

## Çoktan Seçmeli Testlerde İki Kategorili ve Önsel Ağırlıklı Puanlamanın Diferansiyel Madde Fonksiyonuna Etkisi ile İlgili Bir Araştırma

### The Effects of Binary and Priori Weighted Scoring on the Differential Item Functioning Multiple Choice Tests

Devrim Özdemir  
Hacettepe Üniversitesi

#### Öz

Bu araştırmanın amacı, iki kategorili ve önsel ağırlıklı puanlama yöntemleri kullanılarak puanlanan bir testte, cinsiyet grupları ve okul türüne göre diferansiyel madde fonksiyonunu ortaya koymaktır. Çeşitli illerdeki 7'si özel, 8'i devlet okulu olmak üzere 15 ilköğretim okulundan seçilen 780'i erkek, 698'i kız öğrenciden elde edilen madde puanları, MH test istatistiği ile DMH katsayısına ve lojistik regresyon modeline göre analiz edilmiştir. Hem okul türü hem de cinsiyet bakımından maddelerin çoğunda göz ardı edilebilir düzeyde DMF olduğu, bir maddede ise okul türü bakımından özel okullarda okuyan öğrenciler lehine, cinsiyet bakımından da erkek öğrenciler lehine DMF ortaya çıktığı belirlenmiştir.

**Anahtar sözcükler:** Diferansiyel madde fonksiyonu, madde yanlılığı, iki kategorili puanlama, önsel ağırlıklı puanlama.

#### Abstract

The purpose of this study is to determine the differential item functioning of a test scored via binary and priori weighting scoring methods in terms of gender and school type. The item scores collected from 780 male and 698 female students chosen from 15 elementary schools of 7 private and 8 public schools were analyzed using DMH coefficient and MH test statistics and logistic regression models. The research findings showed that most of the items have ignorable DIF in terms of both gender and school type. Only in one item did DIF favor the students attending private schools in terms of school type and female students in terms of gender.

**Key words:** Differential item functioning, item bias, binary scoring, priori weighted scoring.

#### Giriş

Bilindiği gibi, test uygulamalarının temel amacı, bireyler ya da eğitim programları hakkında bilgi edinmektir. Bunun için hatalardan arınık ve geçerliliği yüksek ölçme araçlarına/sonuçlarına ihtiyaç vardır. Ancak, geçerliliği olumsuz yönde etkileyen faktörlerden biri de “yanlı” maddelerdir (biased item). Yanlılık, madde parametrelerinin kestirim değeri ile gerçek değeri arasındaki farklılık olarak ifade edilir ve ölçme sonuçlarına “sistemik hata” olarak karşır. Bu durum ise ölçme aracının yapı ve yordama geçerliklerini düşürür (Osterlind, 1983).

Eğitim alanındaki yanlılık genellikle, testlerin uygulandığı alt gruplardan kaynaklanmaktadır. Madde yanlılığı, aynı yetenek düzeyinde yer almasına karşın, farklı alt gruplara ait bireylerin bir test maddesini doğru yanıtlama olasılıklarındaki farklılık olarak tanımlanabilir (Zumbo 1999).

Yanlılık kavramı genel bir kavramdır. Yanlılığın operasyonel olarak araştırılabilmesi için diferansiyel madde fonksiyonundan (DMF, differential item function) yararlanır. DMF, maddenin ölçmek istediği her bir yetenek düzeyindeki maddeyi doğru yanıtlama olasılıklarının alt gruplara göre farklılıkları gösteren bir bağıntıdır. Bir diğer ifade ile DMF, yanlılığın bir göstergesidir.

Günümüze kadar DMF üzerine bir çok çalışma yapılmıştır. Bu çalışmaların temeli DMF'yi ortaya çıkarma

Dr. Devrim Özdemir, Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, e-posta: devrimo@hacettepe.edu.tr

üzerindedir. Yapılan ilk çalışmalar, DMF'nin tanımı gereği, klasik test kuramında madde güçlük indekslerinin karşılaştırılması biçiminde olmuş; A ve B alt gruplarının i. madde üzerindeki farklılıkları istatistiksel olarak test edilmiştir (Osterlind, 1983):  $H_0: P_{IA} - P_{IB} = 0$

Benzer yaklaşımla, örtük özellikler kuramında madde güçlük parametrelerinin karşılaştırılması yapılmıştır (Hambleton, Swaminathan ve Rogers, 1991):

$$H_0: b_{IA} - b_{IB} = 0$$

Ancak bu çözümler, DMF türlerini ve DMF'nin büyüklüğü nicel olarak ortaya koyamamaktadır.

#### DMF Türleri

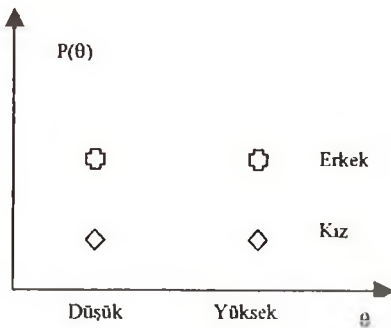
DMF, öncelikle Tekbiçimli DMF (TB-DMF) ve Tekbiçimli Olmayan DMF (TBO-DMF) olmak üzere iki ayrı grupta ele alınmaktadır (Camilli ve Shepard, 1994).

Alt grup olarak cinsiyet ele alındığında (kategoriler: kız-erkek) kategorilerden birinin diğerine göre üstünlüğü değişik yetenek düzeylerinde de devam ediyorsa, böylesi durumlar TB-DMF olarak adlandırılmaktadır (Camilli ve Shepard, 1994; Yu, 2001).

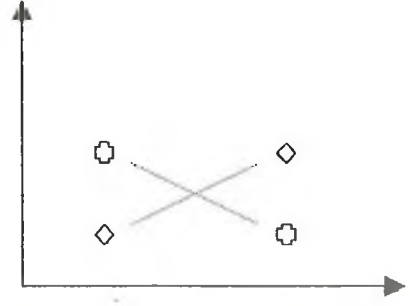
Örneğin düşük yetenek düzeylerinde kızların performansı erkeklere göre daha yüksek iken, aynı durum yüksek yetenek düzeylerinde de geçerli ise, bu durumda TB-DMF'den sözedilebilir. Tam tersi durumda –düşük yetenek düzeyinde kategorilerden birisi üstün, yüksek yetenek düzeyinde ise diğeri üstün- ise TBO-DMF olduğu ifade edilir (Şekil 1 ve 2).

Bu DMF türleri dışında, simetrik ve asimetrik olarak adlandırılan farklı DMF türleri de vardır.

DMF'nin bu kadar çeşitli olması DMF'nu ortaya çıkarma yöntemlerinin de çok sayıda olmasını beraberinde



Şekil 1. İki kategorili alt gruplara ilişkin TB-DMF. (Kesişim yok)



Şekil 2. İki kategorili alt gruplara ilişkin TBO-DMF. (Kesişim var)

de getirmiştir. Bu yöntemlerden bazıları, Mantel-Haenszel (MH) test istatistiği, SIBTEST (simultaneous item bias test), log-linear ve lojistik regresyon modelleri olarak sıralanabilir. Bu yöntemlerden her birinin diğerlerine göre üstün ya da zayıf yönleri bulunmaktadır. Örneğin; MH yöntemi TB-DMF için duyarlı sonuçlar verirken (Hambleton ve Rogers, 1989), TBO-DMF için lojit yöntemler daha duyarlıdır (Fidalgo, Mellenberg ve Muñiz, 2000).

#### Mantel-Haenszel (MH) Yöntemi

Bir test istatistiği olan MH, aynı zamanda da odds oranının (odds ratio) en çok olabilirlik kestiricisidir.

MH test istatistiği, kategorik değişkenlerin sıklıklarının yer aldığı olumsuzluk çizelgeleri ile çalışmaktadır.

2x2'lik bir olumsuzluk çizelgesi için MH test istatistiği (Osterlind, 1983),

$$MH = \hat{\Omega} = \frac{N_{11}N_{22}}{N_{21}N_{12}}$$

şeklinde gösterilebilir. MH test istatistiği alt gruplar arasında bir karşılaştırma sonucu verir. Ele alınan her madde için,  $MH \equiv 1$  ise maddede DMF yoktur;  $MH > 1$  ise madde referans olarak alınan grup lehine çalışmıştır;  $MH < 1$  ise madde diğer grup lehine çalışmıştır.

Bu sonucun daha iyi yorumlanabilmesi ve DMF'nin nicel büyüklüğü hakkında bilgi vermesi için  $\Delta_{MH}$  dönüşümü yapılır.

$$\Delta_{MH} = -(2,35) \cdot \ln(MH)$$

$\Delta_{MH}$  dönüşümünden sonra elde edilen değerlere göre;  $\Delta_{MH}=0$  ise, maddede DMF yoktur;  $\Delta_{MH}<0$  ise madde

	+	-	Toplam
1.Grup	$N_{11}$	$N_{12}$	$N_{+1}$
2.Grup	$N_{21}$	$N_{22}$	$N_{+2}$
Toplam	$N_{1+}$	$N_{2+}$	$N$

---

	+	-	Toplam
1.Grup	$P_{11}$	$P_{12}$	$P_{+1}$
2.Grup	$P_{21}$	$P_{22}$	$P_{+2}$
Toplam	$P_{1+}$	$P_{2+}$	1

Şekil 3. 2X2 Olumsuzluk çizelgesi ve iki değişkenli olasılık dağılımları

referans olarak alınan grup lehine çalışmıştır;  $\Delta_{MH} > 0$  ise madde diğer grup lehine çalışmıştır.  $|\Delta_{MH}|$  ise, DMF'nin düzeyi hakkında bilgi vermektedir.  $|\Delta_{MH}| < 1$  ise DMF yoktur ya da göz ardı edilebilecek bir DMF söz konusudur (A-level);  $1 \leq |\Delta_{MH}| < 1,5$  ise orta düzeyde bir DMF vardır (B-level);  $|\Delta_{MH}| \geq 1,5$  ise yüksek düzeyde bir DMF olduğu ifade edilir (C-level) (Gierl, Jodoin ve Akkerman, 2000).

#### Lojistik Regresyon Yöntemi

Bireylerin, herhangi bir maddeyi doğru yanıtlama olasılığı için geçerli bağıntı aşağıdaki gibidir (Whitmore ve Schumaker, 1999);

$$\Psi(Z) = \frac{e^Z}{1 + e^Z}$$

Yukarıdaki bağıntıda Z,

$$Z = \beta_0 + \beta_1 q + \beta_2 g$$

şeklinde ifade edilir. Bu eşitlikte,  $\beta_1$ : Yetenek katsayısı ve  $\beta_2$ : Grupların etki katsayısı olarak adlandırılır. G değişkeni, birincil grup için 1, ikincil grup için 0 şeklinde kodlanmış bir gösterge değişkendir. g değişkeni için  $\beta_2$ 'nin modele katkısı incelendiğinde, elde edilen sonuç anlamlı ise olası bir DMF'nin varlığından söz edilebilir. Ancak bu şekildeki DMF için TB-DMF olduğu söylenebilir. TBO-DMF için tam regresyon modeline ihtiyaç vardır. Bu model,

$$Z = \beta_0 + \beta_1 q + \beta_2 g + \beta_3 \theta * g$$

şeklinde gösterilebilir. Yukarıdaki eşitlikte  $\beta_3$ : yetenek ve grup değişkenlerinin etkileşim katsayısıdır.  $\beta_3$  katsa-

yısının modele katkısı anlamlı bulunursa bu durumda TBO-DMF olduğu söylenebilir (Whitmore ve Schumaker, 1999).

MH ve lojistik regresyon modelleri iki kategorili (0-1 puanlama) durumlarında kullanılacak yöntemlerdir. Maddelerin çoklu kategoriye göre puanlanmış biçimlerinde ise lojistik regresyonun özel bir hali kullanılabilir (Zumbo, 1999).

Zumbo (1999), iki kategorili puanlanmış testler için lojistik modeli;

$$\ln \left[ \frac{P_i}{1-P_i} \right] = \beta_0 + \beta_1 TP + \beta_2 \text{Grup} + \beta_3 \text{Grup} * TP$$

ve çok kategorili puanlanmış testler için lojistik modeli;

$$\log \left[ \frac{P(Y \leq j)}{P(Y > j)} \right] = \beta_0 + \beta_1 TP + \beta_2 \text{Grup} + \beta_3 \text{Grup} * TP$$

şeklinde tanımlamıştır. Modeldeki test puanı TP ile gösterilmiştir. j ise, c kategori sayısını göstermek üzere  $j=1,2,\dots,(c-1)$  olarak verilmiştir.

#### Araştırmanın Amacı

Bu araştırmanın temel amacı, çoktan seçmeli testlerde iki kategorili ve önsel ağırlıklı puanlamanın DMF'ye etkisini incelemektir.

#### Problem

Türkçe okuduğunu anlama yeteneğini ölçme amacı ile geliştirilmiş çoktan seçmeli bir testte, iki kategorili ve önsel ağırlıklı puanlama yöntemlerini kullanmanın DMF'ye etkisi nasıldır?

#### Alt Problemler

1. İki kategorili ve önsel ağırlıklı puanlama yöntemleri ile elde edilen madde puanları, DMF bakımından okul türüne göre yanlılık göstermekte midir?
2. İki kategorili ve önsel ağırlıklı puanlama yöntemleri ile elde edilen madde puanları, DMF bakımından cinsiyete göre yanlılık göstermekte midir?

*Sayıtlar*

1. Uygulamanın yapıldığı tüm sınıflarda test, yönergeye uygun bir biçimde uygulanmıştır.
2. Öğrenciler test maddelerini cevaplarırken gerçek bilgilerini yansıtmışlardır.

*Sınırlılıklar*

1. Araştırma, uygulamanın yapıldığı ilköğretim okullarının 4, 5, 6 ve 7. sınıflarında okuyan 1478 öğrenciyle sınırlıdır.
2. Araştırma, Türkçe okuduğunu anlama yeteneği ile sınırlıdır.
3. Araştırma, cinsiyet ve okul türü alt grupları ile sınırlıdır.

*Yöntem*

Araştırmanın verileri, 2001-2002 öğretim yılında çeşitli ilköğretim okullarının 4., 5., 6. ve 7. sınıflarında okuyan 1478 öğrenciye uygulanan ve Türkçe okuduğunu anlama yeteneğini ölçme amacı ile araştırmacı tarafından geliştirilen, her biri 4 seçenekli 20 maddelik çoktan seçmeli test yardımı ile elde edilmiştir.

Araştırmada iki kategorili puanlama yöntemi ile önsel ağırlıklı puanlama yönteminden yararlanılmıştır. İki kategorili puanlamada her bir maddeye verilen cevaplar, doğru cevap için 1; yanlış, boş ya da birden fazla cevaplar için 0 puan verilmek suretiyle puanlanmaktadır. Önsel ağırlıklı puanlama yöntemi, her bir maddeye ilişkin seçeneklerin, doğru cevaba yakınlıklarına göre, bir ağırlık atanmak suretiyle puanlanmasını ifade eder. Bu yöntemde seçeneklere verilecek ağırlıkların belirlenmesi, testin uygulanmasından önce konu alanı uzmanları tarafından yapılmaktadır. Bu yöntemde doğru cevap dışındaki seçeneklere de puan atandığından, maddelerin yazılmasında "en doğru cevabı gerektiren" soru formatı kullanılmıştır.

Testin uygulanması ile elde edilen veri, okul türüne (özel okullar ve devlet okulları) ve cinsiyete göre (erkek-kız) düzenlenmiştir. DMF'nin araştırılması bu alt gruplar bazında gerçekleştirilmiştir.

Madde yanlılığı ve DMF'yi belirlemek için iki kategorili puanlama yöntemi ile elde edilen verilerin analizinde MH test istatistiği ile  $\Delta_{MH}$  katsayısından ve lojistik regresyon modellerinden; önsel ağırlıklı puanlama

yöntemi ile elde edilen verilerin analizinde ise Zumbo (1999)'nun önerdiği lojistik regresyon modelinden yararlanılmıştır.

*Bulgular*

Araştırmaya katılan öğrencilerin okul türlerine göre dağılımları Tablo 1'de verilmiştir.

Tablo 1'den görülebileceği gibi, öğrencilerin %45'i özel okullarda okumakta iken, %55'i ise devlet okullarında okumaktadır. Okul türüne göre iki kategorili puanlama ve önsel ağırlıklı puanlama yardımıyla elde edilen test istatistikleri ve model sonuçları Tablo 2'de verilmiştir.

Tablo1.

*Okul Türüne Göre Öğrenci Dağılımı*

Okul Türü	Öğrenci Sayısı	%
Özel Okul	665	45
Devlet Okulu	813	55
Toplam	1478	100

Tablo 2'den de görülebileceği gibi, iki kategorili puanlama sonuçlarına ilişkin MH ve  $\Delta_{MH}$  sonuçları ile lojistik regresyon modelinin kestirim sonuçları incelendiğinde, MH sonuçlarına göre 19. madde dışında kalan maddelerde yanlılık miktarı anlamlı bulunmamıştır.

Tablo 2'den  $\Delta_{MH}$  sonuçları incelendiğinde, 19. maddenin özel okullarda okuyan öğrenciler lehine çalıştığı gözlenmektedir. Özel okullarda okuyan öğrenciler lehine çalışan ancak anlamlı düzeyde yanlılık göstermeyen diğer maddeler sırasıyla; 1, 3, 6, 9, 10, 13, 15, 17 ve 18. maddelerdir. Bu maddeler dışındaki maddeler ise devlet okulunda okuyan öğrenciler lehine çalışmış ancak yine, yanlılık miktarı anlamlı bulunmamıştır.

Lojistik regresyon sonuçları incelendiğinde, iki kategorili puanlama ile önsel ağırlıklı puanlama arasında model uyumu olarak önemli bir fark gözlenmemiştir. Her iki puanlama yönteminin de modele tam uyum gösterdiği görülmektedir.

MH'den elde edilen sonuçlara göre maddelerin hangi alt grubun lehine çalıştığı belirlenebilmekle birlikte, bu durum ancak TB-DMF'nin olduğu durumlar için geçerlidir. Maddelerin farklı yetenek düzeylerinde aynı grup

Tablo 2.  
İki Kategorili ve Önsel Ağırlıklı Puanlamanın Okul Türü Bakımından DMF'ye Etkisine İlişkin Sonuçlar

Madde No	DMF Düzeyi	İki Kategorili Puanlama					Önsel Ağırlıklı Puanlama			
		MH Sonuçları		Lojistik Regresyon Sonuçları			Lojistik Regresyon Sonuçları			Kesişim Etkisi (p)
		MH	$\Delta_{MH}$	Model Uyumu	Grup Etkisi (p)	Kesişim Etkisi (p)	Model Uyumu	Grup Etkisi (p)		
1	A	0,984	0,037	0,002	0,163	0,186	0,000	0,118	0,066	
2	A	1,208	-0,445	0,000	0,164	0,016	0,000	0,099	0,067	
3	A	0,803	0,517	0,000	0,297	0,401	0,000	0,243	0,116	
4	A	1,125	-0,278	0,001	0,636	0,071	0,000	0,586	0,498	
5	A	1,486	0,931	0,000	0,487	0,367	0,000	0,979	0,969	
6	A	0,985	0,259	0,000	0,618	0,676	0,000	0,828	0,649	
7	A	1,081	-0,183	0,000	0,248	0,313	0,000	0,298	0,446	
8	A	1,106	-0,236	0,000	0,197	0,216	0,000	0,432	0,337	
9	A	0,857	0,363	0,000	0,703	0,754	0,000	0,880	0,853	
10	A	0,924	0,186	0,000	0,482	0,464	0,000	0,600	0,508	
11	A	1,325	-0,661	0,000	0,217	0,493	0,000	0,881	0,903	
12	A	1,016	-0,038	0,000	0,535	0,631	0,000	0,282	0,242	
13	A	0,894	0,263	0,000	0,951	0,877	0,000	0,799	0,688	
14	A	1,177	-0,384	0,000	0,981	0,588	0,000	0,765	0,371	
15	A	0,752	0,671	0,000	0,989	0,649	0,000	0,456	0,382	
16	A	1,228	-0,482	0,000	0,900	0,420	0,000	0,521	0,669	
17	A	0,828	0,445	0,000	0,270	0,323	0,000	0,257	0,352	
18	A	0,807	0,503	0,000	0,598	0,345	0,000	0,068	0,066	
19*	B	0,431	1,975	0,000	0,616	0,438	0,000	0,208	0,119	
20	A	1,088	-0,198	0,000	0,639	0,197	0,000	0,578	0,606	

lehine yanlılık gösterip göstermediği lojistik regresyon sonuçlarına bakılarak yorumlanabilir. Tablo 2'de, lojistik regresyon analizinden elde edilen grup etkisi ve kesişim etkileri incelendiğinde, madde yanlılığı olmadığı gözlenmiştir. Bu bulgu, MH sonuçları ile de paralellik göstermektedir. Buna karşın, 19. maddede ortaya çıkan orta düzeydeki (B-level) yanlılık lojistik regresyon sonuçlarında gözlenmemiştir. Bunun bir nedeni, MH test istatistiğinin DMF'nin nicel büyüklüğünü de dikkate alması ile açıklanabilir.

Araştırmaya katılan öğrencilerin cinsiyetlerine göre dağılımları Tablo 3'te verilmiştir. Tablo 3'ten de görülebileceği gibi, araştırmaya katılanların %53'ü erkek, %47'si ise kız öğrencidir.

Cinsiyete göre iki kategorili puanlama ve önsel ağırlıklı puanlama yardımıyla elde edilen test istatistikleri ve model sonuçları Tablo 4'te verilmiştir.

Tablo 3.  
Cinsiyete Göre Öğrenci Dağılımı

Cinsiyet	Öğrenci Sayısı	%
Erkek	780	53
Kız	698	47
Toplam	1478	100

Tablo 4'te iki kategorili puanlama sonuçlarına ilişkin MH ve  $\Delta_{MH}$  sonuçları ile lojistik regresyon modelinin kestirim sonuçları incelendiğinde, okul türü bakımından elde edilen sonuçlarda olduğu gibi, MH sonuçlarına göre 19. madde dışında kalan maddelerde yanlılık miktarı önemli bulunmamıştır. Yine,  $\Delta_{MH}$  sonuçları incelendiğinde, 19. maddenin erkek öğrenciler lehine çalıştığı gözlenmektedir. Erkek öğrenciler lehine çalışsan diğer

Tablo 4.

*İki Kategorili ve Önsel Ağırlıklı Puanlamanın Cinsiyet Bakımından DMF'ye Etkisine İlişkin Sonuçlar*

Madde No	DMF Düzeyi	İki Kategorili Puanlama			Önsel Ağırlıklı Puanlama				
		MH Sonuçları	Δ <sub>MH</sub>	Lojistik Regresyon Sonuçları	Lojistik Regresyon Sonuçları	Lojistik Regresyon Sonuçları	Lojistik Regresyon Sonuçları		
		MH		Model Uyumu	Grup Etkisi (p)	Kesişim Etkisi (p)	Model Uyumu	Grup Etkisi (p)	Kesişim Etkisi (p)
1	A	0,848	0,388	0,006	0,932	0,808	0,000	0,665	0,721
2	A	0,855	0,367	0,000	0,685	0,388	0,000	0,775	0,755
3	A	1,526	-0,994	0,000	0,966	0,397	0,000	0,540	0,629
4	A	1,071	-0,161	0,005	0,539	0,424	0,000	0,264	0,247
5	A	1,021	-0,049	0,000	0,245	0,212	0,000	0,142	0,148
6	A	0,838	0,415	0,000	0,585	0,329	0,000	0,325	0,284
7	A	1,222	-0,471	0,000	0,391	0,169	0,000	0,459	0,479
8	A	0,983	0,041	0,000	0,173	0,137	0,000	0,333	0,290
9	A	0,988	0,029	0,000	0,310	0,344	0,000	0,866	0,882
10	A	0,873	0,318	0,000	0,214	0,350	0,000	0,598	0,607
11	A	0,944	0,135	0,000	0,106	0,062	0,000	0,610	0,609
12	A	1,053	-0,122	0,000	0,594	0,566	0,000	0,864	0,843
13	A	0,954	0,110	0,000	0,382	0,246	0,000	0,275	0,271
14	A	1,192	-0,413	0,000	0,470	0,423	0,000	0,360	0,310
15	A	0,902	0,242	0,000	0,578	0,398	0,000	0,393	0,285
16	A	1,182	-0,393	0,000	0,651	0,370	0,000	0,291	0,236
17	A	1,076	-0,173	0,000	0,444	0,475	0,000	0,684	0,742
18	A	0,935	0,159	0,000	0,168	0,110	0,000	0,383	0,343
19*	B	1,925	-1,539	0,000	0,000	0,001	0,000	0,000	0,001
20	A	0,792	0,549	0,000	0,156	0,182	0,000	0,288	0,218

maddeler sırasıyla 3, 4, 5, 7, 12, 14, 16 ve 17. maddelerdir; ancak bu yanlılık göz ardı edilebilecek düzeyde bulunmuştur (A-level). 1, 2, 6, 8, 9, 10, 11, 13, 15, 18 ve 20. maddeler ise kız öğrenciler lehine çalışmakla birlikte yine yanlılık miktarı anlamlı bulunmamıştır.

Tablo 4'te lojistik regresyon analizine dayalı sonuçlar incelendiğinde, iki kategorili puanlama ve önsel ağırlıklı puanlama yöntemlerinin modele tam uyum gösterdiği görülmektedir. İki kategorili ve önsel ağırlıklı puanlama için yapılan lojistik regresyon analizine dayalı grup etkisi ve kesişim etkileri incelendiğinde ise, 19. madde dışında madde yanlılığı bulunmadığı görülebilir. Lojistik regresyon analizi sonuçları 19. maddedeki yanlılığın farklı yetenek düzeylerinde farklı gruplar lehine olduğu bilgisini verirken, MH test istatistiği bu yanlılığın erkekler lehine ancak göz ardı edilebilecek düzeyde olduğunu göstermektedir.

### Sonuç ve Öneriler

Bu araştırmada, Türkçe okuduğunu anlama yeteneğini ölçme amacıyla geliştirilen bir test, cinsiyet grupları ve okul türüne göre, diferansiyel madde fonksiyonu açısından incelenmiş, iki kategorili ve önsel ağırlıklı puanlama türlerinin DMF'ye etkisi araştırılmıştır.

Araştırma sonucunda, iki kategorili ve önsel ağırlıklı puanlama bakımından teste yer alan maddelerin büyük çoğunluğunun hem okul türüne hem de cinsiyete göre önemli düzeyde yanlılık göstermediği belirlenmiştir. Her iki alt grup için de yalnızca bir maddenin orta düzeyde (B-level) yanlılık gösterdiği ve bu yanlılığın okul türü bakımından özel okullarda okuyan öğrenciler lehine, cinsiyet bakımından da erkek öğrenciler lehine olduğu belirlenmiştir. Bu sonuca göre, testin uygulanmasıyla elde edilen verilerin "yansız" maddeler üzerinden elde edilen veriler olduğu söylenebilir. Bu durum,

geliştirilen testin, dikkate alınan alt gruplardan kaynaklanabilecek sistematik hatalardan arınık olduğunun bir göstergesi olarak kabul edilebilir.

Test geliştirme çalışmalarında genellikle tesadüfi hatalar üzerinde durulmakta ve sistematik hatalar göz ardı edilmektedir. DMF yardımıyla sistematik hatanın varlığı hakkında bilgi elde etmek, geçerliği yüksek testler oluşturmada önemli bir çalışma olacaktır.

Bu araştırmada okul türüne ve cinsiyete göre karşılaştırmalar yapılmıştır. Benzer çalışmalar farklı alt gruplar dikkate alınarak yapılabilir.

Bu araştırmada iki kategorili ve önsel ağırlıklı puanlama türlerinin DMF'ye etkisi belirlenmeye çalışılmıştır. Benzer çalışma, farklı puanlama yöntemleri dikkate alınarak yapılabilir.

#### Kaynakça

- Camilli, G. & Shepard, L.A. (1994). *Methods for identifying biased test items*. London: Sage Publications.
- Fidalgo, A.M. , Mellenberg, G.J. & Muñoz, J. (2000). Effects of amount of DIF, test length, and purification type on robustness and power of Mantel-Haenszel procedures. *Methods of Psychological Research Online*, 5 (3), Retrieved on November 17, 2001 (de indirildi) from the World Wide Web: <http://www.mpr-online.de>

- Gierl, M.J., Jodoin, M.G. & Ackerman, T.A. (24-27 Nisan 2000). Performance of Mantel-Haenszel, simultaneous item bias test, and logistic regression when the proportion of DIF items is large. Annual Meeting of the American Educational Research Association (AERA).
- Hambleton, R.K. & Rogers, H.J. (1989). Detecting potentially biased test items: Comparison of IRT area and Mantel-Haenszel methods. *Applied Measurement in Education*, 2 (4), 313-334.
- Hambleton, R.K., Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of item response theory*. London: Sage Publications.
- Osterlind, J.S. (1983). *Test item bias*. London: Sage Publications.
- Whitmore, M.L. & Schumacker, R.E. (1999). A comparison of logistic regression and analysis of variance differential item functioning detection methods. *Educational and Psychological Measurement*, 59 (6), 910 - 927.
- Yu, C.H. (May 2001). True score model and item response theory. Retrieved on May 25, 2001 (de indirildi) from the World Wide Web: <http://seamonkeyv.ed.asu.edu/~alex>.
- Zumbo, B.D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Geliş	14 Ağustos 2002
İnceleme	12 Eylül 2002
Kabul	3 Şubat 2003