

SBS 2011 TÜRKÇE ALT TESTİNDEKİ MADDELERİN DEĞİŞEN MADDE FONKSİYONU AÇISINDAN FARKLI YÖNTEMLERLE İNCELENMESİ*

Emre TOPRAK

Dr., Erciyes Üniversitesi, etoprak@erciyes.edu.tr

Levent YAKAR

Dr., Hacettepe Üniversitesi, leventyakar@hacettepe.edu.tr

Received: 07.12.2016

Accepted: 21.03.2017

Öz

Bu araştırmada SBS 2011 yılı 8. sınıf Türkçe alt testinde yer alan 23 maddede cinsiyete göre DMF bulunma durumunun farklı DMF belirleme yöntemleriyle incelenmesi amaçlanmıştır. Ayrıca ikinci bir amaç olarak, DMF içeren maddelerin bulunması halinde bunların coğrafi bölgelere göre durumunun ne olduğunun ortaya konulması belirlenmiştir. DMF tespit etmek üzere verilerin analizinde klasik test teorisine dayalı olan Mantel-Haenszel ve Lojistik Regresyon yöntemleri; Madde Tepki Kuramına dayalı olan SIBTEST, Wald istatistiği ve Olabilirlik Oranı yöntemleri kullanılmıştır. Araştırmadan elde edilen bulgulara göre Mantel-Haenszel, SIBTEST, Wald istatistiği ve Olabilirlik Oranı yöntemlerinin ortak maddeler üzerinde DMF varlığını belirlediği, Lojistik Regresyon yönteminin ise farklı bir madde üzerinde DMF varlığını belirlediği görülmüştür. Mantel-Haenszel ve SIBTEST yöntemleriyle sadece 9. ve 21. maddede; Lojistik Regresyon yöntemiyle sadece 2. maddede DMF tespit edilmiştir. Olabilirlik Oranı yöntemi, 9. ve 21. maddeler yanında bu maddelere ek olarak 14. ve 18. maddelerde, Wald İstatistiği ise 9. ve 21. maddeler yanında 14. maddede DMF tespit etmiştir. Lojistik Regresyon yönteminin diğer modellerin tespit ettiği maddelerde sorun görmemesi ve DMF tespit ettiği 2. maddenin diğer yöntemler tarafından sorunsuz görülmesi nedeniyle bu yöntem yorumlarda kullanılmamıştır. DMF tespit edilen maddelerin 2'sinde kızlar lehine, 2'sinde ise erkekler lehine Değişen Madde Fonksiyonu gözlenmiştir.

Anahtar Kelimeler: Değişen madde fonksiyonu, SBS, cinsiyet, coğrafi bölge.

ANALYSIS OF SBS 2011 TURKISH SUBTEST ITEMS IN TERMS OF DIFFERENTIAL ITEM FUNCTIONING BY DIFFERENT METHODS

ABSTRACT

In this study, it was aimed to investigate the presence of DIF according to sex in 23 items in SBS 2011 8th grade Turkish subtest by using different DIF determination methods. In addition, as a second objective, it has been determined that if items are found which containing DIF, their status by geographical region is determined. Mantel-Haenszel and Logistic Regression methods based on the classical test theory and SIBTEST, Wald Statistic and Likelihood Ratio methods based on Item Response Theory are used to detect DIF in the analysis of data. According to the findings obtained from the research, Mantel-Haenszel, SIBTEST, Wald Statistic and Likelihood Ratio methods determined DIF which was on common items and Logistic Regression method

* Bu çalışma 1-3 Eylül 2016 tarihleri arasında gerçekleştirilen 5. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresinde sözlü bildiri olarak sunulmuştur.

determined DIF on a different item. Mantel-Haenszel and SIBTEST methods detect DIF only on 9th and 21st items; Logistic regression method detect DIF only in the 2nd item. The likelihood ratio method determined DIF on 14th and 18th items in addition to 9th and 21st items, Wald Statistic determined DIF on 14th item in addition to 9th and 21st items. Logistic regression method was not used in the interpretation because the Logistic Regression method did not show any problems in items which DIF detected by other methods. DIF was found in two of the items in favor of the girls and two of items in favor of the men.

Keywords: Differential Item functioning, SBS, gender, geographical region.

GİRİŞ

Ölçme ve değerlendirme çalışmalarının temel amacı öğrencilerin ölçülen konuyla ilgili özelliklere sahip olup olmama durumunu geçerli ve güvenilir şekilde belirlemektir. Öğrenciler hakkında verilen kararlar sadece bir test ya da ders başarısı değerlendirmesi olmayıp; aynı zamanda toplumsal bir değer de taşımaktadır. Nitekim bu öğrenciler kendileriyle ilgili yapılan ölçme işlemleri ve değerlendirme çalışmalarına göre toplumsal yaşamdaki konumlarına yerleşmektedir. Bu nedenle öğrenci başarısını belirlemeye yönelik olarak yapılan testler hiçbir şekilde bir kişiye veya gruba avantaj veya dezavantaj yaratmamalıdır. Test maddelerinin, aynı yetenek düzeyinde olan fakat farklı gruplardan gelen bireylerin cevapları açısından farklı özellikler göstermemesi beklenir (Öğretmen ve Doğan, 2004). Adil bir test, tüm birey ve grupların testi doğru olarak yanıtlaması halinde yanıtlayıcıların bilgi ve becerilerini gösterebileceği fırsat eşitliğini sağlamalıdır (Roever, 2005). Test kullanıcılarına göre farklı özelliklere sahip gruplardan elde edilen puanlar birbiriyle karşılaştırılabilir ve bu karşılaştırma sonuçları çeşitli kararlara temel olarak kullanılabilir (Özmen, 2014). Eğer testler içerisindeki maddeler bir gruba cinsiyeti, sosyoekonomik düzeyi, dini vb. çeşitli özelliklerinden dolayı bir avantaj sağlıyorsa sınavın bir gruba karşı yanlı davrandığı söylenebilir (Karakaya, 2012). Sadece testi alan belli bir grubun puanlarında ortaya çıkan farklılık nedeniyle yanlılık, testin sistematik hata içermesi ve geçerliğin düşmesi olarak tanımlanabilir (Camilli ve Shephard, 1994). Madde ve test yanlılığı geçerliğe yönelik en önemli tehditlerden birisidir (Clauser ve Mazor, 1998). Ölçme araçlarının geçerliğinin olumsuz yönde etkilenmemesi için sınavlar ölçülmek istenen özelliği tam ve doğru olarak ölçebilmelidir (Güçlü ve Şanal, 2015).

Günümüzde ölçme araçlarının hemen her alanda kullanılmaya başlanmasından dolayı yanlılık, ölçme süreçlerinin en kritik noktalarından biri haline almıştır. Çünkü bu kavram hem sınava katılan bireyler için bir güvenilirlik sorunu oluşturmakta hem de sınavları uygulayan kişi ya da kurumlar için toplumdaki algıyı ve prestiji temellendirmektedir. Yanlılık kavramının kişi ya da kurumları bu derece etkiliyor olmasının nedenlerinden biri de yanlılığın tesadüfi hataların bir sonucu olmamasıdır. Hatanın tesadüfi olması, ölçme yapılan ortam, ölçme zamanı, ölçme yapan kişinin dikkatsizliği gibi ölçme uygulamalarıyla ilgili olan birçok farklı nedene bağlanabilir. Ancak yanlılıkta belli bir kişiye avantaj veya dezavantaj sağlayan sistematik bir hata söz konusudur. Tesadüfi hataların kaynağı tam olarak tespit edilemeyecekken; sistematik hatalar testler ve test maddeleri analiz edilerek ortaya konulabilir.

Testlerde madde yanlılığını belirlemek için çeşitli istatistiksel yöntemler geliştirilmiştir. Bunlardan en bilineni değişen madde fonksiyonu analizleridir. Değişen madde fonksiyonu terimi farklı gruplarda farklı istatistiksel

özellikler gösteren maddelerin basit bir gözlemini anlatmak için alanyazına girmiştir (Holland ve Wainer, 1993; Akt. Ayan, 2011). Madde ile ölçülmek istenilen psikolojik yapının her bir yetenek düzeyinde maddeyi doğru yanıtlama olasılıklarının alt gruplara göre farklılık gösterdiğini belirleyen bir fonksiyon olan DMF (Zumbo, 1999; Akt. Atalay, 2010), bir maddenin iki farklı grupta psikometrik olarak nasıl farklılaştığını ortaya koyar. Zumbo (2007), DMF'yi bir demografik grupta yer alan bireylerin testteki bir maddeyi, aynı bilgi düzeyine sahip olan ikinci bir demografik grupta yer alan bireylere oranla daha sıklıkla doğru cevapladığı durumları açıklamayı sağlayan istatistiksel bir kavram olarak açıklamıştır (Suna, 2012). Madde yanlılığı belirleme çalışmaları istatistiksel süreç ve eğitim ve psikoloji bulgularına, uzman kanılarına, içerik analizlerine ve olgusal değerlendirmelere dayalı karar verme süreci olmak üzere iki basamaklı bir şekilde yürütülür (Uyar ve Uyanık, 2016).

Genellikle yanlı maddeleri belirlemek için kullanılan DMF, alt gruplar arasındaki gerçek farklılık veya madde yanlılığı nedeniyle ortaya çıkabilmektedir (Camilli ve Shepard, 1994). Bununla birlikte madde yanlılığının bir indeksi olan DMF'nin varlığı, o maddenin yanlı olduğunu belirtmek için yeterli değildir. DMF istatistiksel bir yöntem iken, madde yanlılığı kavramsaldir ve yoruma açıktır (Camilli ve Shepard, 1994; Clauser ve Mazor, 1998; Akt. Atalay, 2010). DMF gösteren maddeler incelenerek, maddenin yanlı olup olmadığına karar verilir. Ayan (2011), DMF çalışmalarındaki bulguları yorumlamanın zorluğu hakkında yaygın olarak farklı DMF belirleme stratejileri kullanıldığında farklı sonuçların gözlemlendiğini ve istatistiksel olarak DMF gösteren maddelerin içeriğine bakıldığında, örneğin basit bir hesaplama probleminin DMF göstermesi için bir sebep bulunmadığını belirtmiş, bir maddede DMF bulmanın o maddenin yanlı, yani gruplardan birine mutlaka adaletsiz davrandığı anlamına gelmediğinin üzerinde durmuştur. Bu durumu Ironson'ın (1982) bir kek pişirmek için gereken sıcaklığı soran örneğinde maddenin istatistiksel olarak erkeklere karşı yanlı olsa bile eğer aşçı seçmek için kullanılan bir testteki madde ise oldukça uygun olabileceğini dile getirdiği örnekle desteklemiştir (Holland ve Wainer,1993; Akt. Ayan, 2011).

Alanyazın incelendiğinde DMF belirlemeye yönelik olarak geliştirilmiş pek çok yöntem olduğu görülmektedir. DMF belirleme yöntemleri bir sınıflamada klasik test kuramına ya da madde tepki kuramına dayalı olmalarına bağlı olarak iki alt grupta; bir diğer sınıflamada gözlenen ya da örtük puan yaklaşımlarına göre yine iki alt grupta sınıflandırılmıştır (Atalay, vd., 2012). Bu yöntemlerin her biri sahip olduğu özelliklere bağlı olarak farklı şekilde işlemektedir. Bu araştırma ile SBS Türkçe alt testinde yer alan maddelerin DMF gösterme durumları farklı yöntemlerle ele alınarak, DMF belirlemede kullanılan bu yöntemlerin hangi maddelerde DMF tespit ettiği karşılaştırılmıştır.

YÖNTEM

Bu araştırma SBS 2011 Türkçe alt testinde cinsiyet ve coğrafi bölge değişkenleri açısından farklı yöntemlerle DMF varlığını belirlemeye yönelik olarak gerçekleştirilen betimsel bir çalışmadır. Araştırmanın amacı, araştırmada kullanılan veriler ve verilerin analizi sırasıyla verilmiştir.

Araştırmanın Amacı

Bu araştırmada SBS Türkçe alt testinde yer alan 23 maddede cinsiyete göre DMF bulunmasının, farklı DMF belirleme yöntemleriyle incelenmesi ve DMF belirleme yöntemlerinin tespit ettiği maddelerin karşılaştırılması amaçlanmıştır. Ayrıca ikinci bir amaç olarak DMF içeren maddelerin bulunması halinde bunların coğrafi bölgelere göre durumunun ne olduğunun ortaya konulması belirlenmiştir.

Araştırma Verileri

Araştırmada kullanılan veriler 2011 yılı 8. sınıf SBS'nin 23 soruluk Türkçe alt testine aittir. Yaklaşık 1 milyon öğrencinin girdiği sınavdan rastgele seçilen 7.000'i erkek ve 7.000'i kız olmak üzere toplam 14.000 öğrencinin verisi analize alınmıştır.

Verilerin Analizi

Değişen madde fonksiyonunu (DMF) tespit etmek üzere verilerin analizinde Mantel-Haenszel, Lojistik Regresyon, SIBTEST, Wald İstatistiği ve Olabilirlik Oranı yöntemlerinden faydalanılmıştır. Mantel-Haenszel ve Lojistik Regresyon yöntemleri için EZDIF programı, SIBTEST yöntemi için aynı adlı olan SIBTEST programı, Wald İstatistiği için IRTPRO programı ve Olabilirlik Oranı yöntemi için IRTLRF programı kullanılmıştır.

Analizlerde erkek öğrenciler odak, kız öğrenciler ise referans olarak alınmıştır. SBS'de şans başarısı düzeltme formülü uygulandığından IRTLRF programında DMF 2 parametreleri lojistik modelde aranmıştır.

BULGULAR

Tablo 1'de öğrencilerin 23 maddeye verdiği yanıtlar sonucu oluşan madde gücü, cinsiyetlere göre madde güçlükleri ve madde test korelasyonları olan klasik test istatistikleri verilmiştir.

Tablo 1: Klasik Test Teorisi İstatistikleri

Maddeler	Madde Güçlüğü	Madde Güçlüğü (Erkek)	Madde Güçlüğü (Kız)	Madde-Test Kor.
1. Madde	0,773	0,775	0,772	0,435
2. Madde	0,593	0,594	0,592	0,513
3. Madde	0,694	0,694	0,694	0,510
4. Madde	0,816	0,813	0,818	0,536
5. Madde	0,545	0,540	0,549	0,472
6. Madde	0,863	0,866	0,861	0,420

7. Madde	0,629	0,628	0,630	0,486
8. Madde	0,511	0,504	0,519	0,466
9. Madde	0,803	0,812	0,795	0,512
10. Madde	0,805	0,806	0,804	0,471
11. Madde	0,198	0,196	0,201	0,161
12. Madde	0,640	0,634	0,647	0,572
13. Madde	0,538	0,537	0,538	0,478
14. Madde	0,652	0,655	0,648	0,529
15. Madde	0,579	0,577	0,581	0,499
16. Madde	0,741	0,745	0,737	0,594
17. Madde	0,213	0,212	0,214	0,202
18. Madde	0,533	0,524	0,543	0,482
19. Madde	0,490	0,486	0,494	0,493
20. Madde	0,576	0,572	0,580	0,455
21. Madde	0,581	0,570	0,592	0,504
22. Madde	0,472	0,470	0,475	0,486
23. Madde	0,693	0,692	0,695	0,528
Test Genel Güçlüğü		0,606		

Tablo 1 incelendiğinde testin genel güçlüğü'nün 0,606 olduğu görülmektedir. Bunun yanında madde güçlüklerinin 11. ve 17. maddeler dışında benzer olduğu ve 0,472 ile 0,863 arasında değiştiği görülmektedir. Öğrencilerin cinsiyetine göre maddelerin güçlük değerlerine bakıldığında da büyük farklılıklar olmadığı görülmektedir. Maddelerin güçlük değerleri ile cinsiyete göre elde edilen madde güçlük değerlerinin birbirine benzer düzeyde olduğu belirlenmiştir. Nokta çift serili madde-test korelasyonları incelediğinde 11. ve 17. Maddeler dışında madde test korelasyonlarının 0,420 ile 0,594 arasında değiştiği ve orta düzeyde olduğu, bunu yanında 11. ve 17. maddelerin madde test korelasyonlarının oldukça düşük olduğu gözlenmiştir. Yine aynı maddelerin güçlük düzeyleri de diğer maddelere göre oldukça düşüktür.

Öğrencilerin 23 maddeye verdiği yanıtlar araştırmada kullanılan Mantel-Haenszel, Lojistik Regresyon, SIBTEST, Wald İstatistiği ve Olabilirlik Oranı yöntemleri ile analiz edildiğinde elde edilen bulgular Tablo 2'de verilmiştir.

Tablo 2: Mantel-Haenszel, Lojistik Regresyon, SIBTEST, Wald istatistiği ve Olabilirlik Oranı Yöntemlerine Göre Değişen Madde Fonksiyonu Analizi Bulguları

Maddeler	M-H D-DIF	M-H (p Değeri)	L-R (p Değeri)	Sibtest (p Değeri)	Olabilirlik Oranı	Wald (p Değeri)
1. Madde	0,05	0,63	0,08	0,45	3,60	0,18
2. Madde	0,09	0,35	0,04*	0,29	3,70	0,18
3. Madde	0,03	0,75	0,74	0,58	0,20	0,95
4. Madde	-0,13	0,29	0,51	0,39	1,80	0,58
5. Madde	-0,05	0,59	0,32	0,61	0,90	0,69
6. Madde	0,10	0,46	0,66	0,36	1,90	0,44
7. Madde	0,01	0,94	0,28	0,85	2,20	0,42
8. Madde	-0,10	0,26	0,28	0,31	2,50	0,34
9. Madde	0,32	0,01*	0,60	0,00*	8,20*	0,04*
10. Madde	0,01	0,96	0,11	0,76	2,30	0,36
11. Madde	-0,04	0,74	0,78	0,79	0,50	0,78
12. Madde	-0,15	0,13	0,54	0,15	3,80	0,27
13. Madde	0,03	0,77	0,67	0,70	0,30	0,88
14. Madde	0,15	0,13	0,26	0,07	8,30*	0,03*
15. Madde	0,01	0,94	0,60	0,86	0,70	0,77
16. Madde	0,15	0,20	0,78	0,12	3,10	0,31
17. Madde	0,03	0,76	0,37	0,58	1,40	0,51

18. Madde	-0,18	0,06	0,93	0,06	3,90*	0,21
19. Madde	-0,02	0,84	0,96	0,93	0,30	0,89
20. Madde	-0,06	0,54	0,66	0,54	0,50	0,82
21. Madde	-0,22	0,02*	0,09	0,03*	7,40*	0,05*
22. Madde	0,01	0,90	0,88	0,82	0,00	0,98
23. Madde	-0,02	0,86	0,86	0,97	0,10	0,98

*Kullanılan yöntemle göre DMF gösteren madde

Tablo 2 incelendiğinde Mantel-Haenszel, SIBTEST, Wald istatistiği ve Olabilirlik Oranı yöntemlerinin ortak maddeler üzerinde DMF olduğunu belirlediği, Lojistik Regresyon yönteminin ise farklı bir madde üzerinde DMF olduğunu belirlediği görülmektedir.

Mantel-Haenszel ve SIBTEST yöntemleri sadece 9. ve 21. maddede DMF tespit ederken; Lojistik Regresyon yöntemiyle sadece 2. maddede DMF tespit edilmiştir. Olabilirlik Oranı yöntemi, Mantel-Haenszel ve SIBTEST yöntemleri ile belirlenen 9. ve 21. Maddeler yanında bu maddelere ek olarak 14. ve 18. maddelerde ve Wald İstatistiği ise 9. ve 21. maddeler yanında 14. maddede DMF tespit etmiştir.

Lojistik Regresyon yönteminin diğer modellerin tespit ettiği maddelerde sorun görmemesi, bunun yanında DMF tespit ettiği 2. maddenin ise diğer yöntemler tarafından sorunsuz görülmesi nedeniyle yorumlarda kullanılmamıştır.

Tablo 3'te cinsiyete göre DMF gösteren maddeler ve bu maddelerin yönü verilmiştir.

Tablo 3: Mantel-Haenszel, SIBTEST, Wald İstatistiği ve Olabilirlik Oranı Yöntemlerine Göre Cinsiyet Açısından DMF Gösteren Maddeler ve DMF Yönü

DMF Gösteren Maddeler	M-H	Sibtest	Olabilirlik Oranı	Wald İstatistiği
9. Madde	Erkekler Lehine	Erkekler Lehine	Erkekler Lehine	Erkekler Lehine
14. Madde	Yok	Yok	Erkekler Lehine	Erkekler Lehine
18. Madde	Yok	Yok	Kızlar Lehine	Yok
21. Madde	Kızlar Lehine	Kızlar Lehine	Kızlar Lehine	Kızlar Lehine

DMF tespit edilen maddelerin 2'sinde kızlar lehine, 2'sinde ise erkekler lehine Değişen Madde Fonksiyonu gözlenmiştir. Yöntemlerin tümü 9. maddede erkekler lehine DMF olduğunu belirlemiştir. 14. maddede DMF olduğunu belirleyen Olabilirlik Oranı ve Wald İstatistiği yöntemleri DMF'nin erkekler lehine olduğunu tespit etmiştir. 18. maddede DMF olduğunu belirleyen tek yöntem Olabilirlik Oranı yöntemi olup, DMF'nin kızlar lehine olduğu sonucuna ulaşmıştır. 21. maddede de 9. maddeye benzer şekilde tüm yöntemler tarafından DMF varlığı belirlenmiş olup, belirlenen DMF kızlar lehinedir.

Test maddelerinde DMF varlığı öğrencilerin yaşadığı coğrafi bölgeler açısından da incelenmiştir. Lojistik regresyon hariç yöntemlerin tümü 9. ve 21. maddede DMF tespit etmiştir. Bu nedenle 9. ve 21. maddelerde coğrafi bölgeler açısından DMF varlığı sadece bu iki maddede DMF tespit eden SIBTEST programı ile analiz edilmiştir. Analiz yapılırken DMF tespit edilen maddeler eşleştirilen maddeler arasına alınmayarak daha iyi sonuçlar elde edilmeye çalışılmıştır.

Tablo 4'te SBS'ye katılan öğrencilerin yaşadığı coğrafi bölgelere göre dağılımı verilmiştir.

Tablo 4: SBS'ye Katılan Öğrencilerin Yaşadığı Coğrafi Bölgelere Göre Dağılımı

Bölge		Erkek	Kız	Toplam
Akdeniz Bölgesi	f	1100	1157	2257
	%	48,74	51,26	16,12
Doğu Anadolu Bölgesi	f	645	568	1213
	%	53,17	46,83	8,66
Ege Bölgesi	f	784	793	1577
	%	49,72	50,28	10,84
İç Anadolu Bölgesi	f	1563	1614	3177
	%	49,20	50,80	22,69
Güneydoğu Anadolu Bölgesi	f	642	665	1307
	%	49,12	50,88	9,34
Karadeniz Bölgesi	f	503	462	965
	%	52,12	47,88	6,89
Marmara Bölgesi	f	1763	1741	3504
	%	50,31	49,69	25,03
Toplam		7000	7000	14000

Tablo 4 incelendiğinde SBS'ye katılan öğrencilerin cinsiyetlerine göre bölgelerdeki dağılımı verilmiştir. Bölgelere göre öğrenci dağılımlarında kız ve erkek öğrencilerin belli bölgelerde çok sık görüldüğü söylenemez. Öğrenciler bölgelerde genel olarak yarı yarıya bir dağılım göstermiştir. Ancak öğrenciler beklendiği üzere bölgeler arasında dengeli dağılmamışlardır. Bölgelerin nüfuslarıyla kıyaslandığında bu sonucun normal bir durum olduğu söylenebilir. Öğrenciler %25,03 ile nüfusun en yoğun olduğu bölge olan Marmara Bölgesinden olurken, en az öğrenci %6,89 ile Karadeniz Bölgesinden olmuştur.

Bölgelerin tümüne göre DMF analizi yapılırken sonuçların bölgedeki öğrenci sayısından etkilenebileceği düşünülmüştür. Bu nedenle bölgelere göre DMF analizi yapılması, DMF tespit edilen maddelerin tüm bölgelerde DMF gösterip göstermemesi açısından önemlidir.

Tablo 5'te 9. ve 21. Maddelerde bölgelere göre SIBTEST yöntemiyle yapılan DMF analizi bulguları verilmiştir.

Tablo 5: 9. ve 21. Maddelerde Bölgelere Göre SIBTEST Yöntemiyle Yapılan DMF Analizi Bulguları

Bölge	9. Madde (p Değeri)	21. Madde (p Değeri)
Akdeniz Bölgesi	0,27	0,14
Doğu Anadolu Bölgesi	0,51	0,20
Ege Bölgesi	0,14	0,76
İç Anadolu Bölgesi	0,13	0,10
Güneydoğu Anadolu Bölgesi	0,36	0,30
Karadeniz Bölgesi	0,26	0,89
Marmara Bölgesi	0,32	0,27

Tablo 5 incelendiğinde SIBTEST yöntemiyle yapılan DMF analizinde bölgelerin hiçbirinde elde edilen değerlerin anlamlı olmadığı ve bölgelere göre DMF bulunmadığı görülmektedir. 9. ve 21. maddelerde tespit edilen

DMF'nin kaynağı bölgelere göre incelendiğinde genelde görülen DMF'nin hiçbir bölge bazında görülmemesi beklenmeyen bir sonuç olarak ortaya çıkmıştır. Bölgelerde görülmeyen DMF'nin genelde bölgelerden elde edilen sonuçların etkileşimi üzerine görüldüğü söylenebilir. 9. maddenin tüm bölgeler için anlamlı olmasa da erkekler lehine, 21. maddenin ise Doğu Anadolu Bölgesi hariç tüm bölgelerde yine anlamlı olmasa da kızlar lehine avantajlı çıkması, sonuçlar genel olarak değerlendirildiğinde anlamlı bir şekilde bir grup lehine anlamlı çıkmasına yol açabilir.

TARTIŞMA ve SONUÇ

Değişen madde fonksiyonu birçok araştırma ve lisansüstü eğitim tez çalışmalarında araştırma konusu olarak çalışılmıştır. Araştırmalarda görüldüğü üzere DMF tespit etmede birçok yöntem kullanılmaktadır (Atalay, 2010; Ayan, 2011; Çepni, 2012; Suna, 2012). Yapılan analizlerde bir yöntemin çok sayıda DMF gösteren madde tespit etmesi o yöntemin daha doğru çalıştığı konusunda bilgi vermez. Aynı şekilde bir yöntemin az sayıda DMF gösteren madde tespit etmesi ilgili yöntemin sonuçlarının yetersizliği anlamına da gelmemelidir. Bu nedenle araştırmada kullanılan yöntemlerin daha iyi ya da kötü olması üzerinde durmaktan çok ortak olarak tespit ettikleri DMF gösteren maddeler üzerinde durulmuştur.

Lojistik regresyon yönteminin Mantel-Haenszel, SIBTEST, Wald istatistiği ve Olabilirlik Oranı yöntemlerinin ortak bir şekilde belirlediği DMF gösteren maddeleri belirleyememesi ve DMF tespit ettiği tek maddenin de yine diğer yöntemlerin hiçbiri tarafından DMF'li olarak tespit edilmemesi yöntemin diğerlerinden daha farklı bir işleyişe sahip olduğunu göstermektedir. Atalay (2010), PISA 2006 öğrenci anketinde yer alan tutum maddelerini değişen madde fonksiyonu açısından incelediği araştırmasında, OLR ve SIBTEST yöntemlerinin farklı maddeler üzerinde DMF bulunduğu sonucuna ulaşmıştır. Suna (2012), TIMSS 2007 fen bilimleri testindeki maddeleri dil ve cinsiyet yanlılığı açısından incelediği araştırmasında, Mantel-Haenszel, SIBTEST ve Lojistik Regresyon yöntemlerini kullanmış ve Mantel-Haenszel ve Lojistik Regresyon yöntemlerinin birbiriyle uyumlu sonuçlar sağladığını, buna karşın SIBTEST yönteminin aynı düzeyde tutarlılık göstermediğini belirlemiştir. Ayan (2011), PISA 2009 fen okuryazarlığı alt testini değişen madde fonksiyonu açısından incelediği araştırmasında, Lojistik Regresyon ve Mantel-Haenszel yöntemlerini kullanmış ve yöntemlerin DMF belirlediği maddelerin birbirinden tamamen farklı sonuçlar ürettiği sonucuna ulaşmıştır. Atalay ve diğerleri (2012), değişen madde fonksiyonunun belirlenmesinde kullanılan farklı yöntemleri bir simülasyon çalışması üzerinde karşılaştırmış ve Lojistik Regresyon yönteminin DMF belirleme performansının düşük, Olabilirlik Oranı yönteminin performansının ise daha yüksek olduğunu tespit etmiştir. Çepni (2011), değişen madde fonksiyonlarını SIBTEST, Mantel-Haenszel, Lojistik Regresyon ve Madde Tepki Kuramına dayalı yöntemlerle incelemiş ve yöntemlerin benzer ve ayrı maddelerde DMF belirlediğini görmüştür. Bekci (2007), Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı'nda değişen madde fonksiyonlarını cinsiyet ve okul türüne göre incelemiş ve Mantel-Haenszel ve Lojistik Regresyon yöntemlerinin DMF belirlemede uyumlu sonuçlar elde etmediğini belirlemiştir. Doğan ve Öğretmen (2008), değişen madde fonksiyonunu belirlemede Mantel-Haenszel, Ki-Kare ve Lojistik Regresyon tekniklerini karşılaştırmış ve yöntemlerin ki-kare değerlerinin büyüklüğü bakımından benzer, DMF verdiği belirlenen madde

sayısı bakımından farklı sonuçlar ürettiğini belirlemiştir. Arıkan ve diğerleri (2016), MIMIC, SIBTEST, Lojistik Regresyon ve Mantel-Haenszel yöntemleriyle bir DMF ve yanlılık çalışması gerçekleştirmiş ve yöntemlerin her birinin hem ortak hem de farklı maddeler üzerinde DMF bulunduğunu belirlemiştir. Araştırmacılar DMF belirleme çalışmalarında farklı DMF belirleme yöntemlerinin bir arada kullanılmasının daha doğru sonuçlar elde edilmesi açısından gerekli olduğunu vurgulamıştır. Araştırma sonuçları alanyazında yer alan çalışmaların bir kısmıyla benzerlik taşıyarak bir kısmından farklılık göstermektedir. Özellikle Lojistik Regresyon yöntemi alanyazında diğer yöntemlere göre düşük performans göstermiştir. Bu çalışmada da Lojistik Regresyon yöntemi hem diğer yöntemlerin belirlediği DMF'li maddeleri belirleyememiş hem de yöntemlerin hiçbirinin DMF görmediği bir maddede DMF olduğu sonucuna ulaşmıştır.

Yöntemlerin tümünün SBS 8. sınıf Türkçe testinde DMF gösteren maddeler bulması testin bu konuda incelenmesi gerektiği sonucunu doğurmaktadır. DMF görülen maddelerin konu uzmanlarınca da değerlendirilmesi DMF gösteren maddelerle karşılaşılması açısından faydalı olabilir. Araştırmalarda DMF gösteren maddeler belirlenmesi testin geçerliğini doğrudan etkileyen önemli bir problemdir (Gök, vd., 2014). Bu nedenle test geliştirme çalışmalarında maddelerin titizlikle hazırlanması ve hiçbir birey ya da gruba yanlılık göstermemesine özen gösterilmelidir.

Genel olarak DMF tespit edilen 2 maddenin bölgesel olarak bakıldığında hiçbir bölgede DMF göstermemesi beklenmeyen bir sonuç olarak görülebilir. Bu çalışmada en azından veride en büyük grupları oluşturan bölgelerden biri ya da birkaçında DMF gözlenmesi beklenmiştir. Ancak bölgeler açısından anlamlı olmasa da DMF'nin yönünün aynı olması genel sonuçta DMF görülmesini açıklayabilir. Benzer şekilde bölgelerde görülebilecek farklı yönlerdeki DMF'ler genel itibarıyla birbirini etkisizleştirerek genel veride DMF olmadığı sonucunu verebilir.

KAYNAKÇA

- Arıkan, Ç. A. (2016). "MIMIC, SIBTEST, Lojistik Regresyon ve Mantel-Haenszel Yöntemleriyle Gerçekleştirilen DMF ve Yanlılık Çalışması" *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(1), 34-52.
- Atalay, K. (2010). *PISA 2006 Öğrenci Anketinde Yer Alan Tutum Maddelerinin Değişen Madde Fonksiyonu Açısından İncelenmesi*. Yüksek Lisans Tezi. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Atalay, K., Gök, B., Kelecioğlu, H. ve Arsan, N. (2012). "Değişen Madde Fonksiyonunun Belirlenmesinde Kullanılan Farklı Yöntemlerin Karşılaştırılması: Bir Simülasyon Çalışması." *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 43, 270-281.
- Ayan, C. (2011). *PISA 2009 Fen Okuryazarlığı Testinin Değişen Madde Fonksiyonu Açısından İncelenmesi*. Yüksek Lisans tezi. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Bekci, B. (2007). *Ortaöğretim Kurumları Öğrenci Seçme Sınavının Değişen Madde Fonksiyonlarının Cinsiyete ve Okul Türüne Göre İncelenmesi*, Yüksek Lisans Tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Camilli, G. & Shepard, L. A. (1994). "Methods For Identifying Biased Test Items." London: Sage publications.

- Clauser, B. E. & Mazor, K. M. (1998). "Using Statistical Procedures To Identify Differentially Functioning Test Items." *Educational Measurement: Issues and Practice*, 17(1), 31-47.
- Çepni, Z. (2011). *Değişen Madde Fonksiyonlarının SIBTEST, Mantel Haenzsel, Lojistik Regresyon ve Madde Tepki Kuramı Yöntemleriyle İncelenmesi*. Doktora Tezi. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Doğan, N. ve Öğretmen, T. (2008). "Değişen Madde Fonksiyonunu Belirlemede Mantel-Haenzsel, Ki-Kare Ve Lojistik Regresyon Tekniklerinin Karşılaştırılması" *Eğitim ve Bilim*, 33(148), 100-112.
- Gök, B., Kabasakal, A. K. ve Kelecioğlu, H. (2014). "PISA 2009 Öğrenci Anketi Tutum Maddelerinin Kültüre Göre Değişen Madde Fonksiyonu Açısından İncelenmesi" *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(1), 72-87.
- Güçlü, M. ve Şanal, M. (2015). "1950 ve 1960 Yılları Arasında Türkiye'de Süreli Yayınlarla Yansıyan Eğitim Sorunları: Eğitim Dergisi Örneği" *Turkish Studies*, 10(15), 451-470.
- Karakaya, İ. (2012). "Seviye Belirleme Sınavındaki Fen Ve Teknoloji İle Matematik Alt Testlerinin Madde Yanlılığı Açısından İncelenmesi" *Kuram ve Uygulamada Eğitim Bilimleri*, 12(1), Kış, 215-229.
- Öğretmen, T. ve Doğan, N. (2004). "OKÖSYS Matematik Alt Testine Ait Maddelerin Yanlılık Analizi", XIII. Ulusal Eğitim Bilimleri Kurultayı, 6-9 Temmuz, İnönü Üniversitesi Eğitim Fakültesi, Malatya.
- Özmen, D. T. (2014). "PISA 2009 Okuma Testi Maddelerinin Yanlılığı Üzerine Bir Çalışma" *Eğitim Bilimleri ve Uygulama*, 13(26), 147-165.
- Roever, C. (2005). "That's not fair!". Fairness. Bias And Differential Item Functioning In Language Testing. SLS Brownbag.
- Suna, H. E. (2012). *TIMMS 2007 Fen Bilimleri Testindeki Maddelerin Dil Ve Cinsiyet Yanlılığı Açısından İncelenmesi*. Yüksek Lisans Tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Uyar, Ş. ve Uyanık, G. K. (2016). "PISA 2012 Bilişsel Maddelerinin Kültüre Göre Değişen Madde Fonksiyonu Bakımından İncelenmesi" *Eğitim ve Öğretim Araştırmaları Dergisi*, 5(3), 230-240.

EXTENDED SUMMARY

Introduction

When the literature is examined, it is seen that there are many methods developed for determining DIF. In this study, the DIF status of the substances in the SBS Turkish subtest were examined by different methods and the DIF was determined in which substances these DIF determination methods were used.

Method

In this study, it was aimed to investigate the presence of DMF according to sex in 23 items of SBS Turkish subtest, to compare them with different DIF determination methods and compare the items determined by DIF determination methods. In addition, if a substance containing DMF is found as a secondary objective, it is determined to determine what the situation is by geographical region. Approximately 1 million students

entered the examination randomly selected from 7,000 men and 7,000 girls, including a total of 14,000 students were analyzed. Mantel-Haenszel, Logistic Regression, SIBTEST, Wald Statistic and Likelihood Ratio methods were used to analyze the data to detect the DIF.

Findings (Results)

Mantel-Haenszel, SIBTEST, Wald Statistic and Likelihood Ratio methods determined that DIF was on common substances and Logistic Regression method was DIF on a different substance. Mantel-Haenszel and SIBTEST methods only detect DIF at 9th and 21st place; Logistic regression method was used to detect DIF only in the 2nd item.

The likelihood ratio method has found DIF at 14th place in addition to these items as well as 9th and 21st items determined by Mantel-Haenszel and SIBTEST methods, in addition to 14th and 18th items and Wald Statistic 9th and 21st items. The Logistic Regression method did not show any problems in the other models, and the second material, which DIF detected, was not used in the interpretation because it was seen by other methods without problems. DIF was found in 2 of the items in favor of the girls, 2 in the case of men in the DIF was observed.

All of the methods identified DIF in favor of men at 9th place. The Likelihood Ratio and Wald Statistic methods that determine DIF at item 14 determine that DIF is in favor of men. The only method that determines DIF at item 18 is the Likelihood Ratio method, which is the result that DIF is in favor of the girls. Similar to item 9 in item 21, DIF has been identified by all methods and the selected DIF girls are favored.

The presence of DIF in the test items was also examined in terms of the geographical regions where the students lived. All methods except the logistic regression detected DIF at 9th and 21st place. For this reason, DIF presence in terms of geographical regions in 9th and 21st items was analyzed only by the SIBTEST program which detects DIF in these two items. When analyzing, it was tried to get better results by not putting the DIF substances into matched materials. When the DIF detected in the 9th and 21st substances is examined according to the source regions, it is unexpected that DIF, which is generally seen, is not seen on the basis of any region. It can be said that the DIF, which is not seen in the regions, is generally seen on the interaction of the results obtained from the regions. The fact that the 9th article is not meaningful for all the regions but the 21st article is advantageous in favor of the girls even if it is not significant in all the regions except the Eastern Anatolian Region can lead to significant meaningful favoring of a group when the results are generally evaluated.

Conclusion and Discussion

Detection of a substance that shows a large number of DIFs in a given analysis does not tell us that it is working correctly. Identifying a substance with a small number of DIFs in the same way should not also imply that the results of the method are inadequate. For this reason, we have focused on the DIF substances that we have

identified more commonly than the methods used in the research to be better or worse. The fact that the logistic regression method cannot identify the DIF data that the Mantel-Haenszel, SIBTEST, Wald statistic and Likelihood Ratio methods jointly determine and DIF finds that the only method has a different function than the other method of not detecting DIF by any other method.

Çepni (2011) has examined DIF with methods based on SIBTEST, Mantel-Haenszel, Logistic Regression and Item Response Theory and found that the methods determine DIF in similar and separate substances. Doğan ve Öğretmen (2008) compared the Mantel-Haenszel, Chi-Square and Logistic Regression techniques for determining the DIF and found that the chi-square values produced different results in terms of the number of substances that gave similar DIF in terms of size. Arıkan et al. (2016) conducted a DIF and bias study using MIMIC, SIBTEST, Logistic Regression and Mantel-Haenszel methods and determined that each method had DIF on both common and different substances. Researchers emphasized that the use of different DIF detection methods in DIF determination studies is necessary to achieve more accurate results. The results of the research differ from some of the studies in the literature and some of the similarities. In particular, the Logistic Regression method showed lower performance than the other methods. In this study, the Logistic Regression method was also able to determine the DIF items determined by the other methods and DIF in a material where none of the methods observed DIF.

Finding DMF substances in SBS 8th grade Turkish test of all methods results in the problem that this test is troublesome. Evaluation of DMF substances by subject experts may also be helpful in avoiding DMF substances. Identification of DMF substances in research is an important problem directly affecting the validity of the test (Gök, et al., 2014). For this reason, care must be taken to ensure that the materials are rigorously prepared in test development studies and that no individual or group is biased.