



Madde Yanlılığının Belirlenmesinde Karma Madde Tepki Kuramının Kullanımı*

Seher YALÇIN**

Makale Bilgisi	ÖZET
Geliş Tarihi: 10.10.2018	<p>Çalışmada, Temel Eğitimden Ortaöğretime Geçiş (TEOG) sınavı Türkçe alt testi maddelerinin yanlılık gösterip göstermediğini incelemek amaçlanmıştır. Bu kapsamda, ilk olarak maddelerin farklılaşan madde fonksiyonu (FMF) olup olmadığı Karma Madde Tepki Kuramı (KMTK) modeli ve Mantel- Haenszel (MH) yöntemine göre belirlenmiştir. İki yöntemde göre de FMF olan maddeler, içerik ve bilişsel düzey bakımından incelenerek yanlılık tespiti için ayrıca uzman görüşüne başvurulmuştur. Araştırmanın örneklemini, TEOG 2015 Türkçe alt testini yanıtlayan öğrencilerden seçkisiz olarak seçilen 5000'i oluşturmaktadır. Analizler sonucu, KMTK'na göre test edilen modellerden, en iyi veriye uyum sağlayan modelin, iki gizil sınıftan oluşan model olduğu görülmüştür. İki yöntemde göre de FMF olarak tespit edilen yedi madde vardır. Bu maddelerin dördü ilk gizil gruba avantaj sağlarken üçü ikinci gizil gruba avantaj sağlamaktadır. İlk gizil gruba avantaj sağlayan dört madde, cümlede ve paragrafta anlam konularında, genellikle kavrama düzeyindeki maddelerdir. İkinci gruba avantaj sağlayan üç madde ise Türkçe dilbilgisi konusunda, bilgi düzeyinde maddeler olduğu tespit edilmiştir. Bu bağlamda, maddelerin bu gibi özelliklerinin bazı gruplara avantaj sağlarken bazılarında dezavantaja neden olduğu öne sürülebilir.</p>
Kabul Tarihi: 18.04.2020	
Erken Görünüm Tarihi: 18.04.2020	
Basım Tarihi: 30.04.2021	
Anahtar Sözcükler: Madde yanlılığı, karma madde tepki kuramı, Mantel- Haenszel	

The Usage of Mixed Item Response Theory in Determining Item Bias

Article Information	ABSTRACT
Received: 10.10.2018	<p>It is aimed in this study to identify whether the items in the Turkish subtest of the Transition from Basic Education to Secondary Education (TEOG) exam are biased. In this regard, whether the items show differential item functioning (DIF) is analyzed primarily according to the Mixed Item Response Theory (MixIRT) model and the Mantel- Haenszel (MH) method. The items that are determined to display DIF according to both methods are analyzed with regards to the subject field and cognitive level, as well as taking expert opinion on including bias. The sample of this study consists of 5000 students who were randomly selected from among the students that took the TEOG 2015 Turkish subtest. Analyses have shown that among the models based on MixIRT, the model with two latent classes is determined to fit the data best. When the items that indicate DIF according to both methods are analyzed in the DIF examination conducted based on the two latent classes that emerged according to the MixIRT, it is observed that four items that are all advantageous for the first latent group, are understanding level questions that contain the subjects of meaning in sentences and paragraphs. The other three items in favor of the second latent class were seen to be knowledge level questions measuring the Turkish grammar subject. In this context, it can be submitted that item features are favored for some groups and disadvantageous for some others.</p>
Accepted: 18.04.2020	
Online First: 18.04.2020	
Published: 30.04.2021	
Keywords: Item bias, mixture item response theory, Mantel- Haenszel	
doi: 10.16986/HUJE.2020059656	Makale Türü (Article Type): Araştırma Makalesi

Kaynakça Gösterimi: Yalçın, S. (2021). Madde yanlılığının belirlenmesinde karma madde tepki kuramının kullanımı. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 36(2), 393-405. doi: 10.16986/HUJE.2020059656

Citation Information: Yalçın, S. (2021). The usage of mixed item response theory in determining item bias. *Hacettepe University Journal of Education*, 36(2), 393-405. doi: 10.16986/HUJE.2020059656

1. GİRİŞ

Öğrenci başarısının belirlenmesi ve öğrencileri seçme ve yerleştirme amacıyla Türkiye'de pek çok sınav uygulanmaktadır. Yapılan sınavlara dayalı olarak da bireyler hakkında bazı kararlar verilmektedir. Verilen kararların doğruluğu ve uygunluğu, uygulanan ölçme aracından elde edilen puanların geçerliği ve güvenilirliğine bağlıdır. Ölçme aracının geçerliği kapsamında

* 11-12 Mayıs 2018 tarihlerinde Marmara Üniversitesi'nin düzenlediği 1. Uluslararası Eğitimde Yeni Arayışlar Kongresi'nde sözlü bildiri olarak sunulmuştur.

** Doç. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Eğitimde Ölçme ve Değerlendirme Ana Bilim Dalı, Ankara-TÜRKİYE. e-posta: yalcins@ankara.edu.tr (ORCID: 0000-0003-0177-6727)

incelenmesi gereken durumlardan biri, ölçme aracındaki maddelerin herhangi bir gruba yanlı olup olmadığıdır (Clauser ve Mazor, 1998). Yanlılık, yapılan bir ölçme işlemine sistematik hata karışması olarak bilinir (Osterlind ve Everson, 2009). Bir testi oluşturan maddelere sistematik hata karışması, testin geçerliğinin düşmesine neden olmaktadır. Bir testteki maddelerin yanlılık açısından incelenmesi için önce farklılaşan madde fonksiyonu (differential item functioning) olma durumuna bakılır. Alan yazında, “differential item functioning” teriminin değişen madde fonksiyonu (DMF), madde işlev farklılığı (MİF) gibi farklı kullanımları olduğu bilinmektedir. Bu çalışmada, farklılaşan madde fonksiyonu teriminin anlam olarak daha doğru bir çeviri olduğu düşüncesiyle bu ifade kullanılmıştır. Farklılaşan madde fonksiyonu (FMF), yetenek düzeyleri aynı olan bireylerin ya da bireyler, yetenek seviyelerine göre eşleştirildikten sonra farklı alt gruplardaki kişilerin bir maddeye doğru yanıt verme olasılığının birbirinden farklılığıdır (Clauser ve Mazor, 1998; Mellenbergh, 1989). FMF belirlenirken iki grup karşılaştırılır. Bu gruplardan referans grup, sıklıkla madde ile ölçülen özellik açısından avantajlı olduğu varsayılan grup olarak alınırken odak gruptaki bireyler ise madde ile ölçülen özellik açısından dezavantajlı olan bireylerdir (De Ayala, 2009; Osterlind ve Everson, 2009).

FMF'nin belirlenmesinde, alan yazında pek çok yöntem önerilmektedir. Bu yöntemlerden sık kullanılan bazıları; Mantel-Haenszel (MH), SIBTEST ve Lojistik regresyonken (Camilli ve Shepard, 1994; Holland ve Wainer, 1993; Osterlind ve Everson, 2009); son yıllarda FMF'nin kaynağını da belirleyebilmesi nedeniyle kullanılan bazı yöntemler ise; Muthen'in MIMIC modeli, açıklayıcı madde tepki modelleri (Wilson ve De Boeck, 2004), çok düzeyli lojistik regresyon (Chaimongkol, Huffer ve Kamata, 2007) ve karma madde tepki kuramı-KMTK (mixed item response theory)'dir (Choi, Alexeev ve Cohen, 2015; Cohen ve Bolt, 2005; Maij-de Meij, Kelderman ve van der Flier, 2010; Samuelsen, 2008). Bu modeller sadece FMF belirlemek amaçlı kullanılmamaktadır ancak çalışmanın kapsamı gereği modellerin FMF tespiti boyutu ele alınmıştır. Ayrıca, alan yazında, bilinen/gözlenen gruplara göre FMF belirlendiğinde, grup homojenliği varsayımını her zaman sağlayamadığı ifade edilmektedir (De Mars ve Lau, 2013; Samuelsen, 2008). Ek olarak, FMF ile ilgili son yapılan çalışmalar, FMF'nin kaynağının genellikle kompleks olduğunu, bilinen gruplarla doğrudan açıklanamadığını belirtmektedir (Cohen ve Bolt, 2005; De Mars ve Lau, 2013; Yalçın, 2018). Bu nedenle, FMF'nin gizil (latent) yani bilinmeyen (unknown) gruplar arasında test edilmesi önerilmektedir (Cohen ve Bolt, 2005; De Ayala, Kim, Stapleton ve Dayton, 2002; De Mars ve Lau, 2013; Finch ve French, 2012; Samuelsen, 2008; Yalçın, 2018). Ayrıca, yapılan simülasyon çalışmalarında da (Maij-de Meij, 2010; Uyar, Kelecioğlu & Doğan, 2017), FMF'yi tanımlamada gizil grup yaklaşımına dayalı KMTK'nın performansının bilinen değişkenleri kullanan FMF belirleme yöntemlerinden daha iyi olduğu görülmüştür. Bu bağlamda, bu çalışmada, FMF'nin kaynağını belirlemede önerilen gizil grup yaklaşımına dayalı KMTK yöntemi kullanılmıştır. Ayrıca FMF tespit edilen maddelerin karşılaştırılması amacıyla geleneksel yöntemlerden en sık kullanılan MH yöntemi kullanılmıştır.

Mantel ve Haenszel tarafından 1959 yılında geliştirilen MH yöntemi, ilk olarak Holland ve Thayer (1988) tarafından FMF belirlemek için ortaya konulmuştur. Non-parametrik bir yöntem olan MH, odak ve referans grup göstergesi tarafından ayrılan doğru ve yanlış tepkilerin sayılarının gösterildiği 2x2'lik çapraz tablolar yardımıyla eşleştirme ölçütüne göre eşleştirilen grupların karşılaştırılmasına dayanır (Holland ve Thayer, 1988). Diğer FMF yöntemleri gibi MH yöntemi, bilinen veya gözlenen gruplar arasında bir maddenin fonksiyonlaşma durumunu karşılaştırır.

Madde tepki kuramı ile gizil sınıf analizinin birlikte kullanılması ise “karma madde tepki kuramını (KMTK) - Mixture item response theory (MixIRT)- ortaya çıkarmaktadır (Cohen ve Bolt, 2005). KMTK modeli, gizil sınıf analizi ve madde tepki kuramını birleştiren güçlü istatistiksel bir kuramdır. KMTK modellerinin (Kelderman ve Macready, 1990; Maij-de Meij et al., 2010), katılımcıların tepkilerindeki nitel farklılıkların türü veya nedeni hakkında bir varsayımı yoktur. KMTK modellerinde (1-parametrelili, 2-parametrelili, 3-parametrelili), gizil sınıflar (homojen alt gruplar) tanımlanır. Her bir gizil sınıf içinde aynı ölçme modelinin bulunduğu, ancak gizil sınıflar arasında farklı parametre tahminleri yapılır. KMTK modeli, evrenin sınırlı sayıda gizil birey sınıfından oluştuğunu varsayar ve bu sınıflar madde tepki örüntülerine dayanarak farklılaşabilir (von Davier ve Rost, 2017). Buna karşılık, bu farklı tepki örüntüleri, kendilerini her bir gruba ilişkili madde tepki modelinin parametrelerinde farklılıklar olarak gösterecektir. İki parametrelili KMTK modelinin formülü şu şekildedir (Finch ve French, 2012):

$$P(U = 1|g, \theta_{ig}) = \frac{e^{(a_{jg}(\theta_{ig}-b_{jg}))}}{1 + e^{(a_{jg}(\theta_{ig}-b_{jg}))}}$$

Formülde “g: 1, 2, ..., G” gizil sınıf üyeliğini, “ b_{jg} ” j maddesi için sınıf içi güçlüğü, “ a_{jg} ” j maddesi için sınıf içi ayırıcılığı ve “ θ_{ig} ” i kişisi için sınıf içinde ölçülen gizil özelliğin düzeyini göstermektedir. KMTK modelleri, FMF'nin belirlenmesinde bilinen değişkenlere göre FMF incelemesi yapmadığından incelemeyi belirli değişkenlerle sınırlanamamaktadır. Bu nedenle FMF'nin kaynağını belirlemede daha uygun olduğu ifade edilmektedir (Maij-de Meij et al., 2010). FMF kaynağının belirlenmesi de, testin yapı geçerliğini pozitif yönde etkileyerek bireylerin yeteneklerinin daha hatasız kestirilmesini sağlar (Ong, Williams ve Lamprianou, 2013). KMTK'na göre FMF belirleme süreci genel olarak şu şekildedir: veriye en iyi uyum sağlayan model KMTK ile belirlenir. Bu belirleme için her bir bireyin ilgili alt teste verdiği yanıtlardan tepki örüntüleri oluşturularak bir gizil sınıflı modelden başlanıp model uyum istatistikleri en iyi değeri verene kadar gizil sınıf sayısı artırılarak analizler tekrarlanır. Veriye en iyi uyum sağlayan model belirlendikten sonra, belirlenen gizil sınıflar arasında FMF olup olmadığı incelenir. FMF incelemesi yapılırken, her bir madde için gizil sınıflar arasındaki güçlük farkları dikkate alınır. Analizler sonucu, her madde için hesaplanan katsayı değeri seçilen manidarlık düzeyine göre (p=.05 veya p=.01) manidar olan maddeler için öğrencilerin gizil yeteneği kontrol edildiğinde, gizil sınıflar arasında FMF içerdiği yorumu yapılır. Ardından, gizil sınıflardaki bireylerin hangi özelliğe sahip olduğu bilinmediğinden oluşan grupların özellikleri tahmin edilmeye çalışılır. Oluşan gizil sınıflar, cinsiyet gibi bilinen bir

değişkeni temsil edebileceği gibi tahmin edilmeyen sosyo-ekonomik düzey, ilgili konuya ilişkin duyuşsal özellikler, maddelerin konu alanı gibi özellikler de olabilir veya kesin sınırları çizilen değişkenler de olmayabilir.

KMTK modellerinde FMF'nin kaynağı tespit edilmeye çalışıldığından FMF'nin belirlenmesi yanında istatistiksel olarak FMF'nin kaynağının da belirlenmesine fırsat tanınması nedeniyle araştırmanın alana katkı sağlayacağı düşünülmektedir. Bu bağlamda, çalışmada ele alınan modellerin alan yazında yaygın kullanılmayışı ve yeterince tanınmayışı sebebiyle araştırmacılara farklı uygulama örnekleri sunması açısından da katkı getirebilir. Bu kapsamda, araştırmada ele alınan modelleri test etmek için Temel Eğitimden Ortaöğretime Geçiş (TEOG) sınavı Türkçe alt testi maddelerinin yanlı olup olmadığı incelenmiştir. TEOG sınavları 2013 yılında uygulamaya başlamış, 2017-2018 eğitim-öğretim yılında ise yerini Liseye Geçiş Sistemine (adrese dayalı geçiş ve sınırlı sayıda okula geçiş için zor bir sınav) bırakmıştır. TEOG gibi seçme ve yerleştirme amaçlı yapılan sınavlarda, bireyler arası farklılıkları adil bir şekilde ortaya koyabilmek önem taşımaktadır. Bu sınavı oluşturan maddelerin herhangi bir grup öğrenciye avantaj sağlamaması ya da dezavantaja neden olmaması yani yanlı olmaması istenir. Maddelerin yanlı olma durumunun incelenmesinde öncelikle farklılaşan madde fonksiyonunun tespit edilmesi gerekir. Bu nedenle, ilk olarak maddelerin FMF gösterip göstermediği gizil gruba dayalı yöntemlerden Karma Madde Tepki Kuramı (KMTK) ve gözlenen grup yöntemlerinden Mantel- Haenszel (MH)'e göre belirlenmiştir. İki yonteme göre de FMF'li olduğu tespit edilen maddelerin yanlı olma durumu için uzman görüşüne başvurulmuştur.

1.1. Araştırmanın Amacı

Bu çalışmada, madde yanlılığının belirlenmesinde gizil sınıflara göre FMF ve kaynağını belirlemede kullanılan Karma Madde Tepki Kuramı modelinin kullanımını göstermek amaçlanmıştır. Bu kapsamda çalışmada yanıt aranan sorular şunlardır:

1. KMTK'na göre veriye en iyi uyum sağlayan model hangisidir? Veriye uyum sağlayan modele göre oluşan gizil sınıflar arasında KMTK'na göre FMF gösteren maddeler nelerdir?
2. Gizil sınıflara göre MH yöntemi ile FMF gösteren maddeler nelerdir?
3. KMTK'na ve MH yöntemine göre FMF gösteren maddeler birbiriyle tutarlı mıdır? Tutarlı olan maddeler için öğrencilerin seçeneklere dağılımı, madde gücülüğü ve uzmanların görüşleri nasıldır?

2. YÖNTEM

2.1. Araştırma Grubu

Araştırmanın örneklemini, 2015 yılı birinci döneminde TEOG Türkçe alt testini yanıtlayan öğrencilerden seçkisiz olarak seçilen 5000 öğrenci oluşturmaktadır. Öğrencilerin %48.5'i kız, %51.5'u erkektir. Ayrıca, iki yonteme göre FMF olduğu saptanan maddeler için beş ölçme ve değerlendirme uzmanının ve iki Türkçe öğretmeninin görüşleri alınmıştır. Uzmanlardan, dördü ölçme ve değerlendirme alanında doktor, biri ise doktora tez aşamasındadır. Uzmanların seçiminde, okuduğunu anlama becerileri ve/veya madde yanlılığı konuları üzerine çalışmaları olmasına dikkat edilmiştir. Öğretmenler ise en az 10 yıllık öğretmenlik deneyimine sahip yüksek lisans öğrencileridir.

2.2. Veri Toplama Aracı ve Verilerin Analizi

Bu çalışmada kullanılan veriler, Milli Eğitim Bakanlığı (MEB) tarafından ulusal düzeyde yapılan TEOG sınavı 2015 yılı birinci dönem sonunda, altı temel dersten merkezi ortak sınavlarla derslerde işlenen müfredata göre yapılan bir uygulamanın verileridir. 2013 yılında uygulamaya başlayan TEOG sınavları, 2017-2018 eğitim-öğretim döneminde yerini başka bir sınava bırakmıştır. Bu çalışmada 20 çoktan seçmeli maddeden oluşan Türkçe testi verileri kullanılmıştır. Veriler, araştırmacının talebi ile MEB Ölçme ve Değerlendirme Sınav Hizmetleri Genel Müdürlüğü'nden yazılı izin ile alınmıştır. Türkçe alt testi çoktan seçmeli 20 madde içermektedir. Analizler öncesi, verilerin varsayımları sağlama durumu (normallik, uç değer, doğrusallık vb.) incelenmiştir (Çokluk, Şekercioğlu ve Büyüköztürk, 2012; Tabachnick ve Fidell, 2001). İlk olarak örneklem büyüklüğünün (N: 5000) analizler için yeterli olduğu değerlendirilmiştir. Ardından, veriler uç değer açısından incelenmiş veri setinde tek ve çok yönlü uç değere rastlanmamıştır. Normallik varsayımı kapsamında çarpıklık ve basıklık katsayıları incelenmiş, sırasıyla -.160 ve -.972 olarak hesaplanmıştır. Elde edilen değerlerin -1 ile +1 sınırları içerisinde yer aldığı görülmüştür. Çok değişkenli normallik varsayımı için, Bartlett Küresellik testi yapılmış, $\chi^2=15824,778$ değerinin manidar olduğu ($p<.05$) görülmüştür. Bu bağlamda verilerin çok değişkenli normallik varsayımını sağladığı ve bunun sonucu olarak değişkenler arasındaki ilişkinin doğrusal olduğu sonucuna da varılmıştır. Varsayımların incelenmesi ardından, Türkçe testi maddelerinin yapı geçerliği için doğrulayıcı faktör analizi (DFA) Mplus-8 paket programında yapılmıştır. Yapılan analiz sonucu, maddelere ilişkin standartlaştırılmış model sonuçlarına ilişkin faktör yükleri Tablo 1'de verilmiştir.

Tablo 1.

Türkçe Alt Testi Maddelerinin Doğrulayıcı Faktör Analizi Sonuçları

Maddeler	Katsayı	Standart hata	Katsayı/SH
m1	0.480	0.018	26.314*
m2	0.401	0.018	22.339*
m3	0.598	0.015	39.501*
m4	0.348	0.018	19.105*
m5	0.686	0.013	51.505*
m6	0.650	0.014	46.354*
m7	0.408	0.018	22.312*
m8	0.679	0.014	50.198*
m9	0.776	0.012	64.409*
m10	0.572	0.015	36.945*
m11	0.706	0.013	54.543*
m12	0.762	0.012	62.586*
m13	0.610	0.016	39.158*
m14	0.495	0.017	28.929*
m15	0.492	0.016	29.958*
m16	0.616	0.015	41.977*
m17	0.500	0.019	26.469*
m18	0.633	0.014	44.561*
m19	0.571	0.016	35.685*
m20	0.650	0.014	45.013*

*: 0.001 düzeyinde manidar olduğunu göstermektedir.

DFA sonucu model uyum istatistikleri (χ^2/sd : 578.771/170=3.40, RMSEA: 0.022, CFI: 0.988) değerlendirildiğinde, χ^2/sd değerinin 5'ten küçük olması yeterli uyuma (Çokluk ve diğ., 2012), RMSEA değerinin sıfıra yakınlığı ve CFI değerinin .95'in üzerinde olması ise modelin mükemmel uyum (Kline, 2005) gösterdiği şeklinde yorumlanabilir. Ayrıca Tablo 1'de de görüldüğü gibi faktör yükleri (0.348-0.776 ranjında) de .35 ve üzerindedir. Bu bağlamda, Türkçe testi maddelerinin tek boyutlu bir yapı gösterdiği ifade edilebilir.

Birinci alt amaç kapsamında, KMTK'na göre veriye en iyi uyum sağlayan modelin ve FMF'nin tespit edilmesinde Mplus-8 programından yararlanılmıştır. Örnek olması açısından iki gizil sınıflı model için KMTK modelinin sentaksı ekte verilmiştir. Mplus parametre kestiriminde, en çok olabilirlik yöntemini kullanmaktadır. KMTK'na dayalı parametre kestiriminde uygun modelin belirlenmesi için alan yazında (Li, Cohen, Kim ve Cho, 2009) önerilen Bayesian Information Criterion (BIC) değeri kullanılmıştır. İkinci alt amacı analiz etmek için gizil sınıflar arasında (analizler sonucu oluşan iki gizil gruptan biri [az öğrenci olan grup Gizil Sınıf (GS)-1] odak; diğeri referans grup (GS-2) olarak alınarak) maddelerin FMF gösterme durumunu tespit etmek için MH yöntemi kullanılmıştır. MH yöntemine göre FMF'nin tespitinde, R yazılım dilinde yer alan kütüphanelerden "difR" kullanılmıştır (Magis, Béland, Tuerlinckx ve De Boeck, 2015). Analizlerde MH yöntemi ile FMF'nin belirlenmesinde iteratif yöntem kullanılmış, 1000 iterasyon yapılmıştır. Analizler sonucu, .05 manidarlık düzeyine göre MH değerleri manidar olanlar FMF'li madde olarak değerlendirilmiştir. MH yönteminde FMF'nin büyüklüğünün belirlenmesinde 'ΔMH' değeri yorumlanmaktadır. Bu değer "0" olduğunda ve "0 ile 1 arasında" 'A: göz ardı edilebilir düzeyde', "1.0" olduğunda ve "1 ile 1.5 arasında" 'B: orta düzeyde', "1.5'e eşit ve büyük" olduğunda 'C: geniş düzeyde' FMF olduğu anlamına gelmektedir (Dorans ve Holland, 1993). MH analizinde odak grup olarak az öğrencinin yer aldığı gizil sınıf-1 alınmıştır.

Üçüncü alt amacı analiz etmek için, KMTK'na ve MH yöntemine göre gizil sınıflar arasında FMF gösteren maddelerin tutarlılık durumu incelenmiştir. Ayrıca, bireylerin seçeneklere dağılımı ve her bir gizil sınıfta maddelerin MTK'na göre madde güçlük değerleri incelenmiştir. Son olarak, her iki yönetime göre FMF olduğu saptanan maddeler için araştırmacı tarafından hazırlanan anket yoluyla beş ölçme ve değerlendirme uzmanının ve iki Türkçe öğretmenin görüşü alınmış, görüşler alıntılarla sunulmuştur. Uzmanların görüşleri "U" koduyla sunulmuştur.

3. BULGULAR

KMTK'na göre test edilen modellerden veriye en iyi uyum sağlayan modelin hangisi olduğunu belirlemek için 1-Gizil Sınıf (GS)'lı modelden, 3GS'lı modele hem iki parametrelili hem üç parametrelili modele göre farklı modeller denenmiştir. BIC değerlerine göre, 2-GS'lı 2PL-KMTK modeli en düşük BIC değerine (112442.227) ve en yüksek entropi değerine (0.691) sahiptir. Bu modele en yakın sonuç veren modellerden 2GS'lı 3PL'li KMTK'da BIC: 112505.694 ve 3GS'lı 2PL'li KMTK için BIC: 112469.609'dur. Bu bağlamda, iki gizil sınıflı modele göre ilk gizil sınıfta 1581 öğrenci vardır, yani öğrencilerin %32'si ilk gizil sınıftadır. İkinci gizil sınıfta ise 3419 öğrenci vardır. Öğrencilerin %68'i de bu sınıftadır. Sonuç olarak, iki gizil sınıflı modele göre FMF analizleri yapılmıştır. İki sınıflı modeldeki öğrenciler, sınıflardaki cinsiyet oranı açısından değerlendirildiğinde, ilk gizil gruptaki bireylerin %61'i kız, %39'u erkek iken; ikinci gizil gruptaki bireylerin %43'ü kız, %57'sinin erkek olduğu görülmüştür.

3.1. Gizil Sınıflar Arasında KMTK Yöntemine göre FMF Gösteren Maddeler

KMTK yöntemine göre yapılan FMF analizi sonuçları Tablo 2’de verilmiştir.

Tablo 2.
KMTK’na Göre FMF Sonuçları

Maddeler	Tahmin	Standart hata	Tahmin/ Standart hata	p
m2	0.500	0.286	1.752	0.080
m3	1.552	0.418	3.711	0.000*
m4	0.246	0.219	1.120	0.263
m5	0.985	0.506	1.948	0.051
m6	0.281	0.410	0.685	0.494
m7	-0.040	0.229	-0.176	0.860
m8	1.256	0.521	2.413	0.016*
m9	2.756	0.882	3.124	0.002*
m10	0.622	0.395	1.575	0.115
m11	0.247	0.435	0.568	0.570
m12	0.719	0.575	1.251	0.211
m13	1.892	0.560	3.381	0.001*
m14	0.396	0.343	1.156	0.248
m15	1.305	0.394	3.313	0.001*
m16	-0.116	0.388	-0.299	0.765
m17	-0.992	0.287	-3.455	0.001*
m18	-1.285	0.352	-3.647	0.000*
m19	-1.176	0.272	-4.323	0.000*
m20	-1.833	0.413	-4.438	0.000*

*<.05

Tablo 2’de görüldüğü gibi, KMTK’na göre dokuz madde (m3, m8, m9, m13, m15, m17, m18, m19 ve m20) .05 düzeyinde FMF içermektedir. Bu dokuz madde, öğrencilerin gizil yeteneği kontrol edildiğinde, gizil sınıflar arasında FMF içermektedir. Bu maddelerin dördü (m17, m18, m19 ve m20) gizil sınıf-2 lehine iken, beşi (m3, m8, m9, m13 ve m15) gizil sınıf-1 lehinedir.

3.2. Gizil Sınıflar Arasında MH Yöntemine göre FMF Gösteren Maddeler

KMTK’daki gizil sınıflara göre yapılan MH testi sonuçları Tablo 3’te sunulmuştur.

Tablo 3.
MH Yöntemiyle Gizil Sınıflara Göre FMF Sonuçları

Maddeler	x^2	p	Alfa MH	Δ MH	Etki Büyüklüğü
m1	103.6193	0.0000 ***	2.9188	-2.5173	C
m2	87.7900	0.0000 ***	2.9993	-2.5812	C
m3	87.5655	0.0000 ***	3.3454	-2.8379	C
m4	146.6906	0.0000 ***	3.3156	-2.8168	C
m5	1.0532	0.3048	0.8070	0.504	A
m6	0.4393	0.5074	0.9089	0.224	A
m7	70.6797	0.0000 ***	2.3292	-1.9870	C
m8	9.4492	0.0021 **	1.5414	-1.0167	B
m9	0.3481	0.5552	0.7733	0.604	A
m10	7.2800	0.0070 **	1.4466	-0.8677	A
m11	12.4612	0.0004 ***	0.6109	1.158	B
m12	7.4888	0.0062 **	0.6762	0.919	A
m13	62.7232	0.0000 ***	3.3402	-2.8341	C
m14	68.4710	0.0000 ***	2.3398	-1.9976	C
m15	134.5610	0.0000 ***	3.4440	-2.9061	C
m16	0.0042	0.9484	0.9856	0.034	A
m17	2.5076	0.1133	1.2003	-0.4290	A
m18	113.5779	0.0000 ***	0.1746	4.101	C
m19	49.5436	0.0000 ***	0.4735	1.757	C
m20	245.9831	0.0000 ***	0.0883	5.705	C

***: 0.001, **: 0.01, *: 0.05 düzeyinde manidar olduğunu göstermektedir.

Tablo 3’te görüldüğü gibi, MH yöntemine göre 13 madde (m1, i2, i3, i4, i7, i8, i11, i13, i14, i15, i18, i19 ve i20) gizil sınıflar arasında FMF göstermektedir. Bu maddelerden 11’i C düzeyinde FMF gösterirken, ikisi B düzeyinde FMF göstermektedir.

3.3. KMTK ve MH Yöntemine göre FMF Gösteren Maddeler Karşılaştırılması ve Ortak Maddeler için Uzman Görüşleri

İki yönetime göre de yedi madde de (m3, m8, m13, m15, m18, m19 ve m20), FMF tespit edilmiştir. Bu maddelerin üçü (soru 18, 19 ve soru20) ilk gizil grup lehine iken dördü (m3, m8, m13 ve m15) ikinci gizil grup lehinedir. Ayrıca yedi madde için bireylerin seçeneklere dağılımı ve her bir gizil sınıfta (GS) maddelerin MTK'na göre madde güçlükleri Tablo 4'te sunulmuştur.

Tablo 4.

Maddelere İlişkin Tepkilerin Seçeneklere Dağılımı ve MTK'ya Göre Güçlük Değerleri

Maddeler	Seçenekler	GS-1		GS-2	
		Frekans	MTK güçlük	Frekans	MTK güçlük
M3	A	121	-2.395	550	-0.601
	B*	1371		2169	
	C	28		348	
	D	19		386	
M8	A	43	-2.513	475	-0.327
	B*	1425		1969	
	C	41		652	
	D	30		358	
M13	A	29	-2.527	406	-0.868
	B*	1426		2379	
	C	75		403	
	D	8		261	
M15	A	29	-1.695	458	-0.117
	B	156		632	
	C*	1192		1778	
	D	162		581	
M18	A	12	-5.152	659	0.255
	B*	1498		1560	
	C	26		773	
	D	3		440	
M19	A	94	-4.276	931	1.929
	B*	1332		1051	
	C	85		833	
	D	27		615	
M20	A	8	-5.687	689	1.344
	B	19		724	
	C*	1495		1130	
	D	17		890	

Tablo 4'te görüldüğü gibi, ilk gizil gruptaki bireylerin çeldiricilere çok düşük oranda yöneldikleri, ancak ikinci gizil gruplardaki bireylerin daha yüksek oranda çeldiricilere yöneldiği görülmektedir. Maddelerin güçlük değerleri incelendiğinde, MTK'ya göre maddeler gizil sınıf-1'deki öğrencilere daha kolay gelmektedir. Maddeler ayrıca, konu alanı ve bilişsel düzeylerine göre incelenmiş ve yanlışlık gösterip göstermedikleri için uzman görüşüne başvurulmuştur. Uzmanların görüşlerinden çıkarılan sonuçlar ve alıntılar aşağıda sunulmuştur.

3.3.1. İlk gizil grup lehine olan maddeler için uzman görüşleri

İlk gizil grup lehine olan maddelerin ekran görüntüsü Şekil 1'de sunulmuştur (MEB, 2015).

3. Yayınevimizin - - - kitapları arasında geçmişimizdeki derin bir yaraya parmak basan bu eser var.

Bu cümlede boş bırakılan yere aşağıdaki-lerden hangisi getirilirse cümle farklı bir anlam kazanır?

- A) öne çıkan
B) kalbe dokunan
C) göze çarpan
D) dikkat çeken

8. Çocuğunuzun istediği oyuncu ya da almak istediğiniz bavulun Çin malı olmama olasılığı çok düşük. Çünkü Çin'den Anadolu'ya uzanan tarihi İpek Yolu'nda, ipek ve baharatın yerini artık çok ucuz olması nedeniyle oyuncaklar, elektronik cihazlar, müzik aletleri ve spor malzemeleri aldı. Türkiye'ye ithal edilen her 100 oyuncak ve bavulun 92'si Çin'den geliyor.

Bu metinde aşağıdaki sorulardan hangisinin cevabı yoktur?

- A) Çin'den çok ürün alınmasının sebebi nedir?
B) Çin'den ithal edilen ürünler kaliteli midir?
C) İpek Yolu ticaret ürünlerinde değişiklik olmuş mudur?
D) Çin'den en çok hangi ürünler ithal edilmektedir?

13. Zaman zaman çevrenizdeki çocukların, kesinlikle sonu gelmeyecekmiş gibi görünen, birbirinden tuhaf sorularına tesadüf etmişsinizdir. Beş altı yaşlarında olan çocukları şöyle bir düşünün, ne çok soru sorular çevrelerinde olup bitenleri anlamak için: "Bu ne, niye, ama neden, nasıl, ne zaman?" Sabırla cevap verirsiniz ve tam bittiğini sandığınız anda başa dönerler: "Bu ne, niye vb.?" Çünkü - - -

Bu metinde boş bırakılan yere düşüncenin akışına göre aşağıdakilerden hangisi getirilmelidir?

- A) soruların tamamı birbirinin aynıdır.
B) sorular hayatı ve dünyayı öğrenmelerine yardımcı olur.
C) soru sormak çocuklar için eğlenceli bir uğraştır.
D) çocuklar sorularıyla büyüklerini bıkırtırlar.

15. I. Aynı salonda daha güzel eserler sergileniyor ama hiçbirinde "Mona Lisa" kadar izdiham yaşanmıyor. Kim olduğu bile bilinmeyen bir kadının resmini mümkün olduğu kadar yakından görebilmek için herkes büyük bir çaba harcıyor.
II. Uzmanların bazıları ressamın kendisini kadın kimliğiyle resmettiğini söylerken, kimileri de Leonardo Da Vinci'nin hiçbir model kullanmadığını, bunun yerine hayalindeki bir kadını resmettiğini düşünüyor.
III. Paris'teki Louvre Müzesinin bir salonunda kalın bir cam vitrinin arkasında sergilenen "Mona Lisa" tablosunu her yıl milyonlarca ziyaretçi büyük bir hayranlıkla izliyor.

Bu paragraflarla mantık akışına göre bir metin oluşturulduğunda sıralama aşağıdakilerden hangisi olur?

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

Şekil 1. Gizil grup-1 lehine FMF'li olan maddeler

Uzmanlardan biri, "Bence okumaktan zevk alan öğrencilerin lehinedir. Çünkü tüm sorularda metine tekrar tekrar geri dönmek gerekiyor. Metni bir kez okumak yetmiyor. Okumaktan sıkılmayan öğrenci, her bir seçeneği okuduktan sonra metne döndüğünde doğru şıkkı bulacaktır." (U5) diye görüş bildirirken, diğer uzman "Soruların konu alanının okuduğunu anlama olması, maddelerin bu gruba avantaj sağlama nedeni olabilir." (U1) diyerek benzer bir görüş belirtmiştir. Diğer uzmanların görüşleri de aşağıda sunulmuştur:

Sorunun bilişsel düzeyi FMF sebebi olabilir. Çünkü ilgili maddelerin PIRLS okuduğunu anlama bilişsel sınıflamasından "Metinde açıkça anlatılan düşünceleri bulma" olabileceği düşünülmektedir. Bu noktada bu sınıflamaya göre değerlendirmek daha uygun olabilir. Çünkü birinci grup lehine olan maddeler de "kavrama" düzeyi olarak değerlendirilebilir. (U2)

İlgili soruların her biri strateji kullanımı ve akıl yürütme becerileri ile çözülebilecek sorular olduğunu düşünüyorum. Örneğin, ilk maddede metni görmeyen bir aday seçenekler arasında farklı olanı rahatlıkla ayırt edebilir. Bu da öğrenciyi doğru yanıtı yönlendirecektir. 8. soru "en çok" ifadesi metinde geçmediğinden, "yanıtı olmayan" seçenek olarak D yanıtına götürmektedir. Soru 13'te ise çünkü ifadesinin arkasından gelecek ifadenin bir açıklama niteliği taşıması gerektiği düşünülerek dahi yapılabilir. Son soruda da, metni okumanıza gerek olmadan, her bir paragrafın ilk üç sözcüğü ile ayırt etmek mümkün. Ayrıca soruların her biri kavrama/akıl yürütme/metinden çıkarımda bulunma becerilerini yoklamaktadır. (U3)

Tüm sorular, seçeneklere dayanmaktadır. Yani sorunun kökü öğrenciyi seçeneklere yönlendirmektedir. Seçeneklerin hepsini incelemeyen (15 hariç) sorunun doğru yanıtı bulunamamaktadır. Metinler de uzun olunca, sürekli geriye dönmek sorunu yaratmış olabilir. Sorunun biçimsel özelliğinden çok aslında sorunun yapısından kaynaklı olabilir. Bu soru tipi de aslında çok doğru bir soru tipi değil. Geriye dönmekte, okuduğunu tekrar okumakta sıkıntı çekmeyen, kafası karışmayan öğrenci soruları doğru yanıtlamıştır. (U4)

Birinci grupta yer alan öğrencilerin okuduğunu anlama becerileri daha yüksek olabilir. Ancak zaten bu sorular okuduğunu anlamaya yönelik. Belki var olan DIF gerçekten maddelerin ilgili özelliğe sahip olan ve olmayanları iyi ayırmasından kaynaklı olabilir. Gruplar arası gerçek farklılıktan kaynaklı olabilir. 15. soruda yer alan yabancı kelimelerin tamamı etkili olabilir. Sorunun sırası Özellikle 15 numaralı madde uzun. Ve son sıralarda yer almakta. Öğrencilerin anlamaları bazı psikolojik faktörlerden kaynaklı olarak düşebilir. Bu nokta da belki bazı psikolojik özellikler de örneğin daha geç sıkılan, daha dikkatli öğrenciler için avantajlı bir soru olmuş olabilir. (U3)

13 ve 15. maddede verilen paragrafların uzun olması, 3. madde dışında kalan maddelerde verilen paragraflardaki metin türlerinin düşünce metni olması, 3. maddenin şıklarında yer alan deyimler göz önüne alındığında maddenin sözlük bilimsel olarak dil seviyesinin zorlayıcı kategoride olması öğrencilerin bu gruplara ayrılmasına neden olmuş olabilir. (U7)

Uzmanlar, ilk gizil gruptaki maddelerin genellikle cümlede ve paragrafta anlam konularına odaklandığını, soruların bilişsel düzeyinin kavrama düzeyi olmasının FMF üzerinde etkili olabileceğini ifade ederken, okumayı seven, okumaktan sıkılmayan ve seçenek analizi yapabilen dikkatli öğrencilerin bu grupta yer alabileceğini ifade etmişlerdir. Ayrıca bir uzman, “Maddeler, sosyal yönü güçlü bir grubun lehinedir. Kitap okumayı alışkanlık edinen, resme ilgi duyan, tarih konularına meraklı, hayatı sorgulayan ve anlamlandırmaya çalışan özellikler grubun özellikleridir.” (U6) şeklinde açıklamıştır. Bu bağlamda, uzmanlara göre bu gruptaki bireyler okumayı seven ve okuma alışkanlığına sahip bireylerdir. İkinci gizil grup lehine olan maddelerin ekran görüntüsü Şekil 2’de sunulmuştur (MEB, 2015).

18. (I) Bambular daha çok tropikal bölgelerde yaşayan çiçekli bitkilerdir. (II) Bambuların bazı türleri çok hızlı büyür. (III) Öyle ki günde bir metre kadar uzayanları vardır. (IV) Bu bitkiler genellikle yıllarca büyüüp geliştikten sonra çiçek açar.

Bu metinde numaralanmış cümlelerin hangisinde fiilimsi kullanılmamıştır?

A) I. B) II. C) III. D) IV.

19. Sesin nerde kaldı, her günkü sesin
Unutulmuş güzel şarkılar için

Bu dizelerdeki fiilimsinin türce özdesi aşağıdakilerin hangisinde vardır?

- A) Bir yer var, biliyorum
Her şeyi söylemek mümkün
B) Yıllardır görmediğimiz kentin
Yürümüşük anısıyla eski cumbalı evlerin
C) Niçin oyun biterken en sonra hatırlanır
Hatıralarımızın en tatlı oyunları
D) Kaç gündür gözlerimi bile kırpmadan

Şekil 2. İkinci gizil grup lehine FMF olan maddeler

İkinci gizil grup lehine olan maddeler için uzmanların görüşleri aşağıda alıntılarla sunulmuştur. Bir uzman, “Her bir soru bir dil bilgisi konusu odaklı, dil bilgisi kurallarını kavramaya dönük sorulardır. Bu nedenle FMF oluşmuş olabilir.” (U1) derken, diğer uzman, “Maddelerin bu gruba avantaj sağlama nedeninin sorunun konu alanı olan dil bilgisi olabilir.” (U2) diye belirtmiştir. Başka bir uzman, diğer uzmana paralel olarak “Dil bilgisi kurallarını bilen öğrencilerin lehinedir. İlgili bilgileri öğrenmiş öğrenciler sorulara doğru yanıt verebilir.” (U3) demiştir. Son olarak bir uzman ise şu görüşü belirtmiştir:

Bu maddeler önceki maddelerle karşılaştırmalı olarak incelendiğinde, farklı olarak akıl yürütmeden çok bazı dil bilgisi terimlerini bilmek ve bu terimlerin verilen metinsel ifade içerisinde bulma becerisi odaklıdır. Akıl yürütme/çıkarımda bulunma gerekli değildir. Bu noktada iki grup maddelerin ayrıştığı görülebilir. (U5)

Tüm uzmanlar, ikinci gizil gruptaki maddelerin dil bilgisi konularına odaklandığını ifade etmiştir. Ayrıca, soruların bilişsel düzeyinin kavrama düzeyi olmamasının, bilgi düzeyi odaklı olmasının FMF üzerinde etkili olabileceğini de uzmanların çoğu ifade etmişlerdir. Ayrıca bir uzman “İki fiilimsi sorusunun arka arkaya gelmesi sorular arası hatırlatıcı etki yaratmaktadır (U6).” diyerek de maddelerin sıralaması ile ilgili durumun da maddeye verilen yanıtı etkileyebileceğini belirtmiştir.

4. SONUÇ, TARTIŞMA VE ÖNERİLER

Bu çalışmada, madde yanlılığının belirlenmesinde gizil sınıflara göre FMF ve kaynağını belirlemede kullanılan Karma Madde Tepki Kuramı modelinin kullanımını göstermek amaçlanmıştır. Bu amacı gerçekleştirmek için TEOG sınavı Türkçe alt testi maddelerinin yanlı olup olmadığı incelenmiştir. TEOG gibi seçme ve yerleştirme amaçlı yapılan sınavlarda, bireyler arası farklılıkları adil bir şekilde ortaya koyabilmek önem taşımaktadır. Maddelerin yanlı olma durumunun incelenmesinde öncelikle farklılaşan madde fonksiyonunun tespit edilmesi gerekir. Bu nedenle, ilk olarak maddelerin FMF gösterip göstermediği gizil gruba dayalı yöntemlerden KMTK ve gözlenen grup yöntemlerinden MH’e göre belirlenmiştir. İki yöntemle göre de FMF’li olduğu tespit edilen maddelerin yanlı olma durumu için uzman görüşüne başvurulmuştur.

KMTK’na dayalı test edilen modellerden, model-veri uyumu sağlanan modeli belirlemek için 1-Gizil Sınıf (GS)’lı modelden, 3GS’lı modele hem iki parametrelili hem üç parametrelili modele göre farklı modeller denenmiştir. Analizler sonucu, iki gizil sınıftan oluşan modelin veriye en iyi uyum sağladığı tespit edilmiştir. Bu gizil sınıflardaki bireyler, tüm testteki başarısı açısından değerlendirildiğinde, ilk gizil gruptaki bireylerin oldukça başarılı olduğu, bu öğrencilere maddelerin daha kolay geldiği, ikinci gizil gruptaki bireylerin ise daha çok zorlandığı görülmüştür. Alan yazında, gizil gruplara göre FMF incelemesi yapan çalışmalarda da bu çalışmada elde edilen bulgulara paralel olarak genellikle iki gizil sınıflı modelin veriye daha iyi uyum sağladığı ve başarı durumlarına göre bireylerin sınıflarda gruplandığı görülmüştür (Choi et al., 2015; Yalçın, 2018).

İki gizil sınıflı modele göre KMTK analizi FMF için yapıldığında, dokuz maddenin FMF içerdiği tespit edilmiştir. Bu dokuz madde, öğrencilerin gizil yeteneği kontrol edildiğinde, gizil sınıflar arasında FMF içermektedir. Bu maddelerin dördü gizil sınıf-2 lehine iken, beşi gizil sınıf-1 lehinedir. MH yöntemine göre 13 madde gizil sınıflar arasında FMF göstermektedir. Ayrıca bu 13 maddeden dördü gizil sınıf-2'nin lehinedir. MH ve KMTK sonuçları karşılaştırıldığında, MH yöntemine göre daha fazla maddenin FMF olarak tespit edildiği görülmüştür. Alan yazında yapılan benzer bir araştırmada bu bulguyu desteklemektedir (Yalçın, 2018). Ayrıca, yapılan simülasyon çalışmalarında da (Maij-de Meij, 2010; Uyar ve diğ., 2017), FMF'yi tanımlamada gizil grup yaklaşımına dayalı KMTK'nın performansının bilinen değişkenleri kullanan FMF belirleme yöntemlerinden daha iyi olduğu görülmüştür. Ancak simülasyon çalışmaları ile yapılacak karşılaştırmalara göre yöntemlerin duyarlılığı ve tutarlılığı hakkında daha doğru yorumlar yapılabilir. İlgilenen araştırmacılara, bu konuda simülasyon modelleri denemeleri önerilmektedir.

Bu çalışmada, alan yazında sıklıkla kullanılan cinsiyete göre FMF incelemesi yerine gizil gruplara göre FMF incelemesi yapılmıştır. Bu nedenle ayrıca gizil grupları tanımlamada cinsiyet dağılımı da göz önünde bulundurulmuştur. İlk gizil gruptaki bireylerin %61'i kız, %39'u erkek iken; ikinci gizil gruptaki bireylerin %43'ü kız, %57'si erkek olduğu görülmüştür. Bu bağlamda, gizil grupların cinsiyetten bağımsız olduğu sonucuna varılmıştır. Bu bulgu, alan yazındaki gizil gruba göre FMF incelemesinde, KMTK kullanan farklı çalışmalarda da görülmektedir (Cohen ve Bolt, 2005; Tay, Newman ve Vermunt, 2011; Yalçın, 2018). Diğer yandan, Türkçe alt testinde cinsiyete göre yanlılık ve/veya FMF incelemesi yapılan farkı yıllardaki ortaöğretimden liseye geçiş sınavları bağlamında yapılan çalışmalarda da, Türkçe alt testi maddelerin cinsiyete göre FMF ve/veya yanlılık içermediği görülmüştür (Kelecioğlu, Karabay ve Karabay, 2009; Satıcı ve Özer-Özkan, 2017; Türkan, 2014; Yurdugül ve Aşkar, 2004). Bu çalışmaya göre elde edilen bulgular birlikte değerlendirildiğinde, cinsiyet dışında bilinmeyen gruplara göre örneğin maddelerin konu alanı, içeriğine göre FMF incelemesi yapıldığında FMF tespit edildiği görülmektedir. Bu nedenle, FMF incelemelerinde gizil gruplara dayalı incelemelerin de yapılması önerilmektedir. Çünkü bir maddenin, bilinen gruplara göre FMF göstermesi durumunda, maddenin ilgili alt gruptaki tüm bireylere avantaj sağladığı veya dezavantaja neden olduğunu söylemek zordur. Yani, bilinen gruplar her zaman grupların homojenliği varsayımını sağlamamaktadır (De Ayala ve diğ., 2002; De Mars ve Lau, 2013). Bu çalışmada, hazır veriler üzerinden çalışıldığı için araştırmanın sınırlılığı olarak bireylere ilişkin özelliklerden sadece cinsiyet ele alınmıştır. Bu bağlamda, MEB'den cinsiyet dışında bireylere ait özelliklere ilişkin veriler alınmadığından daha ayrıntılı incelemeler yapılamamıştır. Sonraki çalışmalarda, bireylere ilişkin başka özellikler ile de gizil sınıfların ilişkileri ve gizil sınıfları açıklama durumları incelenebilir.

Her iki yönteme göre FMF olduğu saptanan yedi madde için gizil gruptaki bireylerin seçeneklere dağılımı incelendiğinde, ilk gizil gruptaki bireylerin çeldiricilere çok düşük oranda yöneldikleri, ancak ikinci gizil gruptaki bireylerin daha yüksek oranda çeldiricilere yöneldiği tespit edilmiştir. Her iki gizil grupta, maddelerin güçlük değerleri incelendiğinde, maddelerin gizil sınıf-1'deki öğrencilere daha kolay geldiği görülmüştür. Yani, ilk gizil gruptaki bireylerin oldukça başarılı olduğu, bu nedenle öğrencilere maddelerin daha kolay geldiği, ikinci gizil gruptaki bireylerin ise daha çok zorlandığı, soruların daha zor geldiği görülmüştür. Bu bağlamda, ilk gruptaki başarılı bireylerin soruların cevaplarını bilen öğrencilerden oluştukları için çeldiricilere yönelme durumları da düşüktür. İkinci gruptaki bireyler ise daha başarısız öğrencilerden oluştuğu için yani soruların cevaplarını net olarak bilmedikleri için çeldiricilere daha fazla yönelmişlerdir. Maddelerin ayrıca, konu alanı ve bilişsel düzeyi tespit edilmiş yanlılık içermeye durumu için uzman görüşüne başvurulmuştur.

Uzmanlar, ilk gizil gruptaki maddelerin genellikle cümlede ve paragrafta anlam konularına odaklandığını, soruların bilişsel düzeyinin kavrama düzeyi olmasının FMF üzerinde etkili olabileceğini ifade ederken, okumayı seven, okumaktan sıkılmayan ve seçenek analizi yapabilen dikkatli öğrencilerin bu grupta yer alabileceğini ifade etmişlerdir. Bu bağlamda, uzmanlara göre bu gruptaki bireyler okumayı seven ve okuma alışkanlığına sahip bireylerdir. Alan yazında yapılan bazı çalışmalarda (Yalçın & Tavşancıl, 2015; Yalçın, 2017), öğrencilerin bir derse yönelik duyuşsal özelliklerinin (sevme, kendine güvenme) o derse ilişkin maddeleri doğru yanıtlama davranışlarıyla ilişkili olduğu tespit edilmiştir. Bu çalışmada, önceki yıllarda yapılan ulusal düzeyde bir sınav verileri kullanıldığı için öğrencilerle görüşme olanağı veya okumayı sevmelerine yönelik bir anket ya da tutum ölçeği kullanılamamıştır. Sonraki araştırmalarda, bu durum göz önünde bulundurularak araştırma kurgusu tasarlanması önerilir. Uzmanların tamamı, ikinci gizil gruptaki maddelerin dilbilgisi konularına odaklandığını belirtmiştir. Ayrıca uzmanların çoğu, soruların bilişsel düzeyinin kavrama düzeyi olmamasının, bilgi düzeyi odaklı olmasının FMF üzerinde etkili olabileceğini ifade etmişlerdir. Alan yazında da, gizil grup yaklaşımına göre FMF incelemesi yapan bazı çalışmalarda, bu çalışmadan elde edilen bulguları destekleyen, maddelerin konu alanının FMF ile ilişkili olabileceği yönünde bulgular elde edilmiştir (Cohen ve Bolt, 2005; Yalçın, 2018). Bu bağlamda, FMF'nin kaynağının tespitinde uzman görüşleri yanında FMF'li maddeler için öğrenciler ile yapılacak sesli düşünmeye (thinking aloud) dayalı görüşmeler de maddelerin FMF içermeye durumuna ilişkin değerli bilgiler sunacaktır. Bu çalışmadan elde edilen bulgular doğrultusunda, araştırmacılara ve madde yazarlarına, aynı yetenek düzeyindeki öğrenciler arasındaki başarı farklılıklarının nedeninin öğrencinin o konuya ilişkin tutumu, özgüveni vb. ve/veya konu alanı (paragrafta anlam veya dilbilgisi gibi) olabileceği belirtilebilir. Ayrıca uzmanların belirttiği, Gizil sınıf-1'de yer alan öğrencilerin seçenek analizi yapabilen öğrencilerden oluştuğu yönündeki tespitinin, bu gruptaki öğrencilerin çok düşük oranda çeldiricilere yönelmiş olmaları da işaret etmektedir. Ancak bu konuda ileriye yönelik çıkarımlar yapabilmek için başka araştırma bulgularıyla sonuçların desteklenmesine ihtiyaç vardır.

Türkçe testinde yer alan tüm maddeler incelendiğinde, KMTK'na göre FMF olarak belirlenen maddelerin dilbilgisi ve okuduğunu anlama boyutlarıyla daha tutarlı olduğu görülmüştür. Sonuçlar KMTK yaklaşımına göre FMF belirlemenin etkili bir yöntem olduğunu göstermektedir. Yapılan simülasyon çalışmalarında da (Maij-de Meij, 2010; Uyar ve diğ., 2017), FMF'yi tanımlamada

KMTK'nın performansının bilinen değişkenleri kullanan FMF belirleme yöntemlerinden daha iyi olduğu görülmüştür. Bu bağlamda, FMF incelemelerinde KMTK yaklaşımının kullanılması önerilmektedir.

Araştırma ve Yayın Etiği Beyanı

Araştırmanın tüm sürecinde etik ilke ve kurallara uyulmuştur.

Yazarların Makaleye Katkı Oranları

Yazar, makaleyi tek başına hazırlamıştır.

Çıkar Beyanı

Yazarın herhangi bir çıkar çatışması bulunmamaktadır.

5. KAYNAKÇA

Camilli, G., & Shepard, L. A. (1994). *MMSS: Methods for identifying biased test items*. Thousand Oaks, CA: Sage.

Chaimongkol, S., Huffer, F. W. & Kamata, A. (2007). An explanatory differential item functioning (DIF) model by the WinBUGS 1.4. *Songklanakarın Journal of Science and Technology*, 29, 449–458.

Choi, Y., Alexeev, N., & Cohen, A. S. (2015). Differential item functioning analysis using a mixture 3-parameter logistic model with a covariate on the TIMSS 2007 mathematics test. *International Journal of Testing*, 15(3), 239-253. doi: 10.1080/15305058.2015.1007241.

Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*, 17(1), 31-44. doi: 10.1111/j.1745-3992.1998.tb00619.x

Cohen, A. S., & Bolt, D. M. (2005). A mixture model analysis of differential item functioning. *Journal of Educational Measurement*, 42(2), 133–148. doi: 10.1111/j.1745-3984.2005.00007.

Çokluk, Ö. S., Şekercioğlu, G. & Büyüköztürk, S. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve lisrel uygulamaları*. Ankara: Pegem Akademi.

De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford Press.

De Ayala, R. J., Kim, S. H., Stapleton, L. M., & Dayton, C. M. (2002). Differential item functioning: A mixture distribution conceptualization. *International Journal of Testing*, 2(3&4), 243-276. doi: 10.1080/15305058.2002.9669495.

De Mars, C. E., & Lau, A. (2011). Differential item functioning detection with latent classes: How accurately can we detect who is responding differentially? *Educational and Psychological Measurement*, 71(4), 597–616. doi: 10.1177/0013164411404221.

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel haenszel and standardization. In P. W. Holland, and H. Wainer, (Eds.), *Differential item functioning* (p. 35–66), New Jersey: USA.

Finch, W. H. & French, B. F. (2012). Parameter estimation with mixture item response theory models: A monte carlo comparison of maximum likelihood and bayesian methods. *Journal of Modern Applied Statistical Methods*, 11(1), 167-178. doi: 10.22237/jmasm/1335845580.

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the mantel-haenszel procedure. In H. Wainer, and H. I. Brown (Eds.), *Test validity* (p. 129-145). Hillsdale, NJ: Lawrence Erlbaum Associates.

Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum.

Kelderman, H., & Macready, G. B. (1990). The use of loglinear models for assessing differential item functioning across manifest and latent examinee groups. *Journal of Educational Measurement*, 27(4), 307–327.

- Kelecioğlu, H., Karabay, B. ve Karabay E. (2014). Seviye belirleme sınavının madde yanlılığı açısından incelenmesi. *İlköğretim Online*, 13(3), 934-953.
- Kline, R. B. (2005). *Methodology in the social sciences. Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.
- Li, F., Cohen, A. S., Kim, S., & Cho, S. (2009). Model selection methods for mixture dichotomous IRT models. *Applied Psychological Measurement*, 33(5), 353-373. doi: 10.1177/0146621608326422.
- Magis, D., Béland, S., Tuerlinckx, F., & De Boeck, P. (2015). A general framework and an R package for the detection of dichotomous differential item functioning. *Behavior Research Methods*, 42(3), 847-862. doi:10.3758/BRM.42.3.847.
- Maij-de Meij, A. M., Kelderman, H. & van der Flier, H. (2010). Improvement in detection of differential item functioning using a mixture item response theory model. *Multivariate Behavioral Research*, 45(6), 975-999. doi:10.1080/00273171.2010.533047.
- Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research*, 13(2), 127-143. doi: 10.1016/0883-0355(89)90002-5.
- MEB. (2015). *TEOG 8. sınıf, 1. dönem merkezi ortak sınavı soruları*. [Çevrim-içi: <http://www.meb.gov.tr/25-26-kasim-2015-8-sinif-1-donem-merkezi-ortak-sinavin-soru-ve-cevap-anahtarlar/duyuru/9978>], Erişim tarihi: 20.12.2017.
- Ong, Y. M., Williams, J., & Lamprianou, I. (2013). Exploring differential bundle functioning in mathematics by gender: The effect of hierarchical modelling. *International Journal of Research & Method in Education*, 36(1), 82-100. doi: 10.1080/1743727X.2012.675263.
- Osterlind, S. J., & Everson, H. T. (2009). *Differential item functioning*. Los Angeles: Sage.
- Samuelsen, K. M. (2008). Examining differential item functioning from a latent mixture perspective. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (p. 177-197). Charlotte, NC: Information Age.
- Satıcı, D. K. ve Özer-Özkan, Y. (2017). Temel eğitimden ortaöğretime geçiş sınavının (2014-kasım) cinsiyet açısından madde yanlılığının incelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 13(1), 254-274.
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics*. Needham Heights, MA: Allyn & Bacon.
- Tay, L., Newman, D. A., Vermunt, J. K. (2011). Using mixed-measurement item response theory with covariates (MM-IRT-C) to ascertain observed and unobserved measurement equivalence. *Organizational Research Methods*, 14, 147-176. doi:10.1177/1094428110366037.
- Uyar, S., Kelecioğlu, H., & Doğan, N. (2017). Comparing differential item functioning based on manifest groups and latent classes. *Educational Sciences: Theory & Practice*, 17(6), 1977-2000. doi: 10.12738/estp.2017.6.0526.
- Türkan, A. (2014). *2012- Seviye belirleme sınavının rasch modeline göre cinsiyet değişkeni açısından yanlılığının incelenmesi*. Yayınlanmamış yüksek lisans Tezi. Gaziantep Üniversitesi Eğitim Bilimleri Enstitüsü, Gaziantep.
- von Davier, M., & Rost, J. (2017). Logistic mixture-distribution response models. In W. J. van der Linden (Ed.), *Handbook of item response theory, volume one: Models* (p. 393-406). Boca Raton: Chapman and Hall/CRC.
- Wilson, M. & De Boeck, P. (2004). Descriptive and explanatory item response models. In P. De Boeck and M. Wilson (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 43-74). New York: Springer-Verlag.
- Yalçın, S., & Tavşancıl, E. (2015). The factors explaining the differential item functioning in the administration of TIMSS 2011 science test according to gender. *Educational Sciences and Practice*, 14(27), 1-21.
- Yalçın, S. (2017). The effect of latent classes formed according to the affective characteristics of students on differential item functioning based on sex. *Elementary Education Online*, 16(4), 1917-1931.

Yalçın, S. (2018). Determining differential item functioning with the mixture item response theory. *Eurasian Journal of Educational Research*, 74, 187-206.

Yurdugül, H. ve Aşkar, P. (2004). Ortaöğretim kurumları öğrenci seçme ve yerleştirme sınavının cinsiyete göre madde yanlılığı açısından incelenmesi. *Eğitim Bilimleri ve Uygulama Dergisi*, 3(5), 3-20.

6. EXTENDED ABSTRACT

Transition from Basic Education to Secondary Education (TEOG) exam started to be performed in 2013 and gave its place to the Transition to High School System in the 2017-2018 school year. It is of vital importance to be able to show the differences between individuals in a fair way in the exams conducted for the purposes of selection and placement such as TEOG. In this context, the items that constitute this exam should not provide advantage or cause disadvantage for a group of students, meaning that they should not involve bias. It is necessary to determine whether there is a differential item functioning (DIF) in order to investigate whether the items that constitute a test are biased. Recent studies conducted in the field of DIF have shown that the causes of the DIF are generally complex and not directly related to defined groups (Cohen & Bolt, 2005; De Mars and Lau, 2013). In this context, it is emphasized that the DIF should be examined among the latent (unknown) groups (Cohen & Bolt, 2005; De Ayala et al., 2002; De Mars and Lau, 2013; Finch and French, 2012; Maij-de Meij et al., 2010; Samuelsen, 2008). It is aimed in this study to identify whether the items in the TEOG exam are biased. In this regard, the state of items' displaying differential item functioning is analyzed primarily according to the Mixed Item Response Theory (MixIRT) model which is based on the latent class method and the Mantel- Haenszel (MH) method based on the observed group approach. The items that are determined to display DIF according to both methods are analyzed with regards to subject field and cognitive level, as well as taking expert opinion on including bias.

Determining the DIF according to two different DIF determination methods and later on taking expert opinions regarding the items that are determined as DIF according to the two methods make it possible to describe the existing situation. 5000 students who were randomly selected from among the students that took the TEOG 2015 Turkish test constitute the sample of this study. In order to analyze the unidimensionality of the items in the Turkish test, confirmatory factor analysis is conducted in the Mplus 8 package. It is seen as a result of the analysis that the items showed a unidimensional construct when the item fit statistics are evaluated. The Mplus-8 package is used in determining the appropriate model based on MixIRT and the DIF (Muthén and Muthén, 2017). The "difR" library in the R software language is employed in determining the DIF according to the MH method (Magis, Béland, Tuerlinckx and De Boeck, 2015).

Among the models based on MixIRT, the model with two latent classes is determined to fit the data best. When the students in the two-class model were evaluated in terms of gender ratio, it is seen that 61% of the individuals in the first latent group were female and 39% were male; while 43% of the individuals in the second latent group were female and 57% were male. It is seen that nine out of 20 Turkish items include DIF after conducting DIF analyses according to the MixIRT model. It is seen that 13 items demonstrate DIF when MH test is carried out according to the latent classes in the MixIRT. Being in common with the two methods, seven items indicate DIF. Among these items, four of them are advantageous for the first latent group while three of them are advantageous for the second latent group. The items are examined in terms of subject field and cognitive level and expert opinion is taken regarding the state of having bias. When the items that indicate DIF according to both methods are analyzed in the DIF examination conducted based on the two latent classes that emerged according to the MixIRT, it is observed that three items which are all advantageous for the second latent group, are knowledge level questions measuring the Turkish grammar subject. Other four items in favor of the first latent class were seen to be understanding level questions which include the subjects of meaning in sentences and paragraphs.

Moreover, the distribution of individuals to choices is examined for seven items. It is seen that individuals in the first latent group have a very low ratio of going towards the distractors, while the individuals in the second latent group tend to have higher rates of going for distractors. As a result of the analyzes, five measurement and evaluation experts' and two teachers' opinions were taken via the questionnaire prepared for seven items which are determined to be with DIF according to both methods. The experts expressed that the questions in favor of the latent class-2 were about gerundial and the elements of the sentence; meaning the grammar subjects, whereas the questions in favor of the latent class-1 were about meaning in sentence and paragraphs. In this regard, it can be put forward that item features are beneficial for some groups and disadvantageous for some others. The results indicate that determining the DIF is an effective method according to the MixIRT approach. In this context, it is recommended to use the MixIRT approach in DIF examinations.

7. EKLER

7.1. İki Gizil Sınıflı Model için KMTK Sentaksı (MPLUS)

```
TITLE: İki sınıflı model için KMTK
DATA: FILE IS turkce.dat;
VARIABLE: NAMES = u1-u20;
USEV ARE u1-u20;
CATEGORICAL = u1-u20;
CLASSES = c (2);
ANALYSIS: TYPE = MIXTURE;
ALGORITHM = INTEGRATION;
INTEGRATION = STANDARD (20);
STARTS = 300 30;
PROCESS = 2;
MODEL: %OVERALL%
      f BY u1-u20;
      %c#1%
      [u2$1-u20$1];
      f@1;
      %c#2%
      [u2$1-u20$1];
      f@1;
OUTPUT: TECH8;
STANDARDIZED;
```