

Deđişen Madde Fonksiyonunu Belirlemede Mantel–Haenszel ve Lojistik Regresyon Tekniklerinin Karşılaştırılması

The Comparison of Mantel-Haenszel and Logistic Regression Techniques in Determining the Differential Item Functioning*

Bilge GÖK**, Hülya KELECİOđLU***, Nuri DOđAN****
Hacettepe Üniversitesi

Öz

Araştırmanın amacı, deđişen madde fonksiyonunu belirlemede kullanılan Mantel-Haenszel (MH) ve lojistik regresyon (LR) tekniklerini karşılaştırmaktır. Bu amaçla araştırmada, 2005 yılında yapılan Ortaöğretim Kurumları Sınavı (OKS)'nın matematik ve fen bilgisi alt testleri kullanılmıştır. Test maddelerinin cinsiyet ve okul türüne göre fonksiyonlarının farklılaşıp farklılaşmadığı incelenmiştir. Elde edilen bulgular doğrultusunda, MH ve LR teknikleri arasında matematik ve fen bilgisi alt testlerinde cinsiyet ve okul türü deđişkenlerine göre genel olarak düşük düzeyde bir uyum görülmüştür.

Anahtar Sözcükler: Deđişen madde fonksiyonu, madde yanlılığı, lojistik regresyon, Mantel-Haenszel.

Abstract

The purpose of the study is to compare Mantel-Haenszel (MH) and logistic regression (LR) techniques which are used to determine the differential item functioning (DIF). In accordance with this purpose, Maths and Science sub tests of Secondary Schools Exam (SSE) were used. It was examined whether the functions of test items varied according to gender and school type. In the light of the findings, it was found that there was a low level of accordance between MH and LR methods in Science and maths sub test according to gender and school type.

Keywords: Differential item functioning, item bias, logistic regression, mantel-haenszel.

Summary

Purpose

The purpose of this study is to compare Mantel-Haenszel (MH) and logistic regression (LR) techniques which are used to determine the differential item functioning (DIF). In accordance with this purpose, the data related to the quantitative part of the Secondary Schools Exam (SSE) applied in 2005 was used. The purpose of determining the features of items in the quantitative part of SSE was in the foreground. This is a descriptive study based on these characteristics.

* Bu çalışma Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Eđitimde Ölçme ve Deđerlendirme Ana bilim Dalı için hazırlanan yüksek lisans tezinin bir kısmından özetlenmiştir.

** Arş. Gör. Bilge GÖK, Hacettepe Üniversitesi, Eđitim Fakültesi, e-posta: bilgeb@hacettepe.edu.tr

*** Doç. Dr. Hülya KALECİOđLU, Hacettepe Üniversitesi, Eđitim Fakültesi, e-posta: hulyaebb@hacettepe.edu.tr

**** Dr. Nuri DOđAN, Hacettepe Üniversitesi, Eđitim Fakültesi, e-posta: nurid@hacettepe.edu.tr

Results

When the findings obtained at in Maths sub test according to the gender were examined, it was found that 9 items out of 25 questions in Maths exam included DIF at the level of "A" according to the MH results. Negligible DIF was found in 3 items in regard to the LR results. According to this, items in 2005 SSE Maths sub test did not include item bias according to the gender. When the findings obtained from two methods were compared, it was seen that DIF was at the level of "A". From this perspective, MH and LR methods were in accordance with each other. However, there was a difference between these two techniques in terms of the number of items including DIF. When the findings obtained according to the school type in Maths sub test were examined, it was seen that 19 items out of 25 questions included DIF - 15 items were at the level of "A", 3 items were at the level of "B" and 1 item was at the level of "C". 1st, 2nd and 7th items included DIF was in favour of the students at private schools, 17th item at public schools. DIF was found in 19 items by means of MH method and in 17 items by means of LR method. It was seen that DIF was available in 16 items out of 20 items according to both MH and LR. According to MH method, DIF was found in the 1st item at the level of "C" and in the 2nd, 7th and 17th items at the level of "B". It was observed that DIF that could be ignored was available in these items according to LR method. From this perspective, when compared the analysis results of MH and the results of LR, it was seen that there was accordance between MH and LR methods at a low level. When the findings obtained at the end of the analysis carried out in Science sub test according to the gender were examined, it was found that 10 items according to MH method and 9 items according to LR method included DIF at the level of "A". In 2 items out of 17 items, DIF was seen according to both MH method and LR method. When the analysis results of MH and the results of LR were compared, it was found that DIF was obtained at the level of "A". Besides the fact that DIF was found to be at the level of A, these two methods were in accordance with each other at a low level. When the findings obtained at the end of analysis according to the school type in Science sub test were examined, it was seen that 15 items out of 25 questions included DIF according to MH results - 14 items were at the level of "A", and 1 item was at the level of "B". However, according to LR results, 7 items included DIF and 18 items did not include DIF. All items in LR analysis included DIF at the level of "A". DIF was found at different levels in 15 items with MH method and in 18 items with LR method. It was found that 11 items out of 21 items included DIF according to both MH and LR methods. When MH analysis results and LR results were compared, it was seen that DIF was at the level of "A" besides the 5th item; however, while 5th item included DIF at the level of "B" according to MH method, it was observed that this item did not include DIF according to LR method. From this perspective, MH method and LR method were not in accordance with each other at a high level according to the school type.

Conclusion

At the end of the study, it was found that although all items in Maths and Science test which included DIF according to the gender and school type through LR method were at a level which can be ignored, in MH method most of the items were found to be at a level which can be ignored. In both methods, the similarity between the items including DIF was at a very low level. In this context, it can be stated that there was only a similarity between MH and LR methods in terms of the fact that they found DIF at a low level.

Giriş

Bir testle ölçülen özellik bakımından aynı yetenek düzeyinde olan fakat cinsiyet, sosyo-ekonomik düzey bakımından ayrı gruplarda yer alan bireylerin test maddelerine verdikleri cevapların benzer olması beklenir. Yetenek düzeyleri aynı ancak farklı alt gruplarda yer alan bireylerin bir maddeye verdikleri cevapların farklılaşması, maddenin bu gruplar için farklı fonksiyon göstermesi sonucunu doğurmaktadır. Bu durumda madde, aynı yetenek düzeyindeki

gruplardan birine avantaj sağlarken diğer grup için dezavantajlı olabilmektedir. Madde yanlılığı olarak adlandırılan bu durum, testin geçerliğini büyük ölçüde zedelemektedir. Test maddelerinin yansızlığının önemi, öğrenciler hakkında önemli kararların alındığı seçme ve yerleştirme amacıyla kullanılan testlerde daha da artmaktadır. Madde yanlılığı, test koşullarından ya da maddenin bazı özelliklerinden dolayı, bir maddeyi bir grubun doğru yanıt olma olasılığının diğer gruba göre daha az ya da çok olmasıdır (Zumbo, 1999). Bu durumda maddeyi doğru cevaplama olasılığı ölçülen yetenek düzeyinden çok, herhangi bir gruba ait olmaya bağlı olmaktadır. Bu durum ölçme aracının hatalı olduğunun bir göstergesi olarak kabul edilmektedir (Osterlind, 1983).

Her bir test maddesindeki yanlılık, testten yapılan çıkarımların anlamını bozan sistematik hata olarak düşünülebilir. Bu nedenle madde yanlılığı belirleme işlemlerinin amacı, madde geçerliğindeki farklılığı ortaya çıkarmaktır (Camilli ve Shepard, 1994). Madde yanlılığının belirlenmesi, test geçerliğini ve güvenilirliğini artırmak için yapılan önemli çalışmalardan biridir. Madde yanlılığı çalışmaları madde fonksiyonunun aynı yetenek düzeyinde ve farklı gruplarda yer alan bireyler için farklılaşıp farklılaşmadığını belirleyen istatistiksel bir süreçle başlar. Aynı yetenek düzeyindeki alt gruplarda fonksiyonu farklılaşan maddeler, değişen madde fonksiyonu (DMF-differential item functioning) gösteren maddeler olarak adlandırılır. Madde fonksiyonlarındaki farklılıkları gruplar arasındaki farklılıklardan ayırt etmek önemlidir. DMF'nin ortaya çıkmasının iki temel nedeni olabilir. Bunlar alt gruplar arasındaki gerçek farklılığı gösteren madde etkisi ve madde yanlılığıdır (Camilli ve Shepard, 1994).

Madde etkisi, farklı gruplardaki cevaplayıcıların bir maddeyi doğru yanıt olma olasılıklarındaki gerçek farklılıkları ifade eder. Bu farklılıklar, bir maddenin ölçmeye çalıştığı belli bir yetenek düzeyinde var olan farklılıklardır. Bu farklılık gruplardan birisinin sahip olduğu bilgi ya da tecrübe ile açıklanmaktadır. Madde etkisi, ölçme aracının uygulandığı alt gruplar arasında ölçmeye konu olan özellik bakımından farklılıkların bulunduğunu ifade etmektedir.

Değişen madde fonksiyonu, grupların yetenek düzeylerindeki farklılıklardan değil de ölçme aracından kaynaklanıyorsa, bu durum yanlılık olarak adlandırılır. Yanlı maddeler ölçme aracının geçerliğini düşürür. DMF gösteren maddelerin yanlı olup olmadıklarını belirlemek için farklılığın kaynağı araştırılır. DMF gösteren bir maddenin, testin yapı ya da kapsam geçerliği bağlamında incelenmesinden sonra yanlı olup olmadığına karar verilir. Bu yaklaşım, yetenek düzeyleri aynı olan farklı gruplardaki üyelerin bir maddede gösterdikleri performans arasında farklılık olduğunda DMF'nin ortaya çıkacağı varsayımına dayanmaktadır (Dorans ve Holland, 1993). DMF, madde yanlılığının bir indeksi olarak tanımlanabilir.

Değişen madde fonksiyonu, tek biçimli (uniform) ve tek biçimli olmayan (non-uniform) fonksiyonlar olarak ortaya çıkabilir (Mellenbergh, 1982). Aynı yetenek düzeylerindeki iki grubun ilgili maddeyi doğru cevaplama olasılığındaki farklılık sabit ise tek biçimli DMF meydana gelir. Gruplar arasında maddeyi doğru yanıt olma olasılığındaki farklılığın, tüm yetenek düzeylerinde tutarlı olmadığına ise tek biçimli olmayan değişen madde fonksiyonu vardır (Camilli ve Shepard, 1994).

Genel olarak madde yanlılığı belirleme teknikleri, klasik test kuramına (KTK) dayalı teknikler ve madde tepki kuramına (MTK) dayalı teknikler olarak sınıflandırılabilir. KTK'ya dayalı teknikler, varyans analizi, lojistik regresyon (LR), Mantel-Haenszel (MH), dönüştürülmüş madde gücü gibi teknikleri kapsamaktadır. MTK modelleriyle elde edilen madde karakteristik eğrisi ve madde parametreleri gruptan gruba değişmediğinden, gruplara ilişkin parametre değerlerinin ya da madde karakteristik eğrilerinin karşılaştırılması, değişen madde fonksiyonu olup olmadığı konusunda bilgi vermektedir (Öğretmen ve Doğan, 2004). Bu çalışmada, klasik test kuramına dayalı tekniklerden Mantel-Haenszel (MH) ve lojistik regresyon (LR) teknikleri kullanılmıştır.

Araştırmanın Amacı ve Önemi

Araştırmanın amacı, değişen madde fonksiyonunu belirlemede kullanılan parametrik bir teknik olan Mantel-Haenszel ile non-parametrik bir teknik olan lojistik regresyon tekniklerini karşılaştırmaktır. Bu amaçla 2005 yılında uygulanan OKS sayısal alt testine ilişkin veriler kullanılmıştır. DMF'yi belirlemede kullanılan iki teknikten elde edilen sonuçların benzerlik ve farklılıklarının ortaya çıkarılması ve teknikler arasındaki uyumun belirlenmesi amaçlanmıştır. DMF gösteren maddeler için birden fazla kanıt toplanması, maddeler hakkında karar vermeyi kolaylaştırabilir. OKS 2008 yılında son kez ilköğretim 8. sınıflara uygulanmış ve 2009 yılından itibaren uygulamadan kaldırılmıştır (<http://oges.meb.gov.tr/>). Bu araştırma sonuçlarının benzer sınavlara ilişkin değişen madde fonksiyonunu belirleme tekniklerinin nasıl işlediğine ilişkin bilgi verebileceği düşünülmektedir. Araştırma sonuçlarının, gelecekte daha geçerli testler oluşturmak için test hazırlayıcılara bilgi sağlaması umulmaktadır.

Yöntem

Araştırmanın Türü

Bu çalışma, değişen madde fonksiyonunu belirlemede kullanılan Mantel-Haenszel ve lojistik regresyon tekniklerinin OKS sayısal alt testleri üzerinde karşılaştırılmasına yöneliktir. Araştırma, kullanılan teknikler hakkında avantaj, dezavantaj, benzerlik ve farklılık vb. yönünden bilgi vereceğinden kuramsal bir çalışmadır.

Evren ve Örneklem

Araştırmanın evrenini OKS'ye giren 777.137 öğrenci oluşturmaktadır. Evrenden tesadüfi yolla örneklem seçilmiştir. Araştırmanın evrenini oluşturan 777.137 öğrencinin %2'si SPSS paket programının tesadüfi örnekleme modülü kullanılarak alınmıştır. Ardından puanı olmayan veya aykırı değer analizi sonucu aykırı değer sayılan veya okul türü ve cinsiyet bilgisi olmayan bireyler analizden çıkarılmıştır. Örneklem seçimi esnasında örneklemdaki maddelerin betimsel istatistikleri hesaplanmış ve örneklemin evreni temsil edecek nitelikte olmasına dikkat edilmiştir. Çalışmada, kategorik veriler için Cochran (1977)'ın örneklem büyüklüğü formülü kullanılarak örneklem büyüklüğünün en az 384 olması gerektiği bulunmuştur. Araştırmada karşılaştırılan gruplar için kullanılan örneklem büyüklüğü 6767 ile 8053 arasında değişmektedir. Ayrıca araştırmada seçilen örneklem büyüklüğünün alt testlere ve gruplara göre örneklemlere ait aritmetik ortalamaların güven aralıkları hesaplanarak örneklemin hatalı seçilip seçilmediği belirlenmiştir ve bu hesaplamalar sonucunda seçilen örneklem büyüklüklerinin uygun olduğu sonucuna varılmıştır. Zieky (1993)'e göre ise, DMF istatistiklerinin kullanımı için örneklem büyüklüğü, karşılaştırılan küçük gruplar için (odak ya da referans gruplar için) en az 200, küçük ve büyük gruplar birlikte toplam en az 600 olmalıdır.

Veri Toplama Araçları

Bu araştırmada, 2005 yılı OKS sayısal alt testine ait veriler kullanılmıştır. Bu sınavın sayısal bölümünü oluşturan matematik ve fen bilgisi alt testinin her birinde 25 madde olmak üzere toplam 50 madde analize dahil edilmiştir.

Verilerin Analizi

Araştırmada değişen madde fonksiyonu analizleri, klasik test kuramı çerçevesinde Mantel-Haenszel ve lojistik regresyon teknikleri kullanılarak yapılmıştır. MH ve LR tekniklerinin araştırma kapsamında kullanılmasının nedeni, bu tekniklerin benzer bileşenlere sahip olmasıdır. Mantel-Haenszel Tekniği: MH tekniği ki-kare üzerine kurulu bir test istatistiğidir (Agresti, 1984). α (odds oranı) odak ve referans gruptaki performans farklılığının derecesini ölçer. MH'nin kolay yorumlanabilmesi için logaritmik bir dönüştürme ile delta katsayısı elde edilir. MH tekniği

ile DMF belirlenirken; Δ -MH=0 ya da $\alpha =1$ ise madde de DMF olmadığı, Δ -MH<0 ya da $\alpha >1$ ise madde- de referans grubun lehine DMF olduğu, Δ -MH> 0 ya da $\alpha <1$ ise maddede odak grubun lehine DMF olduğu şeklinde yorumlanır (Nandakumar, 1993). Bununla birlikte MH analizi sonucunda ki-kare değerinin anlamlı çıkması da maddenin DMF içerdiğini ifade eder. DMF düzeyini belirlerken; $|\Delta$ -MH |<1 ise maddede ihmal edilebilir düzeyde (A düzeyi); $1 \leq |\Delta$ -MH |<1,5 ise maddede orta düzeyde (B düzeyi); $|\Delta$ -MH |≥1,5 ise maddede önemli düzeyde (C düzeyi) DMF olduğu söylenir (Zieky, 1993). Araştırmada cinsiyet değişkeni açısından kız öğrenciler odak grup, erkek öğrenciler referans grup olarak ele alınırken, okul türü değişkeni açısından ise devlet okulu referans grup, odak özel okul odak grup olarak ele alınmıştır.

Lojistik Regresyon Tekniği: Bağımlı değişkenin iki değer alabildiği ve bağımsız değişkenin sürekli bir değişken olduğu durumlarda kullanılan bir modeldir. Lojistik regresyon, hem tek biçimli hem de tek biçimli olmayan DMF'yi belirler. LR tekniği ile DMF belirlenirken, madde etkisi değişkeni anlamlı olduğunda maddede DMF olduğu, grup-madde etkileşim değişkeni anlamlı olduğunda tek biçimli olmayan değişen madde fonksiyonu olduğu sonucuna varılır. Bu çalışmada istatistiksel önemliliğe karar verirken manidarlık düzeyi p=0,05 olarak alınmıştır. Lojistik regresyon modelinde standartlaştırılmış regresyon katsayıları etki büyüklüğü olarak yorumlanır (Gierl, Jodoin & Ackerman, 2000). Standartlaştırılmış regresyon katsayıları (R^2) DMF'nin derecesini verir ve üç düzeyde belirlenir (Jodoin & Gierl, 2001). $R^2 < 0,035$ ise DMF yoktur ya da ihmal edilebilir düzeydedir, $0,035 \leq R^2 < 0,070$ ise orta düzeyde DMF vardır, $R^2 \geq 0,070$ ise önemli düzeyde DMF vardır. Bir maddenin DMF içeren (B veya C düzeyi) bir madde olarak sınıflandırılabilmesi için ki-kare değerinin 0,05'e eşit veya 0,05'den küçük ve R^2 değerinin de en az 0,035 olması gerekir (Zumbo,1999).

Bu çalışmada maddelere ilişkin verileri düzenlemede SPSS 11,5 paket programı kullanılmıştır. DMF'yi belirlemede yararlanılan Mantel-Haenszel ve lojistik regresyon teknikleri için EZDIF paket programı işe koşulmuştur. Lojistik regresyon analizinde DMF veren maddeleri saptamak için Zumbo (1999)'nun $\Delta R^2 = R^2$ (step3)- R^2 (step1) formülü ile hesaplanan sınıflamasından yararlanılmıştır. ΔR^2 sonuçları SPSS'te Zumbo (1999) tarafından yazılan özel bir script ile elde edilmiştir.

Bulgular ve Yorum

Bu bölümde cinsiyet ve okul türü bakımından matematik ve fen bilgisi alt testlerinde DMF olup olmadığı ele alınmıştır. Bu amaçla kullanılan LR ve MH sonuçları verilmiş ve bu tekniklerden elde edilen sonuçlar ve sonuçlar arasındaki uyum incelenmiştir.

Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırmak amacıyla kullanılan matematik alt testinde cinsiyete göre madde fonksiyonları farklılaşmakta mıdır?

Matematik alt testinde cinsiyete göre DMF gösteren maddelere ilişkin MH tekniğine ait sonuçlar Tablo 1'de ve LR tekniğine ait sonuçlar ise Tablo 2'de verilmiştir.

Tablo 1.

Cinsiyet Değişkenine Göre Matematik Alt Testine İlişkin Mantel-Haenszel Analiz Sonuçları

| Soru no | α | χ^2 | p | Δ -MH | SE | DMF düzeyi |
|---------|----------|----------|-------|--------------|-------|------------|
| 2 | 1.161 | 16.106 | 0.000 | -0.351 | 0.087 | A |
| 3 | 1.130 | 11.184 | 0.001 | -0.287 | 0.085 | A |
| 5 | 0.912 | 4.209 | 0.040 | 0.216 | 0.104 | A |
| 6 | 0.892 | 8.184 | 0.004 | 0.269 | 0.094 | A |
| 14 | 0.855 | 14.898 | 0.000 | 0.368 | 0.095 | A |
| 16 | 1.110 | 5.722 | 0.017 | -0.246 | 0.102 | A |
| 18 | 1.087 | 4.141 | 0.042 | -0.196 | 0.095 | A |
| 19 | 1.108 | 6.692 | 0.010 | -0.240 | 0.092 | A |
| 23 | 0.816 | 20.350 | 0.000 | 0.479 | 0.106 | A |

Referans grup: Kız (n=7255)

Odak grup: Erkek (n=8053)

MH sonuçlarına göre, 25 maddeden oluşan matematik alt testinde 9 madde A düzeyinde DMF içermektedir, B ya da C düzeyinde DMF içeren maddeye rastlanmamıştır. Bu sonuca göre, matematik alt testinde sadece ihmal edilebilir düzeyde DMF içeren madde bulunmaktadır. MH analiz sonuçları incelendiğinde; 2, 3, 16, 18 ve 19 numaralı maddelerin kız öğrencilerden yana; 5, 6, 14 ve 23 numaralı maddelerin erkek öğrencilerden yana DMF içerdiği görülmektedir.

Tablo 2.

Matematik Alt Test Maddelerinin Cinsiyet Değişkenine Göre Lojistik Regresyon Analiz Sonuçları

| Madde no | Madde etkisi p | Grup x madde etkileşimi p | DMF | | DMF ΔR^2 | DMF türü | DMF düzeyi |
|----------|-------------------|------------------------------|----------|-------|------------------|----------|------------|
| | | | χ^2 | p | | | |
| 2 | 0.036 | 0.503 | 13,740 | 0,001 | 0,002 | TB | A |
| 3 | 0.048 | 0.652 | 11,928 | 0,003 | 0,001 | TB | A |
| 14 | 0.002 | 0.159 | 18,170 | 0,000 | 0,001 | TB | A |

Tablo 3'te yer alan bilgilerden lojistik regresyon analiz sonuçlarına göre 3 maddede tek biçimli DMF olduğu, 22 maddede ise DMF olmadığı görülmektedir. Ancak bu üç maddede elde edilen DMF'lerin düzeyi göz ardı edilebilir niteliktedir. DMF düzeyini belirlemede kullanılan sınıflamaya göre LR analizinde maddelerin tümü A düzeyinde DMF göstermektedir.

Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırmak amacıyla kullanılan matematik alt testinde okul türüne göre madde fonksiyonları farklılaşmakta mıdır?

Matematik alt testinde okul türüne göre DMF veren maddelere ilişkin MH tekniğine ait sonuçlar Tablo 3'te ve LR tekniğine ait sonuçlar ise Tablo 4'te verilmiştir.

Tablo 3.

Matematik Alt Testine İlişkin Okul Türü Değişkenine Göre Mantel-Haenszel Analiz Sonuçları

| Soru no | α | χ^2 | p | Δ -MH | SE | DMF düzeyi |
|---------|----------|----------|-------|--------------|-------|------------|
| 1 | 0.459 | 369.308 | 0.000 | 1.829 | 0.096 | C |
| 2 | 0.590 | 166.458 | 0.000 | 1.240 | 0.096 | B |
| 3 | 0.657 | 112.917 | 0.000 | 0.987 | 0.093 | A |
| 4 | 0.791 | 33.957 | 0.000 | 0.550 | 0.094 | A |
| 6 | 0.861 | 11.767 | 0.001 | 0.352 | 0.102 | A |
| 7 | 0.646 | 110.197 | 0.000 | 1.025 | 0.098 | B |
| 9 | 0.913 | 4.416 | 0.036 | 0.214 | 0.100 | A |
| 11 | 1.269 | 29.308 | 0.000 | -0.559 | 0.103 | A |
| 12 | 1.290 | 34.067 | 0.000 | -0.599 | 0.103 | A |
| 13 | 1.212 | 19.420 | 0.000 | -0.452 | 0.102 | A |
| 14 | 0.673 | 83.875 | 0.000 | 0.931 | 0.101 | A |
| 16 | 0.895 | 5.523 | 0.019 | 0.260 | 0.109 | A |
| 17 | 1.688 | 124.050 | 0.000 | -1.230 | 0.111 | B |
| 18 | 1.515 | 75.255 | 0.000 | -0.976 | 0.112 | A |
| 19 | 1.395 | 53.510 | 0.000 | -0.783 | 0.107 | A |
| 20 | 1.343 | 36.431 | 0.000 | -0.692 | 0.114 | A |
| 21 | 1.198 | 13.040 | 0.000 | -0.425 | 0.117 | A |
| 24 | 1.367 | 39.527 | 0.000 | -0.735 | 0.117 | A |
| 25 | 0.911 | 4.223 | 0.040 | 0.218 | 0.105 | A |

Referans grup: devlet okulu (n=7343)

Odak grup: özel okul (n=6767)

MH sonuçlarına göre, matematik alt testinde 25 sorudan 15 madde A düzeyinde, 3 madde B düzeyinde, 1 maddede C düzeyinde olmak üzere toplam 19 madde DMF içermektedir. B ve C düzeyinde DMF içeren maddelerin incelenmesi gerekmektedir. MH analiz sonuçları incelendiğinde; 11, 12, 13, 17, 18, 19, 20, 21, 24 ve 25 numaralı maddelerin devlet okulunda

okuyan öğrencilerin lehine; 1, 2, 3, 4, 6, 7, 9, 14, 16 ve 25 numaralı maddelerin özel okulda okuyan öğrencilerin lehine DMF içerdiği görülmektedir.

Tablo 4.

Matematik Alt Test Maddelerinin Okul Türü Değişkenine Göre Lojistik Regresyon Analiz Sonuçları

| Madde no | Madde etkisi | Grup x madde etkileşimi | DMF | | DMF ΔR^2 | DMF türü | DMF düzeyi |
|----------|--------------|-------------------------|----------|---------|------------------|----------|------------|
| | p | p | χ^2 | p | | | |
| 1 | 0,000 | 0,847 | 326,408 | p=0,000 | 0,022 | TB | A |
| 2 | 0,000 | 0,878 | 134,076 | p=0,000 | 0,009 | TB | A |
| 3 | 0,000 | 0,000 | 137,076 | p=0,000 | 0,012 | TBO | A |
| 4 | 0,000 | 0,000 | 78,072 | p=0,000 | 0,007 | TBO | A |
| 7 | 0,000 | 0,688 | 77,956 | p=0,000 | 0,005 | TB | A |
| 11 | 0,010 | 0,240 | 56,971 | p=0,000 | 0,000 | TB | A |
| 12 | 0,000 | 0,012 | 71,926 | p=0,000 | 0,005 | TBO | A |
| 13 | 0,014 | 0,625 | 44,166 | p=0,000 | 0,004 | TB | A |
| 14 | 0,000 | 0,000 | 74,964 | p=0,000 | 0,005 | TBO | A |
| 17 | 0,000 | 0,427 | 189,248 | p=0,000 | 0,017 | TB | A |
| 18 | 0,000 | 0,035 | 131,529 | p=0,000 | 0,012 | TBO | A |
| 19 | 0,000 | 0,103 | 83,723 | p=0,000 | 0,008 | TB | A |
| 20 | 0,000 | 0,000 | 101,848 | p=0,000 | 0,008 | TBO | A |
| 21 | 0,000 | 0,005 | 48,067 | p=0,000 | 0,004 | TBO | A |
| 22 | 0,002 | 0,022 | 18,350 | p=0,000 | 0,001 | TBO | A |
| 24 | 0,000 | 0,378 | 78,131 | p=0,000 | 0,009 | TB | A |
| 25 | 0,035 | 0,078 | 1,845 | p=0,398 | 0,000 | TB | A |

Tablo 4'te elde edilen LR analiz sonuçlarına göre, 8 maddede DMF olmadığı, 17 maddede ise DMF olduğu görülmektedir. DMF olduğu gözlenen maddelerden 1, 2, 7, 11, 13, 17, 19, 24 ve 25 numaralı maddelerde tek biçimli (TB) DMF, diğer maddelerde ise tek biçimli olmayan (TBO) DMF vardır. DMF düzeyini belirlemede kullanılan sınıflamaya göre, LR analizinde tüm maddeler A düzeyinde DMF içermektedir.

Matematik alt testinde yapılan analizlerde iki teknik arasında uyum var mıdır?

Matematik alt testinde cinsiyet değişkenine göre yapılan analizlerde, MH tekniği ile 9 maddede, LR tekniği ile 3 maddede A düzeyinde DMF bulunmuştur. 9 maddenin 3'ünde hem MH hem de LR teknikleri ile DMF olduğu görülmüştür. Bu maddeler; 2, 3 ve 14 numaralı maddelerdir. DMF düzeyi bakımından uyum gözlenirken DMF gösteren madde sayısı bakımından MH ve LR teknikleri çok düşük uyum göstermektedir.

Matematik alt testinde okul türü değişkenine göre yapılan analizlerde, MH tekniği ile 19 maddede, LR tekniği ile 17 maddede aynı düzeyde DMF bulunmuştur. Yirmi beş maddenin 16'sında hem MH hem de LR teknikleri ile DMF olduğu görülmüştür. Bu maddeler; 1, 2, 3, 4, 7, 11, 12, 13, 14, 17, 18, 19, 20, 21, 24 ve 25 numaralı maddelerdir. DMF veren madde sayısı bakımından yüksek bir uyum gözlenmektedir. Bu maddelerden MH tekniğine göre 1. madde C düzeyinde, 2. 7. ve 17. maddeler B düzeyinde DMF göstermektedir; fakat LR tekniğine göre bu maddelerde göz ardı edilebilir düzeyde DMF olduğu gözlenmektedir. Diğer maddelerde ise her iki tekniğe göre de A düzeyinde DMF olduğu görülmektedir. Maddelerin DMF düzeyleri dikkate alındığında MH tekniği ile LR tekniği arasında düşük düzeyde bir uyum olduğu söylenebilir.

MH ve LR tekniklerinden elde edilen χ^2 değerleri arasında önemli bir ilişki olup olmadığını incelemek için Spearman'ın sıra farkları korelasyon katsayısı kullanılmış ve matematik alt testinde sırasıyla cinsiyet ve okul türü değişkenlerine göre iki teknik için elde edilen χ^2 değerleri arasında önemli bir ilişki olmadığı ($r=-0,50$, $p=0,67$; $r=-0,43$, $p=0,08$) görülmüştür.

Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırmak amacıyla kullanılan fen bilgisi alt testinde cinsiyete göre madde fonksiyonları farklılaşmakta mıdır?

Fen bilgisi alt testinde cinsiyete göre DMF veren maddelere ilişkin MH tekniğine ait sonuçlar Tablo 5'te ve LR tekniğine ait sonuçlar Tablo 6'da verilmiştir.

Tablo 5.

Fen Bilgisi Alt Testine İlişkin Cinsiyet Değişkenine Göre Mantel-Haenszel Analiz Sonuçları

| Soru no | α | χ^2 | p | Δ -MH | SE | DMF düzeyi |
|---------|----------|----------|-------|--------------|-------|------------|
| 2 | 1.091 | 5.887 | 0.015 | -0.205 | 0.084 | A |
| 10 | 0.913 | 5.268 | 0.022 | 0.214 | 0.092 | A |
| 11 | 1.117 | 8.416 | 0.004 | -0.259 | 0.089 | A |
| 12 | 0.896 | 8.967 | 0.003 | 0.259 | 0.086 | A |
| 14 | 0.914 | 5.833 | 0.016 | 0.211 | 0.087 | A |
| 15 | 0.799 | 29.145 | 0.000 | 0.527 | 0.097 | A |
| 16 | 0.896 | 7.112 | 0.008 | 0.258 | 0.096 | A |
| 21 | 1.245 | 37.478 | 0.000 | -0.514 | 0.084 | A |
| 22 | 1.119 | 9.718 | 0.002 | -0.264 | 0.084 | A |
| 23 | 1.148 | 13.667 | 0.000 | -0.325 | 0.088 | A |

Referans grup: kız (n=7255)

Odak grup: erkek (n= 8053)

Tablo 5'te verilen MH sonuçlarına göre, fen bilgisi alt testindeki 25 sorudan 10 soru DMF içermektedir. DMF içeren maddelerin tümü A düzeyinde DMF içermektedir. Fen bilgisi alt testinde cinsiyete göre maddelerin tamamı ihmal edilebilir düzeyde DMF içermektedir. MH sonuçları incelendiğinde; 2, 11, 21, 22 ve 23 numaralı maddelerin kız öğrencilerin lehine; 12, 14, 15 ve 16 numaralı soruların erkek öğrencilerin lehine DMF içerdiği görülmektedir.

Tablo 6.

Fen Bilgisi Alt Test Maddelerinin Cinsiyet Değişkenine Göre Lojistik Regresyon Analiz Sonuçları

| Madde no | Madde etkisi | Grup x madde etkileşimi | DMF | | DMF ΔR^2 | DMF türü | DMF düzeyi |
|----------|--------------|-------------------------|----------|---------|------------------|----------|------------|
| | p | | χ^2 | p | | | |
| 4 | 0.018 | 0,009 | 6,885 | p=0,032 | 0,000 | TBO | A |
| 6 | 0,000 | 0,000 | 14,157 | p=0,001 | 0,001 | TBO | A |
| 7 | 0.043 | 0.044 | 4,191 | p=0,123 | 0,001 | TBO | A |
| 13 | 0,000 | 0,000 | 16,784 | p=0,000 | 0,002 | TBO | A |
| 15 | 0,006 | 0.715 | 35,466 | p=0,000 | 0,003 | TB | A |
| 19 | 0.015 | 0,009 | 6,480 | p=0,039 | 0,001 | TBO | A |
| 20 | 0,002 | 0,003 | 9,655 | p=0,008 | 0,001 | TBO | A |
| 22 | 0,000 | 0,005 | 15,805 | p=0,000 | 0,001 | TBO | A |
| 25 | 0,009 | 0.031 | 7,362 | p=0,025 | 0,001 | TBO | A |

Tablo 6'da verilen lojistik regresyon sonuçlarına göre, 9. maddede DMF olduğu görülmektedir. Maddelerin tümünün DMF düzeyi göz ardı edilebilir niteliktedir. DMF olduğu gözlenen sorulardan 15. maddede tek biçimli (TB) DMF, diğer maddelerde ise tek biçimli olmayan (TBO) DMF vardır.

Mantel-Haenszel ve lojistik regresyon tekniklerini karşılaştırmak amacıyla kullanılan fen bilgisi alt testinde okul türüne göre madde fonksiyonları farklılaşmakta mıdır?

Matematik alt testinde okul türüne göre DMF veren maddelere ilişkin MH tekniğine ait sonuçlar Tablo 7'de ve LR tekniğine ait sonuçlar Tablo 8'de verilmiştir.

Tablo 7.

Fen Bilgisi Alt Testine İlişkin Okul Türü Değişkenine Göre Mantel-Haenszel Analiz Sonuçları

| Soru no | α | χ^2 | p | Δ -MH | SE | DMF düzeyi |
|---------|----------|----------|-------|--------------|-------|------------|
| 1 | 1.140 | 9.696 | 0.002 | -0.308 | 0.098 | A |
| 2 | 0.743 | 48.542 | 0.000 | 0.698 | 0.100 | A |
| 5 | 0.613 | 112.344 | 0.000 | 1.152 | 0.108 | B |
| 6 | 1.118 | 7.489 | 0.006 | -0.261 | 0.095 | A |
| 9 | 0.898 | 6.044 | 0.014 | 0.252 | 0.101 | A |
| 10 | 0.876 | 8.906 | 0.003 | 0.310 | 0.103 | A |
| 11 | 0.850 | 13.839 | 0.000 | 0.382 | 0.102 | A |
| 12 | 1.213 | 20.109 | 0.000 | -0.453 | 0.101 | A |
| 14 | 0.738 | 51.070 | 0.000 | 0.715 | 0.100 | A |
| 16 | 1.124 | 5.952 | 0.015 | -0.275 | 0.111 | A |
| 18 | 0.892 | 7.205 | 0.007 | 0.268 | 0.099 | A |
| 19 | 0.759 | 37.399 | 0.000 | 0.648 | 0.105 | A |
| 22 | 0.835 | 17.436 | 0.000 | 0.423 | 0.101 | A |
| 23 | 0.882 | 8.127 | 0.004 | 0.294 | 0.102 | A |
| 25 | 0.916 | 3.969 | 0.046 | 0.206 | 0.102 | A |

Referans grup: devlet okulu (n=7343) Odak grup: özel okul (n=6767)

MH sonuçlarına göre, 25 sorudan 14 madde A düzeyinde, 1 madde B düzeyinde olmak üzere toplam 15 madde DMF içermektedir. Ancak B düzeyinde yanlılık içeren maddenin incelenmesi gerekmektedir. MH analiz sonuçları incelendiğinde; 1, 6, 12 ve 16 numaralı maddelerin devlet okullarında okuyan öğrencilerin lehine; 2, 5, 9, 10, 11, 14, 18, 19, 22, 23 ve 25 numaralı maddelerin özel okullarda okuyan öğrencilerin lehine DMF içerdiği görülmektedir.

Tablo 8.

Fen Bilgisi Alt Test Maddelerinin Okul Türü Değişkenine Göre Lojistik Regresyon Analiz Sonuçları

| Madde no | Madde etkisi | Grup x madde etkileşimi | | DMF | | DMF ΔR^2 | DMF türü | DMF düzeyi |
|----------|--------------|-------------------------|----------|---------|-------|------------------|----------|------------|
| | p | p | χ^2 | p | | | | |
| 1 | 0,000 | 0,000 | 69,751 | p=0,000 | 0,005 | TBO | A | |
| 2 | 0,000 | 0,000 | 58,70 | p=0,000 | 0,004 | TBO | A | |
| 3 | 0,000 | 0,000 | 27,807 | p=0,000 | 0,002 | TBO | A | |
| 5 | 0,000 | 0,775 | 84,743 | p=0,000 | 0,006 | TB | A | |
| 6 | 0,000 | 0,003 | 37,255 | p=0,000 | 0,003 | TBO | A | |
| 7 | 0,034 | 0,458 | 12,421 | p=0,002 | 0,001 | TB | A | |
| 11 | 0,001 | 0,000 | 20,917 | p=0,000 | 0,002 | TBO | A | |
| 12 | 0,000 | 0,261 | 45,173 | p=0,000 | 0,005 | TB | A | |
| 14 | 0,000 | 0,000 | 51,227 | p=0,000 | 0,003 | TBO | A | |
| 15 | 0,000 | 0,000 | 107,661 | p=0,000 | 0,001 | TBO | A | |
| 16 | 0,000 | 0,000 | 65,982 | p=0,000 | 0,006 | TBO | A | |
| 17 | 0,025 | 0,223 | 9,091 | p=0,011 | 0,001 | TB | A | |
| 18 | 0,000 | 0,000 | 13,360 | p=0,001 | 0,001 | TBO | A | |
| 19 | 0,000 | 0,005 | 19,174 | p=0,000 | 0,002 | TBO | A | |
| 20 | 0,000 | 0,000 | 31,803 | p=0,000 | 0,003 | TBO | A | |
| 23 | 0,007 | 0,008 | 7,417 | p=0,025 | 0,000 | TBO | A | |
| 24 | 0,000 | 0,000 | 121,184 | p=0,000 | 0,009 | TBO | A | |
| 25 | 0,012 | 0,006 | 6,910 | p=0,032 | 0,000 | TBO | A | |

Tablo 8’de elde edilen lojistik regresyon analiz sonuçlarına göre, 18 maddede DMF olduğu görülmektedir. DMF olduğu gözlenen maddelerden 5, 7, 12 ve 17. maddelerde tek biçimli DMF, diğer maddelerde ise tek biçimli olmayan DMF vardır. DMF düzeyini belirlemede kullanılan sınıflamaya göre LR analizinde maddelerin tümü A düzeyinde DMF göstermektedir.

MH tekniği ile 15 maddede, LR tekniği ile 18 maddede farklı düzeylerde DMF bulunmuştur. 21 maddenin 11’inde hem MH hem de LR teknikleri ile DMF olduğu görülmüştür. Bu maddeler; 1, 2, 5, 6, 11, 12, 14, 16, 18, 19, 23 ve 25 numaralı maddelerdir.

Fen bilgisi alt testinde yapılan analizlerde iki teknik arasında uyum var mıdır?

Fen bilgisi alt testinde cinsiyet değişkenine göre yapılan analizlerde, MH tekniği ile 10 maddede, LR tekniği ile 9 maddede A düzeyinde DMF bulunmuştur. Tablo 5 ile Tablo 6 karşılaştırıldığında, 17 maddenin 2’sinde hem MH tekniği hem de LR tekniği ile DMF olduğu görülmüştür. Bu maddeler; 15. ve 22. maddelerdir. Buna göre DMF gösteren maddelerin çoğu iki teknikte de farklılık göstermektedir. Tablo 5’teki MH sonuçları ile Tablo 6’daki LR sonuçları karşılaştırıldığında, iki teknikte de DMF’nin A düzeyinde elde edilmesi dışında iki teknik düşük düzeyde bir uyum göstermektedir.

Fen bilgisi alt testinde okul türü değişkenine göre yapılan analizlerde, MH sonuçları ile LR sonuçları karşılaştırıldığında, MH tekniğine göre 5. madde B düzeyinde DMF gösterirken, LR tekniğine göre bu maddede DMF olmadığı gözlenmiştir. İki teknikte de DMF’nin A düzeyinde elde edilmesi dışında tam bir uyum gözlenmemektedir.

MH ve LR tekniklerinden elde edilen χ^2 değerleri arasında önemli bir ilişki olup olmadığını incelemek için Spearman’ın sıra farkları korelasyon katsayısı kullanılmış ve fen bilgisi alt testinde sırasıyla cinsiyet ve okul türü değişkenlerine göre iki teknik için elde edilen χ^2 değerleri arasında önemli bir ilişki olmadığı ($r = -0,20$, $p=0,61$; $r = -0,29$, $p=0,30$) görülmüştür.

Sonuç ve Öneriler

Sonuçlar

Matematik alt testinde cinsiyete dayalı olarak yapılan analizlerden elde edilen sonuçlar incelendiğinde, MH sonuçlarına göre 25 sorudan 9’u A düzeyinde DMF içermektedir, B ya da C düzeyinde DMF’ye rastlanmamıştır. Matematik alt testinde DMF içeren 9 maddenin tamamı ihmal edilebilir düzeyde iken, LR analiz sonuçlarına göre, 3 maddede göz ardı edilebilir düzeyde DMF olduğu görülmüştür. Yurdugül (2003)’ün çalışmasında, 2001 yılı OKS matematik alt testinde yer alan bir maddenin cinsiyet değişkenine göre B düzeyinde DMF içerdiği görülmüştür. İki teknikten elde edilen sonuçlar karşılaştırıldığında, DMF’nin A düzeyinde elde edildiği görülmektedir. Bu açıdan MH ve LR teknikleri uyum göstermektedir. Ancak DMF gösteren madde sayısı bakımından iki teknik arasında fark vardır. Bu açıdan çalışmada elde edilen sonuç Yurdugül (2003)’ün çalışmasında elde ettiği sonuç ile tutarlılık göstermektedir.

Matematik alt testinde okul türüne dayalı olarak yapılan analizlerden elde edilen sonuçlar incelendiğinde, MH sonuçlarına göre 25 sorudan 15’i A düzeyinde, 3’ü B düzeyinde, 1’i C düzeyinde olmak üzere toplam 19 maddenin DMF içerdiği görülmüştür. B ve C düzeyinde DMF içeren maddelerin incelenmesi gerekmektedir. Bu maddelerden 1. madde C düzeyinde, 2. ve 7. maddeler B düzeyinde DMF göstermekte olup özel okulda okuyan öğrencilerin lehine işlemiştir. 17. madde ise B düzeyinde DMF göstermekte olup devlet okulunda okuyan öğrencilerin lehine işlemiştir. Diğer maddeler ihmal edilebilir düzeyde DMF içermektedir. MH tekniği ile 19 maddede, LR tekniği ile 17 maddede aynı düzeyde DMF bulunmuştur. 20 maddenin 16’sında hem MH hem de LR teknikleri ile DMF olduğu görülmüştür. Bu maddelerden MH tekniğine göre 1. madde C düzeyinde, 2. 7. ve 17. maddeler B düzeyinde DMF göstermektedir. LR tekniğine göre bu maddelerde göz ardı edilebilir düzeyde DMF olduğu gözlenmektedir. Araştırmada elde edilen bu sonuç MH tekniğinin LR tekniğine göre daha fazla sayıda B düzeyinde DMF içeren madde tespit

ettiğini gösterirken, bu durum bazı araştırmalarda elde edilen sonuçlar ile çelişmektedir (Wiberg, 2009). Diğer maddelerde ise her iki tekniğe göre de A düzeyinde DMF olduğu görülmektedir. Bu açıdan MH analiz sonuçları ile LR analiz sonuçları karşılaştırıldığında, MH tekniği ile LR tekniği arasında düşük düzeyde bir uyum olduğu görülmüştür. Araştırmadan elde edilen bu sonuçlar bazı araştırmalarla tutarlılık gösterirken (Gierl, Khaliq & Boughton 1999), Gierl, Rogers & Klinger (1999) ve Bertrand & Bouteau (2003)'ün yaptıkları çalışmalarda MH ve LR teknikleri arasında yüksek düzeyde uyum bulduklarına dair araştırma sonuçlarıyla çelişmektedir. Bununla birlikte araştırmada elde edilen sonuç, Gierl, Khaliq & Boughton (1999) ve Hidalgo & Lopez-Pina (2004) 'ün yaptığı çalışmada MH tekniğinin LR tekniğine göre daha az sayıda DMF'li madde belirlediği ve bu konuda daha tutucu olduğu sonucu ile çelişmektedir.

Fen bilgisi alt testinde cinsiyete dayalı olarak yapılan analizlerden elde edilen sonuçlar incelendiğinde, MH tekniği ile 10 maddede, LR tekniği ile 9 maddede A düzeyinde DMF bulunmuştur. 17 maddenin 2'sinde hem MH tekniği hem de LR tekniği ile DMF olduğu görülmüştür. MH analiz sonuçları ile LR analiz sonuçları karşılaştırıldığında DMF'nin A düzeyinde elde edildiği görülmektedir. DMF'nin A düzeyinde elde edilmesi dışında iki teknik düşük düzeyde bir uyum göstermektedir. Bu sonuç, Gierl, Khaliq & Boughton (1999)'in bir fen bilgisi testinin cinsiyet değişkeni üzerinde yaptığı çalışmada elde ettiği DMF'li maddelerin farklı tekniklere göre tutarlı sonuçlar göstermemesi bulgusu ile benzerlik göstermektedir.

Fen bilgisi alt testinde okul türüne dayalı olarak yapılan analizlerden elde edilen sonuçlar incelendiğinde, MH sonuçlarına göre 25 sorudan 14 madde A düzeyinde, 1 madde B düzeyinde olmak üzere toplam 15 madde DMF içermektedir. LR sonuçlarına göre ise, 7 maddede DMF olmadığı, 18 maddede ise DMF olduğu görülmektedir. DMF düzeyini belirlemede kullanılan sınıflamaya göre LR analizinde maddelerin tümü A düzeyinde DMF göstermektedir. MH tekniği ile 15 maddede, LR tekniği ile 18 maddede farklı düzeylerde DMF bulunmuştur. 21 maddenin 11'inde hem MH hem de LR tekniği ile DMF olduğu görülmüştür. MH analiz sonuçları ile LR analiz sonuçları karşılaştırıldığında, 5. madde dışında DMF'nin A düzeyinde elde edildiği görülmektedir. Fakat MH tekniğine göre 5. madde B düzeyinde DMF gösterirken, LR tekniğine göre bu maddede DMF olmadığı gözlenmiştir. Bu açıdan fen bilgisi alt testi okul türü değişkenine göre MH tekniği ile LR tekniği yüksek düzeyde bir uyum göstermemektedir. Araştırmada elde edilen bu sonuç Zheng, Gierl & Cui (2007) tarafından MH ile LR teknikleri arasında bulunan yüksek uyum ile tutarlılık göstermemektedir.

Araştırmanın sonucunda, cinsiyet ve okul türü bakımından matematik ve fen bilgisi testlerinde LR tekniği ile DMF bulunan maddelerin tamamı ihmal edilebilir düzeyde iken, MH tekniğinde çoğunluğu ihmal edilebilir düzeyde bulunmuştur. İki teknikte DMF bulunan maddeler arasındaki benzerlik çok düşük düzeydedir. Bu bakımdan, MH ve LR teknikleri arasında sadece düşük düzeyde DMF bulmaları bakımından benzerlik olduğu söylenebilir. Araştırmada elde edilen bu sonuç, DMF'nin büyüklüğü ile ilgili olarak Wiberg (2009) tarafından yapılan çalışmada elde edilen sonuçlar ile tutarlılık gösterirken, DMF'li maddeleri belirleme açısından ise Swaminathan & Rogers (1990) ve Wiberg (2009)'un yaptıkları çalışmalarda elde ettikleri sonuç ile tutarlılık göstermemektedir. Bununla birlikte araştırmada elde edilen sonuçlar incelendiğinde, fen bilgisi testi okul türü değişkeni dışında genel olarak MH tekniğine göre elde edilen DMF'li madde sayısının, LR tekniğinden daha fazla olduğu görülmüştür. Bu sonuca göre MH tekniğinin LR tekniğinden daha duyarlı olduğu söylenebilir. Bunun nedenleri arasında, MH tekniğinde odak ve referans grupların madde parametre tahminleri aynı anda yapıldığı için toplam örneklem sayısı LR'ye göre daha fazla olabilmektedir. Bu açıdan LR'nin daha az sayıda DMF'li madde bulmasının sebebi örneklem büyüklüğüne bağlı olabilir; fakat LR ve MH tekniklerinin MTK'ya dayalı tekniklere göre daha küçük örneklem gerektirmesi de bu tekniklerin bir avantajı olarak görülebilir (Penfield & Camilli, 2007). Ayrıca araştırmada fen bilgisi alt testinin matematik alt testinden daha fazla DMF'li maddeye sahip olduğu görülmüştür. Araştırmada elde edilen bu sonuçlar, Greer (2004) tarafından yapılan çalışmada elde edilen sonuçlar ile tutarlılık göstermektedir.

Bununla birlikte alanyazındaki yapılan çalışmalar incelendiğinde, DMF'li madde sayısı bakımından teknikler arasında bir uyum görülmemesinin nedenleri arasında maddelerin güçlük ve ayırt ediciliklerinin farklı olması, örneklem büyüklükleri, grup ortalamalarındaki farklılıklar, yetenek farklılıkları gibi faktörler gösterilmektedir (Hidalgo ve Pina, 2004; Narayanan ve Swaminathan, 1996; Clauser, Mazor & Hambleton, 1991; Bertrant & Boiteau, 2003; Swaminathan & Rogers, 1990; Mazor, Clauser & Hambleton, 1992; Fidalgo, Mellenbergh & Muñiz, 2000). MH ve LR teknikleri ortalamalar arası grup farklılıklarından büyük ölçüde etkilenmektedir. Dolayısıyla test maddelerindeki ortalamaların farklı olması sonuçları etkileyeceğinden, bu tür durumlarda her iki teknik arasında uyum görülmeyebilir. Bununla birlikte MH tekniği çok yüksek ayırtıcılığa sahip maddeleri tanımlarken, çok zor maddelerdeki DMF'yi tespit edemeyebilir (Clauser, Mazor & Hambleton, 1991) ve yetenek dağılımlarının eşit olmaması durumunda da LR tekniği I. Tip hata oranını artırmaktadır (Narayanan ve Swaminathan, 1996).

Araştırma sonuçlarına bakıldığında, MH istatistiğinin tek biçimli olmayan DMF'yi belirleyemediği bulunmuştur. Araştırmanın bu sonucu diğer araştırma sonuçlarıyla destelenmektedir (Mellenbergh, 1982; Hambleton & Rogers, 1989; Narayanan & Swaminathan, 1996). Araştırmada LR tekniğinin avantajının tek biçimli olan ve olmayan DMF'yi tanımlamada kullanılabilen modele dayalı bir yaklaşım sağlaması olduğundan, LR tekniğinin sadece DMF'nin bulunduğu maddeleri belirlemede kalmadığı, aynı zamanda DMF'nin doğasına ilişkin çok daha iyi bilgi elde edilmesini sağladığı görülmüştür. Bununla birlikte bu çalışmada ve yapılan diğer araştırmalarda LR'nin tekbiçimli olmayan DMF'yi belirlemede MH'den daha güçlü olduğu, tek biçimli DMF'yi belirlemede de MH kadar güçlü olduğu gösterilmiştir (Swaminathan & Rogers, 1990). Sonuç olarak iki teknik de tekbiçimli DMF'yi belirlemede benzer performans göstermiş, fakat sadece LR tek biçimli olmayan DMF'yi belirleyebilmiştir. Bununla birlikte bazı araştırmacılar tek biçimli olmayan DMF'nin büyük ölçekli testlerde nadiren bulunduğunu iddia etmektedir (Camilli & Shepard, 1994).

Genel olarak araştırmadan elde edilen sonuçlar ve bu konuda yapılmış diğer araştırma sonuçları incelendiğinde, MH ya da LR tekniklerinin her durumda daha fazla ya da daha az DMF'li madde vermediği görülmüştür. Alanyazında pek çok çalışma MH tekniğinin pratik ve güçlü bir yol olduğunu ve daha tutarlı sonuçlar verdiğini belirtmekle birlikte (Narayanan & Swaminathan, 1994; Hambleton & Rogers, 1989; Clauser, Mazor, & Hambleton, 1993; Huang, 1998; Bertrant ve Boiteau, 2003), bazı araştırmalar ise LR tekniğinin en etkili ve tavsiye edilen tekniklerden biri olduğunu ifade etmektedir (Clauser & Mazor, 1998; Swaminathan & Rogers, 1990; Wiberg, 2007). Bu açıdan bakıldığında ve tekniklerin birbirlerine göre pek çok üstünlüğü ya da sınırlılığı olduğu göz önüne alındığında, bir tekniğin diğerine göre "teknik olarak tercih edilir" olması görüşünün, durumu belirlemede kesin bir yol olmadığı sonucuna varılabilir. Araştırmada elde edilen bu sonucun bazı araştırmalarla da desteklendiği görülmüştür (Holland & Thayer, 1988). Ayrıca bazı araştırmacılar bu konuyla ilgili olarak DMF'yi belirlemede hangi teknik ve tekniğin seçileceğinin çok önemli olmadığını ifade ederken (Robitzsch & Rupp, 2009), bazı araştırmacılar ise bu araştırmada kullanılan Wiberg (2007) tarafından belirtilen kriterlere göre seçilmesi gerektiğini savunmaktadırlar.

Son olarak, bu araştırmadan elde edilen sonuçlar ve konuyla ilgili diğer araştırma sonuçları birlikte değerlendirildiğinde, daha güvenilir sonuçlar elde etmek için birden çok DMF belirleme tekniğinin bir arada kullanılması ve farklı tekniklerin hepsinde tespit edilen maddelerin DMF'li madde olarak karşılaştırılması gerektiği ifade edilebilir. Araştırmada elde edilen bu sonucun tekniklerin karşılaştırılmasını amaçlayan diğer çalışmalarla da desteklendiği görülmüştür (Camilli & Shepard, 1994; Wiberg, 2007; Wiberg, 2009; Hidalgo & Pina, 2004). DMF çalışmalarında gerçek veri ile birlikte simulasyon verisinin kullanılması da daha güvenilir sonuçlar elde etmek, sonuçların doğruluğunu kontrol etmek ve hatalı olarak belirlenen maddeleri tespit etmek açısından önemli olabilir.

Öneriler

Araştırmanın sonuçlarına göre aşağıdaki önerilerde bulunulabilir:

Öncelikle, bu çalışmada DMF belirlemede MH ve LR teknikleri kullanılmıştır. Bunun dışında DMF'yi ortaya çıkarmaya yönelik diğer teknikler (IRT, SIBTEST, alan indeksleri yaklaşımı gibi) çalışmaya katılarak incelenebilir. Farklı tekniklerden elde edilen sonuçların karşılaştırılması yoluna gidilebilir.

Diğer yandan çalışmada bağımsız değişken olarak cinsiyet ve okul türü alınmıştır. Farklı değişkenler (sosyoekonomik düzey gibi) göz önüne alınarak incelemeler yapılabilir.

Kaynakça

- Agresti, A. (1984). *Analysis of ordinal categorical data*. New York: John Wiley & Sons.
- Bertrand, R. & Boiteau, N. (2003). Comparing the stability of IRT-Based and non IRT-based DIF methods in different cultural contexts using TIMSS data (Eric Document Reproduction No. ED 476 924).
- Camilli, G. & Shepard, L. A. (1994). *Methods for identifying biased test items*. London: Sage Publications.
- Clauser, B. E., Mazor, K. & Hambleton, R. K. (1991). *Examination of various influences on the Mantel-Haenszel statistic*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Clauser, B. E., Mazor, K. & Hambleton, R. K. (1993). The effects of purification of the matching criterion on the identification of DIF using the Mantel-Haenszel procedure. *Applied Measurement in Education*, 6(4), 269–279.
- Clauser, B. E., Mazor, K. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*. 17(1), 31–44.
- Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley and Sons.
- Dorans, N. J. & Holland, P. W. (1993). DIF detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35–66) Hillsdale, NJ: Lawrence Erlbaum.
- Fidalgo, A. M., Mellenbergh, G. J. & Muñiz, J. (2000). Spain effects of amount of DIF, test length, and purification. *Methods of Psychological Research*, 5(3) 43-53.
- Gierl, M. J., Jodoin, M. & Ackerman T. (2000). *Performance of mantel-haenszel, simultaneous item bias test and logistic regression when the proportion of dif items is large*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, USA.
- Gierl, M. H., Khaliq, S. N. & Boughton, K. (1999, June). *Gender differential item functioning in mathematics and science: prevalence and policy implications*. Paper presented at the Annual Meeting of the Canadian Society for the Study of Education, Canada.
- Gierl, M. J., Rogers, W. T. & Klinger, D. A. (1999). Statistical and judgmental reviews to identify and interpret translation differential item functioning. *Alberta Journal of Educational Research*, 45(4), 353-376.
- Greer, T. G. (2004). *Detection of differential item functioning (DIF) on the SATV: A comparison of four methods: Mantel-haenszel, logistic regression, simultaneous item bias and likelihood ratio test*. Unpublished PhD Dissertation, University of Houston.
- Hambleton, K. R. & Rogers, J. H. (1989). Detecting potentially biased test items: comparison of IRT area and MH methods. *Applied Measurement in Education*, 2(4), 313–334.
- Hidalgo, M. D. & Lopez-Pina, J. (2004). Differential item functioning detection and effect-size: A comparison between LR and MH procedures. *Educational and Psychological Measurement*, 64(6), 903–915.

- Holland, P. W. & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test Validity* (pp. 129-145). Hillsdale, NJ: Lawrence Erlbaum.
- Huang, C. Y. (1998). *Factors influencing the reliability of DIF detection methods*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.
- Jodoin, M. G. & Gierl, M. J. (2001). Evaluating power and type I error rates using an effect size with the logistic regression procedure for DIF. *Applied Measurement in Education, 14*, 329-349.
- Mazor, K. M., Clauser, B. E. & Hambleton, R. K. (1992). The effect of sample size on the functioning of the Mantel-Haenszel statistic. *Educational and Psychological Measurement, 52*, 443-452.
- Mellenbergh, G. J. (1982). Contingency table models for assessing item bias. *Journal of Educational Statistics, 7*, 105-118.
- Milli Eğitim Bakanlığı. (2009). *Ortaöğretim kurumlarına geçiş sistemi*. 12.03.2009 tarihinde <http://oges.meb.gov.tr/> adresinden alınmıştır.
- Nandakumar, R. (1993). A fortran 77 program for detecting differential item functioning through the mantel-haenszel statistic. *Educational and Psychological Measurement, 53*, 679-684.
- Narayanan, P. & Swaminathan, H. (1996). Identification of items that nonuniform DIF. *Applied Psychological Measurement, 20*(3), 257-274.
- Osterlind, J. S. (1983). *Test item bias*. London: Sage Publications.
- Öğretmen, T. ve Doğan, N. (2004). OKÖSYS matematik alt testine ait maddelerin yanlılık analizi. *İnönü Üniversitesi Eğitim Fakültesi Dergisi, 5*(8). SAYFA NUMARALARI YOK.
- Penfield, R. D. & Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao & S. Sinharay (Eds.), *Handbook of Statistics Psychometrics* (26, pp. 125-167). Amsterdam: Elsevier.
- Robitzsch, A. & Rupp, A. A. (2009). Impact of missing data on the detection of differential item functioning: the case of mantel-haenszel and logistic regression analysis. *Educational and Psychological Measurement, 69*(1), 18-34.
- Swaminathan, H. & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*, 361-370.
- Wiberg, M. (2007). *Measuring and detecting differential item functioning in criterion-referenced licensing test: a theoretic comparison of methods* (EM No. 60). Umea, Sweden: Umea University, Department of Educational Measurement.
- Wiberg, M. (2009). Differential item functioning in mastery tests: A comparison of three methods using real data. *International Journal of Testing, 9*, 41-59.
- Yurdugül, H. (2003). *Ortaöğretim kurumları seçme ve yerleştirme sınavının madde yanlılığı açısından incelenmesi*, Doktora tezi, Hacettepe Üniversitesi, Ankara.
- Zheng, Y., Gierl, M. J., & Cui, Y. (2007). *Using real data to compare DIF detection and effect size measures among Mantel-Haenszel, SIBTEST and logistic regression procedures*. Paper presented at NCME 2007, Chicago.
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland, and H. Wainer (Eds.), *Differential item functioning* (pp. 337-347). Hillsdale, NJ: Lawrence Erlbaum.
- Zumbo, B. D. A. (1999). *Handbook on the theory and methods of differential item functioning: Logistic regression modeling as a unitary framework for binary and likert-type item scores*. Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense.