrencilerin Türkçe testine ait ortalama DMF değerlerini göstermektedir. Bir maddenin cinsiyete göre anlamlı düzeyde DMF içerebilmesi için kız ve erkek öğrenci değerlerinin arasında 0.5 değerinde farklılık olması gerekmektedir. Bütün maddeler incelendiğinde hiçbir maddede cinsiyete göre DMF değişimi 0.5 olmamıştır. Bu ise herhangi bir maddenin anlamlı düzeyde DMF içermediğinin kanıtıdır. Şekil 5 incelendiğinde en yüksek DMF değişimi 7.,8.,17. ve 19. maddelerde gözlenmektedir. Bir diğer alt test olan matematik testine ait kontrast değerleri Tablo 8'de verilmiştir.

**Tablo 8. Matematik Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değeri**

Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	T	Madde
Erkek	.58	Kız	.58	-.11	-3.56	1
Erkek	-.23	Kız	-.23	.00	.00	2
Erkek	-.28	Kız	-.28	.07	2.99	3
Erkek	-.18	Kız	-.18	.00	.00	4
Erkek	-.15	Kız	-.15	.14	5.84	5
Erkek	-.79	Kız	-.79	.10	4.56	6
Erkek	-.55	Kız	-.55	.13	6.16	7
Erkek	.53	Kız	.53	-.15	-5.07	8
Erkek	-.09	Kız	-.09	.00	.00	9
Erkek	.22	Kız	.22	-.07	-2.54	10
Erkek	.19	Kız	.19	-.05	-2.01	11
Erkek	.20	Kız	.20	-.05	-1.90	12
Erkek	.02	Kız	.02	-.08	-3.23	13
Erkek	-.23	Kız	-.23	.00	.00	14
Erkek	.02	Kız	.02	.00	.00	15
Erkek	.63	Kız	.63	.00	.00	16
Erkek	-.10	Kız	-.10	.00	.00	17
Erkek	-.11	Kız	-.11	.00	.00	18
Erkek	.04	Kız	.04	-.06	-2.36	19
Erkek	.17	Kız	.17	-.13	-4.91	20

Tablo 8 incelendiğinde matematik testine verilen cevaplara göre erkek katılımcıların DMF ölçüm değerleri -0.79 ile 0.63 logit değeri aralığında değişmektedir. Matematik testinde erkek katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -0.79 logit ölçümü ile 6. maddedir. En yüksek DMF değeri 0.63 logit ölçümü ile 16. maddedir. Matematik testine verilen cevaplara göre kız katılımcıların DMF ölçüm değerleri -0.79 ile 0.63 logit değeri aralığında değişmektedir. Matematik testinde kız katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -0.79 logit ölçümü ile 6. maddedir. En yüksek DMF değeri 0.63 logit ölçümü ile 16. maddedir.

Tablo 8'de matematik testinin DMF'ye ilişkin analiz sonuçları incelendiğinde maddelerin DMF kontrast değerleri -0.13 ile 0.14 logit değerleri arasında değerler almaktadır. Matematik testinin en az logit değerini alan maddesi -0.13 ile 20.madde olurken, en çok logit değerini alan maddesi 0.14 ile 5. maddedir. Uygun DMF kontrast değerleri -0.5 ile 0.5 logit değerleri arasındadır. Matematik testine ilişkin t ölçümleri -5.07 ile 6.16 değerleri arasında değişmektedir. En düşük t değerini -5.07 ile 8. madde alırken; en yüksek t değerini 6.16 ile 7. madde almaktadır. Matematik testine ait olan DMF ölçüm değerleri Şekil 6'da verilmiştir.



**Şekil 6.** Matematik Testinin Cinsiyete Göre DMF Değişimi

Şekil 6'da matematik testine ait olan bütün maddelerin DMF ölçüm değerleri cinsiyetlere göre farklı kategorilerde gösterilmiştir. Şekil 6 incelendiğinde yeşil renk ile ifade edilen ölçümler ortalama olması gereken DMF değerini göstermekte iken, kırmızı renk ile gösterilen ölçümler kız öğrencilerin matematik testine ait ortalama DMF değerlerini göstermekte ve siyah renkli ölçümler ise erkek öğrencilerin matematik testine ait ortalama DMF değerlerini göstermektedir. Bir maddenin cinsiyete göre anlamlı düzeyde DMF içerebilmesi için kız ve erkek öğrenci değerlerinin arasında 0.5 değerinde farklılık olması gerekmektedir. Bütün maddeler incelendiğinde hiçbir maddede cinsiyete göre DMF değişimi 0.5 olmamıştır. Bu ise herhangi bir maddenin anlamlı düzeyde DMF içermediğinin kanıtıdır. Şekil 6 incelendiğinde en yüksek DMF değişimi 5., 7. ve 20.maddelerde gözlenmektedir. Fen ve teknoloji testine ait kontrast değerleri Tablo 9'da verilmiştir.

**Tablo 9.** Fen ve Teknoloji Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değeri

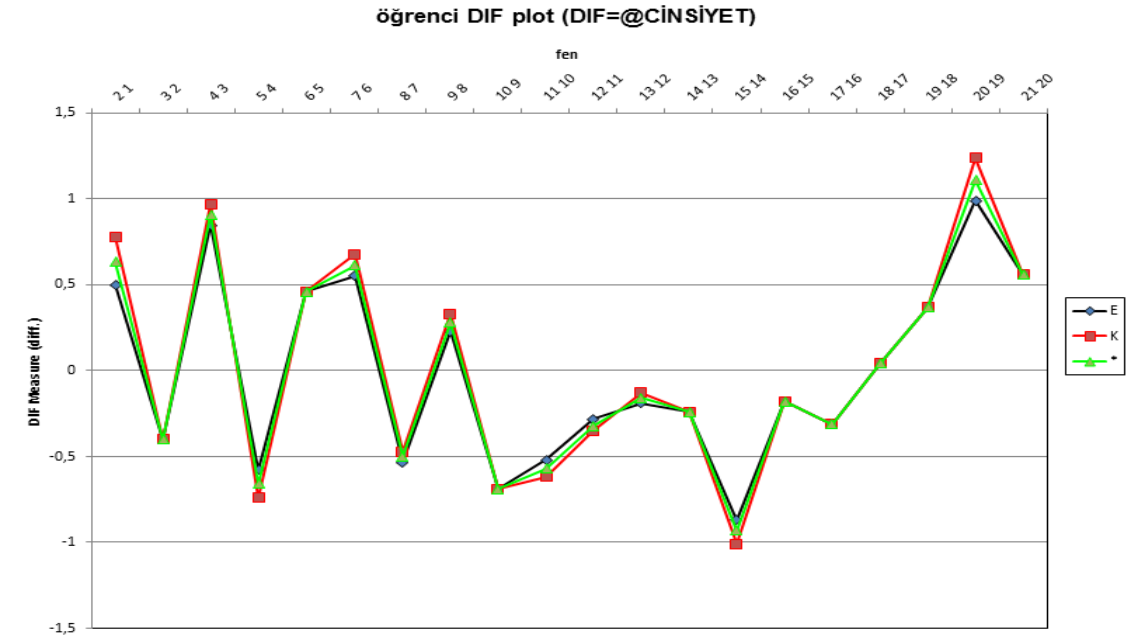
Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	T	Madde
Erkek	.50	Kız	.78	-.29	-9.86	1
Erkek	-.40	Kız	-.40	.00	.00	2
Erkek	.85	Kız	.97	-.12	-3.85	3
Erkek	-.58	Kız	-.74	.17	7.82	4
Erkek	.46	Kız	.46	.00	.00	5
Erkek	.55	Kız	.68	-.13	-4.65	6
Erkek	.23	Kız	.33	-.10	-3.83	8
Erkek	-.69	Kız	-.69	.00	.00	9
Erkek	-.28	Kız	-.35	.07	3.02	11
Erkek	-.19	Kız	-.13	-.05	-2.36	12
Erkek	-.24	Kız	-.24	.00	.00	13
Erkek	-.87	Kız	-1.01	.14	6.51	14
Erkek	-.18	Kız	-.18	.00	.00	15
Erkek	-.31	Kız	-.31	.00	.00	16
Erkek	.04	Kız	.04	.00	.00	17

Erkek	.37	Kız	.37	.00	.00	18
Erkek	.99	Kız	1.24	-.24	-7.07	19
Erkek	.56	Kız	.56	.00	.00	20

Tablo 9'da öğrencilerin maddelere vermiş oldukları cevaplardan faydalanılarak oluşturulan fen ve teknoloji testinin DMF ölçüm değerleri verilmiştir. Tabloda erkeklerin DMF ölçüm istatistikleri, kızların DMF ölçüm istatistikleri, toplam grubun ölçüm DMF değerleri, t değerleri ve maddelerin sıra numarası verilmiştir.

Tablo 9 incelendiğinde fen ve teknoloji testine verilen cevaplara göre erkek katılımcıların DMF ölçüm değerleri -0.87 ile 0.99 logit değeri aralığında değişmektedir. Fen ve teknoloji testinde erkek katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -0.87 logit ölçümü ile 14. maddedir. En yüksek DMF değeri 0.99 logit ölçümü ile 19. maddedir. Fen ve teknoloji testine verilen cevaplara göre kız katılımcıların DMF ölçüm değerleri -1.01 ile 1.24 logit değeri aralığında değişmektedir. Fen ve teknoloji testinde kız katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -1.01 logit ölçümü ile 14. maddedir. En yüksek DMF değeri 1.24 logit ölçümü ile 19. maddedir.

Tablo 9'da fen ve teknoloji testinin DMF'ye ilişkin analiz sonuçları incelendiğinde maddelerin DMF kontrast değerleri -0.29 ile 0.17 logit değerleri arasında değerler almaktadır. Fen ve teknoloji testinin en az logit değerini alan maddesi -0.29 ile 1. madde olurken, en çok logit değerini alan maddesi 0.17 ile 4. maddedir. DMF kontrast değerinin negatif yönlü olması yanlılığın odak grup (erkek) lehine olduğunu göstermektedir. Uygun DMF kontrast değerleri -0.5 ile 0.5 logit değerleri arasındadır. Fen ve teknoloji testine ilişkin t ölçümleri -9.86 ile 7.82 değerleri arasında değişmektedir. En düşük t değerini -9.86 ile 1. madde alırken; en yüksek t değerini 7.82 ile 4. madde almaktadır. Fen ve teknoloji testine ait DMF ölçüm değerleri Şekil 7'de verilmiştir.



**Şekil 7.** Fen ve Teknoloji Testinin Cinsiyete Göre DMF Değişimi

Şekil 7'de fen ve teknoloji testine ait olan bütün maddelerin DMF ölçüm değerleri cinsiyetlere göre farklı kategorilerde gösterilmiştir. Şekil 7 incelendiğinde yeşil renk ile ifade edilen ölçümler ortalama olması gereken DMF değerini göstermekte iken, kırmızı renk ile gösterilen ölçümler kız öğrencilerin fen ve teknoloji testine ait ortalama DMF değerlerini göstermekte ve siyah renkli ölçümler ise erkek öğrencilerin fen ve teknoloji testine ait ortalama DMF değerlerini göstermektedir. Bir maddenin cinsiyete göre anlamlı düzeyde DMF

içerebilmesi için kız ve erkek öğrenci değerlerinin arasında 0.5 değerinde farklılık olması gerekmektedir. Bütün maddeler incelendiğinde hiçbir maddede cinsiyete göre DMF değişimi 0.5 olmamıştır. Bu ise herhangi bir maddenin anlamlı düzeyde DMF içermediğinin kanıtıdır. Şekil 7 incelendiğinde en yüksek DMF değişimi 1., 6. ve 19.maddelerde gözlenmektedir. Son olarak T.C. İnkılap Tarihi ve Atatürkçülük testine ait kontrast değerleri Tablo 10'da verilmiştir.

**Tablo 10.** T.C. İnkılap Tarihi ve Atatürkçülük Testine İlişkin Değişen Madde Fonksiyonu Kontrast Değeri

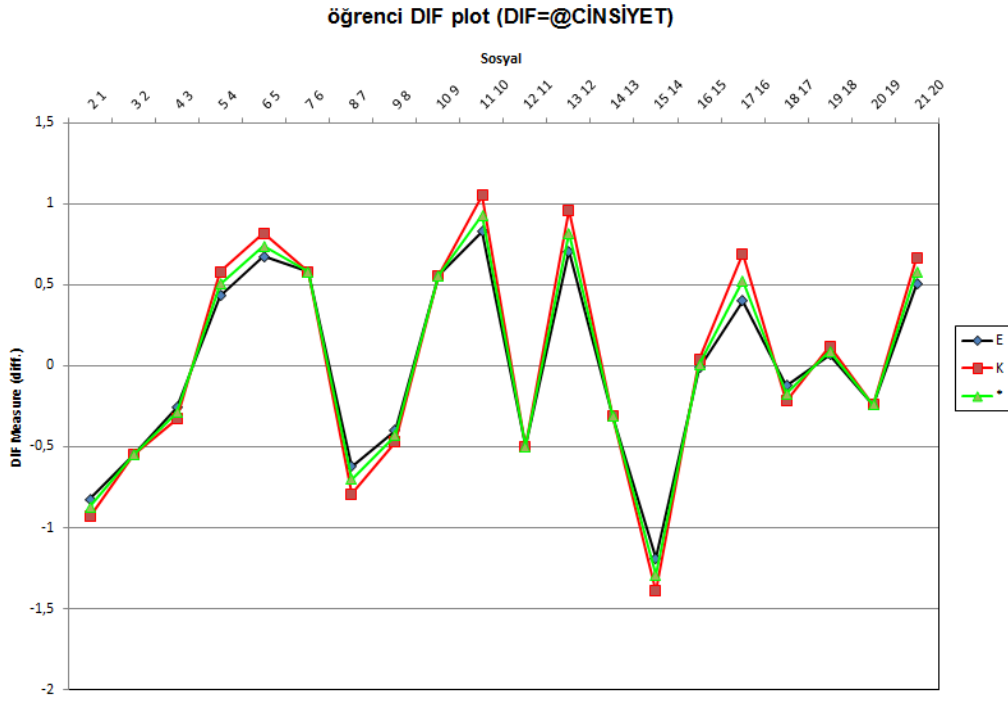
Odak Grup	DMF Ölçüm	Referans Grup	DMF Ölçüm	DMF Kontrast	t	Maddeler
Erkek	-0.26	Kız	-0.34	.08	3.29	3
Erkek	.45	Kız	.60	-.15	-5.27	4
Erkek	.60	Kız	.60	.00	.00	6
Erkek	-.63	Kız	-.81	.18	8.25	7
Erkek	.86	Kız	1.07	-.22	-6.51	10
Erkek	-.51	Kız	-.51	.00	.00	11
Erkek	.73	Kız	.98	-.25	-7.89	12
Erkek	-.32	Kız	-.32	.00	.00	13
Erkek	-1.23	Kız	-1.43	.20	9.34	14
Erkek	-.01	Kız	.05	-.06	-2.30	15
Erkek	.42	Kız	.70	-.29	-9.86	16
Erkek	-.12	Kız	-.22	.11	4.44	17
Erkek	.08	Kız	.12	-.04	-1.62	18
Erkek	-.25	Kız	-.25	.00	.00	19

Tablo 10 incelendiğinde T.C. İnkılap Tarihi ve Atatürkçülük testine verilen cevaplara göre erkek katılımcıların DMF ölçüm değerleri -1.23 ile 0.86 logit değeri aralığında değişmektedir. T.C. İnkılap Tarihi ve Atatürkçülük testinde erkek katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -1.23 logit ölçümü ile 14. maddedir. en yüksek DMF değeri 0.86 logit ölçümü ile 10.maddedir. T.C. İnkılap Tarihi ve Atatürkçülük testine verilen cevaplara göre kız katılımcıların DMF ölçüm değerleri -1.43 ile 1.07 logit değeri aralığında değişmektedir. T.C. İnkılap Tarihi ve Atatürkçülük testinde kız katılımcıların vermiş olduğu cevaplara göre en düşük DMF değeri -1.43 logit ölçümü ile 14. maddedir. en yüksek DMF değeri 1.07 logit ölçümü ile 10. maddedir.

Tablo 10'da T.C. İnkılap Tarihi ve Atatürkçülük testinin DMF'ye ilişkin analiz sonuçları incelendiğinde maddelerin DMF kontrast değerleri -0.29 ile 0.18 logit değerleri arasında değerler almaktadır. T.C. İnkılap Tarihi ve Atatürkçülük testinin en az logit değerini alan maddesi -0.29 ile 16.madde olurken, en çok logit değerini alan maddesi 0.18 ile 7. maddedir. DMF kontrast değerinin negatif yönlü olması yanlılığın odak grup (erkek) lehine olduğunu göstermektedir. Uygun DMF kontrast değerleri -0.5 ile 0.5 logit değerleri arasındadır. T.C. İnkılap Tarihi ve Atatürkçülük testine ilişkin t ölçümleri -9.80 ile 9.27 değerleri arasında değişmektedir. En düşük t değerini -9.80 ile 16. madde alırken; en yüksek t değerini 9.27 ile 14. madde almaktadır. T.C. İnkılap Tarihi ve Atatürkçülük testine ait olan DMF ölçüm değerleri Şekil 8'de verilmiştir.

Şekil 8'de T.C. İnkılap Tarihi ve Atatürkçülük testine ait olan bütün maddelerin DMF ölçüm değerleri cinsiyetlere göre farklı kategorilerde gösterilmiştir. Şekil 8 incelendiğinde yeşil renk ile ifade edilen ölçümler ortalama olması gereken DMF değerini göstermekte iken, kırmızı renk ile gösterilen ölçümler kız öğrencilerin T.C. İnkılap Tarihi ve Atatürkçülük testine ait ortalama DMF değerlerini göstermektedir. Siyah renkli ölçümler ise erkek öğrencilerin T.C. İnkılap Tarihi ve Atatürkçülük testine ait ortalama DMF değerlerini göstermektedir. Bir maddenin cinsiyete göre anlamlı düzeyde DMF içerebilmesi için kız ve erkek öğrenci değerlerinin arasında 0.5 değerinde farklılık olması gerekmektedir. Bütün maddeler

incelendiğinde hiçbir maddede cinsiyete göre DMF değişimi 0.5 olmamıştır. Bu ise herhangi bir maddenin anlamlı düzeyde DMF içermediğinin kanıtıdır. Şekil 8 incelendiğinde en yüksek DMF değişimi 5.,10.,12.ve 16.maddelerde gözlenmektedir.



**Şekil 8.** T.C. İnkılap Tarihi ve Atatürkçülük Testinin Cinsiyete Göre DMF Değişimi

### Tartışma / Sonuç ve Öneriler

Bu araştırmada ortaöğretim kurumlarına öğrenci seçme ve yerleştirme amacıyla uygulanan TEOG sınavının alt testlerinde yer alan maddelerin cinsiyete göre DMF içerip içermediklerinin MTK'ya dayalı bir yöntem olan Rasch modeli ile belirlenmesi amaçlanmıştır. Maddelerin DMF tespitine geçilmeden önce yapılan analiz işlemlerinin Rasch modelinin sayıltılarını karşılayıp karşılamadığını belirlemek için Türkçe, matematik, fen ve teknoloji ve T.C. İnkılap Tarihi ve Atatürkçülük testlerinin normal dağılım, tek boyutluluk ve test maddelerinin yerel bağımsızlık özelliklerini sağlayıp sağlamadığı incelenmiştir. Çarpıklık katsayısının -1 ile +1 arasında olmasından dolayı dört alt testteki bütün maddelerin normal dağıldığı gözlenmiştir.

Normallik sayıltısının ardından Rasch analizi için gerekli koşullardan biri olan verinin tek boyutlu yapıya sahip olup olmadığının incelenmesi için maddelere DFA uygulanmıştır. DFA sonucunda Türkçe testinin genel yapısını bozan 1, 2, 3, 6, 10, 14, 16 ve 18; fen ve teknoloji dersinde 7 ve 10 ile T.C. İnkılap Tarihi ve Atatürkçülük dersinde 1, 2, 5, 8, 9 ve 20 numaralı maddeler testten atılmış ve tek boyutluluk sağlanmıştır. Tek boyutlu olduğu görülen Türkçe alt testinde 12, matematik alt testinde 20, fen ve teknoloji alt testinde 18 ve T.C. İnkılap Tarihi ve Atatürkçülük alt testinde 14 madde olmak üzere toplam 64 madde için DMF analizleri yapılmıştır.

Türkçe testine ilişkin DMF kontrast değerleri incelendiğinde DMF kontrast değerinin negatif değer alması yanlılığın odak grup lehine olduğunu göstermektedir. Kontrast değeri -0.50 ile 0.50 arasında değerler alması beklenmektedir. Analiz sonuçları incelendiğinde atılmamış olan maddelerin kiminin negatif kiminin pozitif kontrast değerleri almasına karşın -0.50 ile 0.50 aralığının dışında değer alan maddeye rastlanmamıştır. Sonuç olarak Türkçe testinin tek boyutluluğu ihlal eden (1., 2., 3., 6., 10., 14., 16. ve 18.) maddeleri hariç tutularak yapılan analizlerde DMF içeren madde bulunmamıştır. Bu ise Türkçe alt testindeki 12 maddenin herhangi bir gruba avantaj veya dezavantaj sağlamadığı kanısına ulaştırmıştır. Bu durum Kelecioğlu, Karabay ve Karabay'ın (2009) 2009 yılı SBS üzerine yapmış oldukları çalışmanın

bulgularıyla paralellik göstermektedir. İlgili çalışmada cinsiyete göre yapılan DMF analizlerinde Türkçe testinde bir maddede DMF'ye rastlanılmış fakat alınan uzman görüşünde DMF'nin yanlışlıktan kaynaklanmadığı sonucuna ulaşılmıştır. Benzer şekilde Türkan (2014) tarafından 2012 SBS üzerine yapılan çalışmada da Türkçe maddelerinin cinsiyete göre DMF içermediği sonucuna varılmıştır. Yurdugül ve Aşkar (2004) da yapmış oldukları çalışmada MH tekniği ile Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı'ndaki maddelerin öğrencilerin cinsiyetleri açısından DMF içerip içermediği araştırılmıştır. Araştırma sonucunda Türkçe ve Sosyal Bilimler testlerinde anlamlı düzeyde DMF'ye rastlanmamıştır.

Bir diğer test olan matematik testine ilişkin DMF kontrast değerleri incelendiğinde maddelerin kiminin negatif kiminin pozitif kontrast değerleri almasına karşın -0.50 ile 0.50 aralığının dışında değer alan maddeye rastlanmamıştır. Sonuç olarak matematik testinin hiçbir maddesinde DMF bulgusuna rastlanmamıştır. Bu ise matematik alt testindeki 20 maddenin herhangi bir gruba avantaj veya dezavantaj sağlamadığı kanısına varılmıştır. Karakaya (2012) da 2009 SBS'nin matematik testindeki maddelerin cinsiyete göre yanlışlığı incelediği çalışmasında benzer sonuçlara ulaşmıştır.

Fen ve teknoloji testine ilişkin DMF kontrast değerleri incelendiğinde atılmamış olan maddelerin kiminin negatif kiminin pozitif kontrast değerleri almasına karşın -0.50 ile 0.50 aralığının dışında değer alan maddeye rastlanmamıştır. Sonuç olarak fen ve teknoloji testinin tek boyutluluğu ihlal eden (7. ve 10.) maddeleri hariç tutularak yapılan analizlerde DMF içeren madde bulunmamıştır. Bu ise fen ve teknoloji testindeki 18. maddenin herhangi bir gruba avantaj veya dezavantaj sağlamadığı kanısına varılmıştır. Bu sonuç Türkan'ın (2014) yapmış olduğu çalışmanın bulgularıyla paralellik göstermektedir. Araştırmacı ilgili çalışmada 2012 SBS fen ve teknoloji maddelerinin cinsiyete göre DMF içermediği sonucunu bulmuştur.

T.C. İnkılap Tarihi ve Atatürkçülük testine ilişkin DMF kontrast değerleri incelendiğinde atılmamış olan maddelerin kiminin negatif kiminin pozitif kontrast değerleri almasına karşın -0.50 ile 0.50 aralığının dışında değer alan maddeye rastlanmamıştır. Sonuç olarak T.C. İnkılap Tarihi ve Atatürkçülük testinin tek boyutluluğu ihlal eden (1., 2., 5., 8., 9. ve 20.) maddeleri hariç tutularak yapılan analizlerde DMF içeren madde bulunmamıştır. Bu ise T.C. İnkılap Tarihi ve Atatürkçülük testindeki 14 maddenin herhangi bir gruba avantaj veya dezavantaj sağlamadığı kanısına varılmıştır. Bu sonuç Kelecioğlu, Karabay ve Karabay'ın (2009) yapmış oldukları 2009 yılı SBS'nin cinsiyete göre madde yanlışlığı çalışması sonucunda tarih testi maddelerinin yanlışlık sonuçları ile örtüşmektedir.

Bu araştırmanın sonucu ile birlikte verilen alan yazın çalışmaları da dikkate alındığında Türkiye'de ortaöğretime geçişte uygulanan geniş ölçekli testlerin alt testlerinde genel olarak cinsiyet düzeyinde bir yanlışlığın olmadığı söylenebilir. Geniş ölçekli sınavlar, Türkiye'de çoğunlukla öğrencileri seçme ve yerleştirme amacıyla yapılmakta olup sonuçları itibarıyla eğitimin paydaşlarının büyük oranda etkilemektedir. Bu nedenle herhangi bir gruba yanlış bir sonuç üretmemesi beklenmektedir. Eğitim, toplumsal sınıf ayrımının bireylerin çabaları doğrultusunda giderilmesi noktasında önemli bir işleve sahiptir. Aynı zamanda statülerin yeniden dağıtımında öncelikli bir toplumsal alt sistemdir. Dolayısıyla eğitimde eşitliğin sağlanması, bir taraftan insan kaynaklarının gelişimine katkı sağlarken diğer taraftan toplumsal uzlaşmanın önemli bir noktası olabilmektedir. Bu doğrultuda bireylerin çabalarının ürüne dönüşmesinde belirleyici, eğitimin bir alt boyutu olan ölçme ve değerlendirmede, verilecek kararları etkileyebilecek her türlü yanlışlıktan arındırılmış sonuçların elde edilmesi gerekmektedir. Bu çalışma da Türkiye'de geniş ölçekli sınavlardan biri olan 2014 Kasım ayında uygulanan TEOG sınavının cinsiyet temelinde bir eşitsizlik oluşturmadığının bir göstergesidir. Araştırma sonucunda diğer araştırmacılara aşağıdaki öneriler sunulabilir:

- Yapılmış olan bu çalışmada tek boyutluluğu sağlamak için t değeri anlamlı olmayan maddeler analizden çıkarılmıştır. Bu maddeleri analizden çıkarmak yerine, araştırmacı testleri alt boyutlara ayırarak analiz yapmayı deneyebilir. Örneğin; matematik testini aritmetik ve geometrik olarak iki alt teste ayırıp tek boyutluluk analizi uygulanabilir ve tek boyutluluk sağlanırsa iki alt test için ayrı DMF analizleri yapılabilir. Böylece testten madde atmadan yanlışlık çalışması gerçekleştirilebilir.



- 2014-Kasım TEOG sınavı KTK'ye uygun olarak hazırlanmış bir sınavdır. Yapılan analizler ise MTK'ye dayalı Rasch Modeli ile yapılmıştır. KTK'ye dayalı diğer yöntemler kullanılarak analizlerin tekrarlanması fayda sağlayacağı düşünülmektedir.

### Kaynaklar

- Akalın, Ş. (2014). *Kamu Personeli Seçme Sınavı genel yetenek testinin madde yanlılığı açısından incelenmesi*. Yayınlanmamış doktora tezi. Ankara: Ankara Üniversitesi Eğitim Bilimleri Enstitüsü Ölçme ve Değerlendirme Anabilim Dalı.
- Bakan Kalaycıoğlu, D. (2008). *Öğrenci Seçme Sınavı'nın madde yanlılığı açısından incelenmesi*. Yayınlanmamış doktora tezi. Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulanması*. Ankara: ÖSYM Yayınları.
- Bakan Kalaycıoğlu, D. ve Kelecioğlu, H. (2011). Öğrenci Seçme Sınavının madde yanlılığı açısından incelenmesi. *Eğitim ve Bilim*, 36(161), 3-13.
- Bekçi, B. (2007). *Ortaöğretim Kurumları Öğrenci Seçme Sınavının değişen madde fonksiyonlarının cinsiyete ve okul türüne göre incelenmesi*. Yayınlanmamış yüksek lisans tezi. Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Berberoğlu, G. (1995). Differential item functioning analysis of computation, word problem and geometry questions across gender and SES groups. *Studies in Educational Evaluation*, 21, 439-456.
- Bond, T.G., & Fox, C.M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences. Second edition*. New Jersey: Lawrence Erlbaum Associates Publishers London.
- Büyüköztürk, Ş. (2002). *Sosyal bilimler için veri analizi el kitabı*. Ankara: PegemA Yayıncılık.
- Camilli, G., & Shepard, L.A. (1994). *Methods for identifying biased test items*. Newbury Park, CA: Sage
- Demircioğlu, G. (2009). *Geçerlik ve güvenirlik, ölçme ve değerlendirme*. Ankara: Pegem A Yayıncılık.
- Deveci Ateşok, N. (2008). *Üniversitelerarası kurul yabancı dil sınavının madde yanlılığı bakımından incelenmesi*. Yayınlanmamış doktora tezi. Ankara: Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates.
- Ercikan, K., & Kim, K. (2009). Examining the construct comparability of the English and french versions of TIMSS. *International Journal of Testing*, 5(1), 23-35.
- Gamer, M., & Engelhard Jr, G. (1999). Gender differences in performance on multiple-choice and constructed response mathematics items. *Applied Measurement in Education*, 12(1), 29-51.
- Gao, F. (1997). DIMTEST enhancements and some parametric IRT asymptotics. *Unpublished Doctoral Dissertation. University of Illinois at Urbana-Champaign, Department of Statistics*.
- Gelbal, S. (1994). *P madde güçlük indeksi ile Rasch modelinin b parametresi ve bunlara dayalı yetenek ölçüleri üzerine bir karşılaştırma*. Yayınlanmamış doktora tezi. Ankara: Hacettepe Üniversitesi Eğitim Bilimleri Bölümü.
- Gök, B., Atalay Kabasakal K. ve Kelecioğlu, H. (2014). PISA 2009 öğrenci anketi tutum maddelerinin kültüre göre değişen madde fonksiyonu açısından incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(1), 72-87.
- Görmez, M. ve Coşkun, İ. (2015). *1. Yılında temel eğitimden ortaöğretime geçiş reformunun değerlendirilmesi.(Analiz No. 114)*. Ankara: Siyaset, Ekonomi ve Toplum Araştırmaları Vakfı.
- Hambleton, R.K., & Swaminathan, H. (1985). *Item response theory: Principle sand applications*. Boston: Kluwer Academic Publishers.

- Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory*. California: Sage Publications Inc.
- Harris, A.M. & Carlton, S. T. (1993). Patterns of gender differences on mathematics items on the scholastic aptitude test. *Applied Measurement In Education*, 6(2), 137-151.
- Kan, A., Sünbül, Ö. ve Ömür, S.(2013). 6.-8. Sınıf seviye belirleme sınavları alt testlerinin çeşitli yöntemlere göre değişen madde fonksiyonlarının incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 9(2), 207-222.
- Kaptan, F. (1994). Rasch modeli madde parametrelerini kullanarak en yüksek olabilirlik yöntemiyle yeteneğin 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, 11 (2).
- Karasar, N. (2008). *Bilimsel araştırma yöntemleri*. Ankara: Nobel Yayıncılık.
- Karakaya, İ. (2012). Seviye belirleme sınavındaki fen ve teknoloji ile matematik alt testlerinin madde yanlılığı açısından incelenmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(1), 215-229.
- Karakaya, İ. ve Kutlu, Ö. (2012). Seviye belirleme sınavındaki Türkçe alt testlerinin madde yanlılığının incelenmesi. *Eğitim ve Bilim*, 37(165).
- Karasar, N. (2008). *Bilimsel Araştırma Yöntemleri*. Ankara, Nobel Yayıncılık.
- Kelecioğlu, H., Karabay, B. ve Karabay E. (2014). Seviye Belirleme Sınavı'nın madde yanlılığı açısından incelenmesi, *İlköğretim Online*, 13(3), 934-953.
- Lai, J.S., & Eton, D.T. (2002). Clinically meaning full gaps. *Rasch Measurement Transactions*, 15(4), 850.
- Le, L.T. (2009). Investigating gender differential item functioning across countries and test languages for PISA science items. *International Journal of Testing*, 9(2), 122-133. doi: 10.1080/15305050902880769.
- Linacre, J.M. (2002). *Differential item and test functioning*. 17 Temmuz 2015 tarihinde <http://www.rasch.org/rmt/rmt163g.htm>. adresinden alınmıştır.
- 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.
- Lyons-Thomas, J., Sandilands, D., & Ercikan, K. (2014). Gender differential item functioning in mathematics in four international jurisdictions. *Education and Science Large-Scale Assessment Special Issue*, 39(172), 20-32.
- MEB. (2011). *Türk milli eğitim sisteminin örgütlenmesi-2011*. 11 Mayıs 2015, <http://sgb.meb.gov.tr/eurydice/index.htm>.
- MEB. (2012). *SBS başvuru kılavuzu*. 13 Mayıs 2015 tarihinde <http://oges.meb.gov.tr/> adresinden alınmıştır.
- MEB. (2014). *26/27 Kasım 2014 tarihli 8. sınıf I. dönem ortak sınavı soruları ve cevap anahtarı*. 21 Haziran 2016 tarihinde [http://www.meb.gov.tr/meb\\_uyuru/](http://www.meb.gov.tr/meb_uyuru/) adresinden alınmıştır.
- Micheels, W.J., & Karnes, M.R. (1950). *Measuring education achievement*. Mc Graw-Hill.
- Öğretmen T. ve Doğan, N. (2004). OKÖSYS Matematik alt testine ait maddelerin yanlılık analizi. *İnönü Eğitim Fakültesi Dergisi*, 5(8).
- Özmen, D.T. (2014). PISA 2009 okuma testi maddelerinin yanlılığı üzerine bir çalışma. *Eğitim Bilimleri ve Uygulama Dergisi*, 13(26), 147-165.
- Rasch, G. (1961). On general law sand the meaning of measurement in psychology. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*. Berkeley, CA: University of California Press.
- Schermelleh-Engel, K. & Moosbrugger, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23-74.

- Suna, H.E. (2012). *TIMSS 2007 Fen bilimleri testindeki maddelerin dil ve cinsiyet yanlılığı açısından incelenmesi*. Yayınlanmamış yüksek lisans tezi. Ankara: Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Taylor, C.S., & 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* (17. Baskı). Ankara: Yargı Yayınevi.
- Turgut, M.F. (1995). *Eğitimde ölçme değerlendirme metotları* (10. Baskı). Ankara: Yargıcı Yayıncılık.
- Turgut, M.F. ve Baykul, Y. (2010). *Eğitimde ölçme değerlendirme*. Ankara: Pegem Yayıncılık.
- Türkan, A. (2014). *2012- Seviye Belirleme Sınavının Rasch modeline göre cinsiyet değişkeni açısından yanlılığının incelenmesi*. Yayınlanmamış yüksek lisans Tezi. Gaziantep: Gaziantep Üniversitesi Eğitim Bilimleri Enstitüsü.
- Ulutaş, S. (2012). *PISA 2006 fen okuryazarlığı testindeki maddelerin yanlılık bakımından araştırılması*. Yayınlanmamış doktora tezi. Ankara: Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Yalçın, S, ve Tavşancıl, E. (2015). TIMSS 2011 Fen Uygulamasında Cinsiyete Göre Farklılaşan Madde Fonksiyonunu Açıklayan Değişkenler. *Eğitim Bilimleri ve Uygulama*, 14(27), 1-21.
- Yılmaz, H. (1998). *Eğitimde ölçme ve değerlendirme*. Konya: Mikro Yayınevi.
- Yurdugül, H. ve Aşkar, P. (2004a). Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavının cinsiyete göre madde yanlılığı açısından incelenmesi. *Eğitim Bilimleri ve Uygulama Dergisi*, 3(5), 3-20.
- Yurdugül, H. ve Aşkar, P. (2004b). Ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavının, öğrencilerin yerleşim yerlerine göre, diferansiyel madde fonksiyonu açısından incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 27, 268-275.
- Zenisky, A.L., Hambleton, R.K., & Robin, F. (2004). DIF detection and interpretation in large-scale science assessments: Informing item writing practices. *Educational Assessment*, 9(1-2), 61-78.
- Zumbo, B.D. (1999). *A hand book on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary frame work for binary and Likert-type (ordinal) itemscores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP. *History Assessment Journal of Educational Measurement*, 26(1), 55-66.
- Wu, M.L., Adams, R.J., Wilson, M.R., & Haldane, S.A. (2007). *ACER ConQuest Version 2: Generalised item response modelling software [computer program]*. Camberwell: Australian Council for Educational Research.

## Extended Abstract

### Introduction

Tests and items are expected not to show different characteristics in terms of the answers of individuals who are at the same ability level but from different groups. In other words, it is a necessity that test items should not be working on behalf of or to the detriment of any individual that is at the same level with those of different gender, statistical area and socio-economical groups. It is a problem to solve if tests or items with a similar bias in terms of validity of assessment instruments. It is especially essential that selection and placement tests which are used to take important decisions about students should have no bias. In Turkey, many different kinds of examinations have been applied for many different purposes such as level determination, selection and placement in national or international scales. One of these national assessments is TEOG (Transition Exam from Primary to Secondary Education) which is applied to select students for an upper grade, so it should not have any kind of bias on behalf of any group. For this reason, it is extremely important to make it clear whether TEOG has any bias or

not. The objective of this study was to indicate whether Turkish, math, science and technology and Revolution History and Kemalism courses put forth any difference due to the gender of the individuals or not in TEOG examination.

### **Method**

The study was based on a descriptive model, as the objective was to describe a situation as it exists. The research includes all students who took part in the TEOG examination in November in 2014. The sample of the research consists of 40000 students that were chosen via random sampling method: 52.14% percentage of which consists of male and 47.86 of female. It was determined whether the items contains Differential Item Function (DIF) or not using the Rasch model and via the Winstep package software. Prior to determining DIF identification, it was examined whether Turkish, math, Revolution History and Kemalism and science and technology courses provide normal distribution, one dimensionality and local independency or not in order that the analysis processes can fulfill the requirements of the Rasch model.


### **Findings**

Coefficients of kurtosis and skewness were examined for normal distribution prior to the determination of DIF in RASCH model and it was observed that the obtained value was between -1 and +1 indicating that the premise has been supported at all subtests. Confirmative Factor Analysis (CFA) was applied to the data for testing the premise of unidimensionality. Following CFA, it was identified that some items affected the general structure of the test and those items were removed to provide unidimensionality. The removed items were: items 1, 2, 3, 6, 10, 14, 16 and 18 in the Turkish course test; items 7 and 10 in science and technology and items 1, 2, 5, 8, 9, 20 in Revolution History and Kemalism course tests. It was accepted that localized independency was provided in the tests that were accepted as unidimensional. For this reason, no analyses were required for localized independency condition. Following the acceptance of unidimensionality 12 items in Turkish subtests, 20 in math subtests, 18 in science and technology subtests and 14 in Revolution History and Kemalism subtest were subject to DIF analysis making up a total of 64 items. Negative DIF contrast values during the interpretation of the analysis via Winstep software indicate that the bias is in favor of the focal group. The condition for finding DIF is that the contrast values should be between -0,5 and 0,5 logit values. It was observed when the values related to DIF were analyzed according to gender difference that DIF contrast value of science and technology test item varies between -0.29-0.29 logits, whereas the values vary between -0.15 - 0.15 for math, -0.45-0.34 for Turkish and between -0.28 - 0.28 logits for Revolution History and Kemalism test item. No required contrast value was determined for any subtest item in the analyses applied for subtests. In other words, no item was determined in subtests which function differently with regard to gender.

### **Conclusion**

It is required that such an important examination which is applied at the national scale with the aim of taking important decisions about the students and selecting them for an upper grade should have reliable and valid consequences. According to the applied analysis, the items that were applied in TEOG in November in 2014 show no bias in terms of gender. This result indicates that there is no item that works for the behalf of or to the detriment of students from different gender groups.

**EK 1. Verilerin Milli Eğitim Bakanlığında alınmasına dair dilekçe**



T.C.  
MİLLÎ EĞİTİM BAKANLIĞI  
Ölçme, Değerlendirme ve Sınav Hizmetleri  
Genel Müdürlüğü

Sayı : 57750415-480.99-E.6209654  
Konu : Veri Talebi

16.06.2015

GAZİANTEP ÜNİVERSİTESİNE  
(Eğitim Bilimleri Enstitüsü)

İlgi : Gaziantep Üniversitesi Eğitim Bilimleri Enstitüsünün 62927161/302.08.01-799 sayılı ve 12/05/2015 tarihli yazısı.

İlgi yazınız ile Genel Müdürlüğümüzden istemiş olduğunuz veriler CD ortamında hazırlanarak yazımız ekinde gönderilmiştir.  
Bilgilerinizi ve gereğini rica ederim.

Kemal BÜLBÜL  
Bakan a.  
Daire Başkanı

EK: Veri CD'si (1 adet)

*Bilgi: EB ASD Bshlym  
falan 26/06/2015*

Güvenli Elektronik İmza  
Aslı ile Aynıdır.  
17/06/2015  
Şefik GÖ