

DİFERANSİYEL MADDE FONKSİYONUNA GENEL BİR BAKIŞ

Dr. Özlem Yeşim ÖZBEK
Gaziosmanpaşa Üniversitesi
Eğitim Fakültesi
Eğitimde Ölçme ve Değerlendirme

ÖZET

Geçerlik, belirli bir kavramsal yapının belirli bir testten elde edilen puanların altında yatan açıklamalarının potansiyel kaynağı olarak hipotez edilmesi şeklinde düşünülebilir. Testin ölçmüs olduğu ana kavramsal yapı yanında, test sonucunu etkileyebilecek diğer kavramsal yapıları da test puanlarının açıklamalarının potansiyel kaynakları olarak görmek yada bunları önceden tahmin etmek ve bu alternatif hipotezlerin kabul edilebilirliğini araştırmak geçerlik sürecinin vazgeçilmez bir parçasıdır. Son yıllarda Diferansiyel Madde Fonksiyonu (DMF) analizleri geçerlik çalışmalarında kavramsal yapıyla ilgili ya da ilgisiz etkenleri bulmak için gelecek vaad eden bir metod olarak tavsiye edilmiştir. Bu makale DMF metodları hakkında istatistiksel detaylara girmeden genel bilgi vermektedir.

Anahtar Kelimeler: Diferansiyel madde fonksiyonu, geçerlik, madde yanlılığı

SUMMARY

Validation can be thought of as hypothesizing a certain construct as a potential source of plausible explanations of scores on a particular test. However, recognizing or foreseeing other constructs as potential sources of these explanations of scores on the test and investigating the tenability of these alternative hypotheses are invaluable part of the validation process. Recently, Differential Item Functioning (DIF) analysis has been suggested as a promising method for the validity investigation to study construct relevant as well as irrelevant sources. This article provides general information on DIF methodology without going into statistical details.

Key Words: Differential item functioning, validity, item bias

GİRİŞ

Psikometrik testlerde geçerlik eğitimde ölçme ve değerlendirmenin en eski ve en tartışmalı konularından biridir. Yıllar boyunca geçerliliği bazı araştırmacılar (örneğin Messick, 1989) test geliştirmenin dışsal bir süreci olarak görürken bazıları da (örneğin Cleary, 1968) test geliştirmenin içsel bir süreci olarak görmüşlerdir. Geçerliliğin çeşitleri ve anlamı konusundaki tartışmaların yıllar geçmesine rağmen sonuçlanmamasına ve psikometrik test standartlarının 1960'ların başından beri sürekli değişmesine rağmen test geliştirmede geçerliliğin önemi hiç bir zaman azalmamıştır. Aksine, 1966, 1974, 1985 ve 1999 Eğitim ve Psikolojik Test Standartlarında (Standards of Educational and Psychological Testing: AERA, APA, NCME, 1966; 1974; 1985; 1999) testlerde geçerlilik çalışmalarının önemini sürekli vurgulanmıştır.

Geçerlik, belli bir kavramsal yapının belirli bir testten elde edilen puanların altında yatan açıklamaların potansiyel kaynağı olarak hipotez edilmesi şeklinde düşünülebilir. Testin ölçmüs olduğu ana kavramsal yapı yanında, test sonucunu etkileyebilecek diğer kavramsal yapıları da test puanlarının açıklamalarının potansiyel kaynakları olarak görmek ve ya bunları önceden tahmin etmek ve bu alternatif hipotezlerin kabul edilebilirliğini araştırmak geçerlik sürecinin vazgeçilemez bir parçasıdır. Son yıllarda Diferansiyel Madde Fonksiyonu (DMF) (Differential Item Functioning) analizleri geçerlik çalışmalarında kavramsal yapıyla ilgili ya da ilgisiz etkenleri bulmak için gelecek vaad eden bir metod olarak tavsiye edilmiştir (Roussos & Stout, 1996; Walker & Beretvas, 2001). Batıdaki gelişmelere rağmen ülkemizde DMF alanındaki kaynakların ve çalışmaların yok denecek kadar azlığı dikkati çekmiş ve bu makale DMF alanındaki türkçe kaynak sıkıntısını gidermek dileğiyle kaleme alınmıştır. Makalede sırasıyla DMF tanıtılmış, DMF analizlerinde kullanılan önemli metodlardan bahsedilmiş ve son yıllarda bu alanda olan gelişmeler tartışılmıştır.

DİFERANSİYEL MADDE FONKSİYONU

DMF aynı düzeyde yetenek / bilgi / beceri gösteren bireylerin çeşitli gruplara ayrıldıklarında maddeyi doğru cevaplama olasılıklarının eşit olmaması durumuna denir. Madde yanlılığı ile DMF aynı şeyler değildir. Ancak DMF gösteren maddelerin incelenmesinden ve DMF'nin ortaya çıkma nedenlerinin araştırılmasından sonra, ve DMF'nin testi geliştirirken hedeflenen davranışsal yapıdan kaynaklanmayan başka bir sebepten dolayı ortaya çıktığının tespit edilmesi durumunda madde yanlılığından söz edilir. Bazen DMF davranışsal yapı dizininin iyi

tanımlanamamasından da kaynaklanabilir. Eger gruplar farklı çok boyutlu yetenek dağılımları göstermekteyse ve testteki maddeler bu yetenekler arasındaki farklılıkları görmek için seçilmişse, tek boyutlu puanlama modelleri bu maddeleri DMF olarak teşhis edebilir(Ackerman, 1992). Madde bazında DMF miktarı az olsa da daha sonra bunların birlikte etkileri rahatsız edici bir seviyeye ulaşabilir. Bu durum genellikle *DMF genişlemesi* olarak bilinir. Eger DMF genişlemesi doğrulanırsa *diferansiyel deste fonksiyonununun* (DDF) olduğundan söz edilir. Bu durumlarda, DMF gösterdiği düşünülen maddelerin eş zamanlı olarak madde grupları düzeyinde birlikte analiz edilmesi tavsiye edilir (Roussos, & Stout, 1996).

DMF gösteren maddelerin testteki oranı çok düşükse ve gruplar arasında eşit dağılımısa, DMF'nin ortaya çıkması çok sorun oluşturmaz. Ancak, testteki yanlış maddelerin sayısının artması ve yanlış maddelerin gruplar arasında eşit dağılmaması durumunda, testin grupları eşit olarak ölçüp ölçmediğinden başka bir deyişle geçerliliğinden şüpheye düşülür. Bu nedenle test geliştirme sürecinde DMF analizlerinin yapılması ve DMF gösteren maddelerin bulunarak, DMF'ye sebep olan etkenlerin araştırılması ve bu maddelerin ölçülmesi hedeflenen özellik dışında başka bir faktör yüzünden DMF gösterdiğinin düşünülmesi durumunda maddelerin düzeltme yoluna gidilmesi ve düzeltilmelerinin mümkün olmaması durumunda da testin çıkarılması gerekmektedir. Fakat DMF gösteren maddeleri testten çıkarırken, çıkarılan maddelerin testin yapı geçerliliğini etkilememeye ve çıkarılan maddeler yerine DMF göstermeyen maddeler eklemeye dikkat edilmelidir, aksi takdirde testteki madde sayısı önemli ölçüde azalabileceği için, testin güvenilirlik ve geçerliği olumsuz yönde etkilenebilir.

DMF konusunda daha ileri detaylara girmeden önce, DMF terminolojisinin kısa bir özeti konuyu daha iyi kavrayabilmemiz açısından gereklidir. Boyut (dimension) testi alan kişinin maddeyi doğru olarak yanıtlama şansını etkileyen herhangi bir özelliğidir. Testin ölçmek için hazırlandığı temel boyut (construct) birincil boyut (primary dimension) ve bunun yanında isteyerek ya da istemeyerek ölçülen diğer boyut ise ikincil boyut (secondary dimension) olarak tanımlanır. Eger ikincil boyut teste istenerek yerleştirilmiş ve testin amaçları içinde var ise yardımcı boyut (auxiliary dimension) olarak düşünülür. Eger ikincil boyut test geliştirme aşamasında planlanmamış ve yine de teste kendini göstermişse, gereksiz boyut (nuisance dimension) olarak adlandırılır. DMF gerekli ya da gereksiz boyut tarafından oluşup oluşmadığına bakılarak yararlı ve zararlı olmak üzere iki gruba ayrılır. DMF teste hedeflenen kavramsal yapı yüzünden oluşmuşsa yararlı (benign), hedeflenmeyen bir kavramsal yapı yüzünden oluşmuş ise zararlı (adverse) olarak tanımlanır (Roussos, &

Stout, 1996). Ayrıca, DMF gruplar ve yetenek düzeyleri arasındaki etkileşime göre düzenli ve düzensiz olarak gruplandırılabilir. Düzenli DMF grup üyeliği ile yetenek düzeyleri arasında hiç bir etkileşim olmaması durumunda ortaya çıkarken, düzensiz DMF bu faktörler arasında etkileşimin olması halinde oluşmaktadır. Baska bir deyişle düzenli DMF'de madde karakteristik eğrileri birbirleri ile kesişmez ve bir grup her zaman diğerinden üstündür, oysa düzensiz DMF'de madde karakteristik eğrileri kesişebilir ve gruplardan her biri yetenek düzeylerinin çeşitli kısımlarında bir diğerinden üstün olabilir(Hambleton & Swaminathan, 1985).

DMF ANALİZİNDE KULLANILAN METODLAR

DMF'yi teşhis etmek için önerilen metodların sayısı oldukça fazla olup bu metodları klasik test kuramına dayanan metodlar yada modern test kuramına dayanan metodlar olarak sınıflamak mümkündür. DMF'yi ortaya çıkarmak için kullanılan metodların diğer bir sınıflamasını da parametrik (Ki-kare) yada nonparametrik (SIBTEST, Mantel-Haenszel) olarak yapmak mümkündür. Klasik test teorisine dayanan metodlar, klasik test istatistiklerini (korelasyon yada grupların maddeye verdikleri doğru cevaplar arasındaki farkı) kullanarak yanlış maddeleri bulmaya çalışır. Klasik test istatistikleri örnekleme bağlı olduklarından klasik test teorisine dayanan metodlar da ölçülen özelliğin gruplar arası değişkenliğinden etkilenirler. Ayrıca klasik test istatistiklerinin birden çok örneklemeden elde edilmesi durumunda, farklı gruplar arasında ölçüklerin eşitlenmesi de oldukça zordur (Embretson & Reise, 2000). Öte yandan madde yanıt kuramına dayanan metodlar örnekleme bağlı değildir ancak ölçülen kavramsal yapının tek boyutlu olmaması durumunda bu metodlar yetersiz kalabilir. Makalenin bu bölümde DMF'yi ortaya çıkarmak için sıklıkla kullanılan metodlardan bazıları istatistiksel detaylara girmeden incelenecektir, metodların daha kapsamlı bir tanıtımı için Millsap ve Everson'un (1993) makalesine ve her metodun altında verilen kaynaklara başvurulabilir. Makalenin geri kalan bölümünde sırasıyla, Madde Güçlük İndexine dayanan metodlar, Madde Yanıt Kuramına dayanan metodlar, Ki-kare metodları, Mantel Haenszel (MH) metodu, Loglinear metodları, Lojistic Regrasyon (LR) metodu ve Simultaneous Item Bias (SIBTEST) metodu tartışılacaktır.

Madde Güçlük İndex'ine Dayanan Metodlar

Madde Güçlük İndexine dayanan metoda göre farklı iki grubun madde güçlük indexleri birbiriyle mükemmel korelasyon gösteriyorsa, yada farklı iki grubun madde güçlükleri arasındaki fark eşitse maddenin

yansız olduğundan söz edilebilir. Bu metodlar ya madde güçlük indexini (p) yada madde güçlük indexinden elde edilen başka indexleri kullanır. Bu gruba giren metodlar arasında en çok tanınanları Angoff ve Ford tarafından geliştirilen (1973) Delta indexi ($\hat{\Delta}_g$), ve önce Cardall ile Coffman (1964) daha sonra da Cleary ile Hilton(1968) tarafından geliştirilen varyans analizi teknikleri sayılabilir. Gözlenemeyen kavramsal yapı dağılımının gruplar arasında farklılık göstermesi halinde, bu gruptaki metodlar iki grup arasındaki yetenek farklılıklarını DMF olarak gösterme riskini taşır.

Madde Yanıt Kuramına Dayanan Metodlar

Madde Yanıt Kuramına (MYK) göre DMF iki grubun madde karakteristik eğrilerinin aynı olmaması ya da iki grubun ölçekleri eşitlendikten sonra aynı parametrelere sahip olmaması durumunda görülür (Hambleton & Swaminathan, 1985). Bu yüzden Madde Yanıt Kuramına dayanan indexler madde karakteristik eğrilerinin birbirinden ne kadar çok farklılaştığını ölçmeye yöneliktir. Bu indexlerin hesaplanması madde parametrelerinin her bir grup için ayrı ayrı hesaplanması ve farklı gruplardan elde edilen parametrelerin ölçeklerinin eşitlenmesinden oluşan iki aşamalı bir süreçtir. Madde Yanıt Kuramına dayanan metodlar tek boyutlu kavramsal yapılarda yüksek performans gösterirken, çok boyutlu kavramsal yapılarda yetersiz kalmakta ve çok boyutluluğu DMF gibi gösterebilmektedir (Crocker & Algina, 1986).

Ki-kare Metodları

Scheuneman (1979) ve Camilli (1979) madde yanıt kuramında kullanılan metodlara yakın denilebilecek metodları geliştirmişlerdir. Bu metodların madde yanıt kuramına dayalı metodlara göre uygulaması daha kolaydır. Ki-kare metodları Madde Yanıt Kuramına dayanan metodlardan yanlı maddenin tanımlanması sırasında gözlenemeyen kavramsal yapının yerine onun hatalı bir ölçümü kabul edilen gözlenen test puanının konulması ile ayrılır. Ki-kare metodlarında gözlenen test puanları çeşitli puan dilimlerine bölünür ve her puan dilimindeki gruplar maddeyi doğru cevaplama oranları açısından karşılaştırılır. Doğru cevaplama oranlarının farklı gruplar arasında farklılık göstermesi ise DMF'nin bir göstergesi olarak düşünülür. Ki-kare metodlarının en önemli yanı puan dilimlerini belirlemede kullanılacak yöntemdir, zira puan dilimlerini ayırmak için kullanılan farklı kriterler Camilli ve Scheuneman tarafından gruplar

arasındaki farkı ortaya çıkarmak için öne sürülen istatistiksel testlerin sonuçlarını da etkiler. Ki-kare metodunda gizil kavramsal yapı yerine gözlenen puanın kullanılması, DMF'nin ölçme hatalarından kaynaklanma riskini de beraberinde getirir (Crocker & Algina, 1986).

Mantel-Haenszel İstatistigi

Kontenjan tablolarına dayanan ve ilk olarak Mantel ve Haenszel (1959) tarafından geliştirilen bu metod daha sonra Hollan (1985) ve arkadaşları (Hollan,1985; Holand & Thayer, 1988) tarafından DMF metodu olarak kullanılmaya başlamıştır (Millsap & Everson, 1993). Mantel-Haenszel (MH) İstatistigi, ki-kare metodunun daha genişletilmiş bir hali olup, kullanımının ve anlaşılmasının kolay olması bakımından madde düzeyindeki DMF'İ teşhis etmekte oldukça yaygın olarak tercih edilen bir methodur.

MH metoduyla referans ve odak grubu adı verilen, iki grubun başarı düzeyleri her bir madde için tek tek karşılaştırılır. DMF'in oluşmasıyla dezavantajlı konuma düşeceği düşünülen grup odak grubu ve DMF'i teşhis etmek için kullanılacak karşılaştırma grubu ise referans grubu olarak tanımlanır. Genellikle gruplar arasında karşılaştırmayı sağlamak için toplam test puanı baz alınır. MH metodunda kontenjan tablolarına yerleştirilen veriler üzerinden MH χ^2 istatistigi hesaplanarak ve χ^2 dağılımını referans olarak DMF hipotezinin manidarlığını test etmek mümkündür (bk., Holand & Thayer, 1988). Bu metod hem null hipotezinin test edilmesine izin verdiği hem de DMF'in büyüklüğünü gösteren bir index sağladığı için sıklıkla kullanılmaktadır (Millsap & Everson, 1993).

Loglinear Modelleri

Kontenjan tablolarına dayanan loglinear modelleri geleneksel ki-kare metodunun daha genelleştirilmiş bir halidir (Alderman & Holland, 1981, Mellenbergh, 1982). Bu metodla üç boyutlu kontenjan tabloları (puan dilimleri / gruplar / madde yanıtları) kullanılarak ilgili loglinear modelin parametreleri kestirilir. Modelin veriye uygun olup olmadığı örneklemeden elde edilen frekanslardan beklenen frekansların tahmin edilmesi ve G^2 istatistiginin (bk., Fienberg, 1980) hesaplanması ile olur.

Loglinear modelleri araştırmacılar arasında diğer modellere göre daha esnek olması, interval maddelerine uygulanabilmesi, aynı anda birden çok grup ve madde üzerinde DMF analizine imkan vermesi bakımından

yaygın olarak kullanılır. Maddelerin önemli bir kısmının yanlı olması durumunda grupları eşitlemekte kullanılan toplam test puanının da yanlı olması söz konusu olacağından, araştırmacılar yanlı maddelerin sırayla test çıkarılmasını ve daha sonra analizin yapılmasını önermektedirler. Loglineer modeller tek boyutlu verilere uygun olup, çok boyutlu ve kompleks verilerde altta yatan kavramsal yapıyı yeterince temsil edemediği için bu durumlarda kullanımına dikkat edilmelidir (Millsap & Everson, 1993).

Lojistik regrasyon Modelleri

Swaminathan ve Rogers (1990) lojistik regrasyona dayanan bir DMF metodu geliştirdiler. Lojistik regrasyon modelinde bireyin maddeye doğru cevap verebilmesinin koşullu olasılığı maximum likelihood metoduyla bulunur (daha fazla bilgi için bk., Bock, 1975; Swaminathan & Rogers, 1990).

Loglineer yöntemlerine pek çok yönden benzeyen, bu metod birden çok gruba ve interval maddelerine uygulanabilmekte ve de düzenli ve düzensiz DMF'i teşhis edebilmektedir (Agesti, 1990; Millsap & Everson, 1993). Lojistik regrasyon modelleri ile DMF'nin oluşumunda katkısı olduğu düşünülen diğer ilgili faktörlerin de (test anxiety gibi) DMF üzerindeki etkisinin araştırılması mümkündür (Miller & Spray, 1993). Ayrıca, lojistik modeller regrasyon katsayısını sağladıkları için bunların grafiklerinin çizilmesi durumunda puan dilimleri üzerinde DMF'nin sorunlu olduğu bölgeler bulunabilir (Millsap & Everson, 1993). Madde yanıt modellerinin çok parametreliliği bir yapı göstermesi durumunda loglineer modelleri gibi, lojistik modeller de sorunlar yasar ve DMF olmadığı halde varmış gibi gösterebilirler.

SIBTEST Metodu

Shealy ve Stout (1993) güçlü bir DMF matematik modeli (MMD; Shealy, & Stout, 1993) geliştirdiler. Shealy ve Stout (1993) modellerinde DMF'in oluşma nedenlerini çok boyutluluk yaklaşımı ile açıklamışlardır. DMF'nin her zaman çok boyutlulukla ilgili bir konu olduğunu ileri süren Shealy ve Stout, DMF'nin iki koşulun (a) madde ölçülmek istenen ana kavramsal yapıya θ ve aynı zamanda yardımcı kavramsal yapıya η duyarlı, ve (b) η 'nin sabit bir θ değeri için şartlı dağılımının ($\eta|\theta$) farklılaşması halinde görüldüğünü belirtmişlerdir. Shealy ve Stout (1993)'ün çok boyutluluğa dayanan DMF modelinin çoğunlukla uygulanan ve bireysel madde analizine dönük DMF metodlarından bir çok

yönden daha avantajlı olduğu söylenir: (a) testin sonuçlarını etkileyen temel kavramsal yapıların incelenmesi grupları eşitlemek için kullanılacak kriterin daha iyi belirlenmesine yol açar; (b) test maddelerinin kavramsal yapı yönünden benzerlik gösteren gruplar / desteler halinde test edilmesi DMF'nin manidarlığını artırır; (c) testin çok boyutlu kavramsal yapısını anlamak, DMF'ye sebep olan nedenleri daha iyi anlamamıza yol açar.

SIBTEST (Shealy, 1989; Shealy & Stout, 1991), Shealy ve Stout'un çok boyutluluğa dayanan DMF matematiksel modellerinin uygulanmasını sağlayan bilgisayar programı olup birden çok maddenin es zamanlı DMF analizini mümkün kılmaktadır. SIBTEST hem düzenli hem de düzensiz DMF'i teşhis edebilmekte, madde bazında ya da madde grupları bazında DMF'i teşhis edebilmekte, DMF'nin büyüklüğünü ölçebilmek için bir indek ve null hipotezini test etmek içinde bir istatistik sağlamaktadır (Shealy & Stout, 1993). SIBTEST, önce sadece teste hedeflenen davranışı ölçen maddelerden oluşan geçerli bir alttest oluşturur, ve bu alttest üzerinden her bir birey için toplam puan (Z) hesaplanarak, bireyler puan dilimlerine göre ayrılır. Daha sonra incelenen puan düzeyindeki (Y) gruplar için düzeltilmiş ortalamalar (adjusted means) hesaplanır ve gruplar arasındaki farkların ağırlıklı ortalaması bir özet istatistiği olarak bulunurak, null hipotezinin manidarlığı z testi kullanılarak test edilir (daha fazla bilgi için bk., Shealy, Stout, & Rossi, 1991; Shealy, & Stout, 1993).

SONUÇLAR

DMF metodlarının geçerlik sürecindeki rolü uç madde altında özetlenebilir: (a) belirli bir testi alan farklı grupların test sonuçlarının farklı faktörler tarafından etkilenip etkilenmediğini bulmak; (b) farklı grupların test sonuçlarının aynı faktörler tarafından etkilendiğinin bulunması durumunda bu faktörlerin testi planlarken hedeflenen ve testi etkilemesi düşünülen faktörler olup olmadığına karar vermek; ve (c) testi hazırlama sürecinde planlanmayan ve testle doğrudan ilgili olmayan faktörlerin bulunması halinde bunların bir gruba haksız avantaj sağlayıp sağlamadığını bulmak.

DMF gösteren maddeleri ortaya çıkarmak için çeşitli metodlar kullanılmaktadır. Madde Yanıt Kuramına dayanmayan ve geleneksel Ki-kare metodlarının daha genişletilmiş halleri olan Mantel-Haenszel, Lojistik Regrasyon Modelleri, Loglineer Modelleri esneklikleri ve kullanım kolaylıkları yönünden sıklıkla tercih edilen metodlar arasındadır (Crocker & Algina, 1986). SIBTEST de son yıllarda kullanımı yaygınlaşan ve bu alanda gelecek vaad eden bir metod olarak göze çarpmaktadır (Millsap & Everson, 1993). 1960'ların başından beri DMF çalışmalarının

sürdürülmesine ve DMF metodlarındaki çeşitlenmesine rağmen, halen DMF'ye sebep olan faktörler kesin olarak bilinmiyor (Messick, 1989; Walker & Beretvas, 2001). Pratikte, DMF'yi tanımlamakta kullanılan istatistiksel ve betimsel metodlar birbiri ile iç içe nadiren yürütülmekte (Roussos, & Stout, 1996) ve bu yüzden DMF çalışmaları bir çok açılardan yetersiz kalmaktadır. Son zamanlarda DMF çalışmalarında görünen bu başarısızlığa dayanarak, araştırmacılar her hangi bir DMF çalışması yapmadan önce testin altında yatan gözlenemeyen kavramsal yapı dağılımını (latent trait distribution) tanımanın gerekli olduğunu vurgulamışlardır (Messick, 1989). Bu uyarılar doğrultusunda son yıllarda hem betimsel hem de istatistiksel metodları kullanarak, DMF'e sebep olan etkenleri ortaya çıkarmaya çalışan başarılı çalışmalar (bk., Gierl et al., 2001; Ryan & Fan, 1996; Ozbek, 2004; Walker & Beretvas, 2001) ümit vericidir.

KAYNAKLAR

- Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, 15, 1, 13-24.
- Agresti, A. (1990). *Categorical Data Analysis*. New York : Wiley.
- Alderman, D.L., & Holland, P.W. (1981). *Item performance across native language groups on the Test of English as a Foreign Language* (ETS Research Rep. No. 81-16). Princeton NJ: Educational Testing Service.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- American Psychological Association, American Educational Research Association & National Council on Measurement in Education. (1966). *Standards for educational and psychological test and manuals*. Washington, DC: American Psychological Association.
- American Psychological Association, American Educational Research Association & National Council on Measurement in Education. (1974). *Standards for educational and psychological test and manuals*. Washington, DC: American Psychological Association.

- Angoff, W. H., & Ford, S. F. (1973). Item race interaction on a test of scholastic aptitude. *Journal of Educational Measurement*, 10, 95-105.
- Bock, R. D. (1975). *Multivariate Statistical Methods*. New York: McGraw-Hill.
- Camilli, G. (1979). *A critique of the chi square method for assessing item bias*. Unpublished manuscript, University of Colorado, Laboratory of Educational Research, Boulder.
- Cardall, C., & Coffman, W. E. (1964). *A method for comparing the performance of different groups on the items in a test* (College Entrance Examination Board Research and Development Rep. 64-5 No. 9; ETS Research Bulletin 64-61). Princeton NJ: Educational Testing Service.
- Cleary, T. A. (1968). Test bias: prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement*, 5, 115-124.
- Cleary, T. A., & Hilton, T. L. (1968). *An investigation of item bias*. *Educational and Psychological Measurement*, 28, 61-75.
- Crocker, L., & Algina, J. (1986). *Introduction to Classical Modern Test Theory*. Rinehard and Winston Inc. United States.
- Embretson, S. E., & Reise, S. P. (2000). *Item Reponse Theory for Psychologists*. Lawrence Erlbaum Associates Inc. Mahway, New Jersey.
- Fienberg, S. E. (1980). *The analysis of cross-classified categorical data*. Cambridge MA: MIT Press.
- Gierl, M. J., Bisanz, J., Bisanz, G., Boughton, K.A., & Khaliq, S. N. (2003). Identifying content and cognitive skills that produce gender differences in mathematics: a demonstration of the multidimensionality-based DIF analysis. *Journal of Educational Measurement*, 40, 281-306.
- Hambleton, K. R., & Swaminathan, H. (1985). *Item Response Theory Principles and Application*. Nijhoff Publishing.
- Hollan, P. W. (1985). On the study of differential item performance without IRT. *Proceedings of the 27th Annual Conference of the Military Testing Association* (Vol. 1; pp.282-287). San Diego CA: Navy Personnel Research and Development Center.
- 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: Erlbaum.
- Mantel, N., & Haenszel, W. (1959). Statistical aspect of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719-748. Mellenbergh, 1982
- Messick, S. (1989). Validity. Linn, Robert L. (ED). *Educational Measurement (3rd ed.)*. The American council on education / Macmillan series on higher education. (pp. 13-103).

- Miller, T. R., & Spray, J. A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of Educational Measurement*, 30, 107-122.
- Millsap, R. E., & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing measurement bias. *Applied Psychological Measurement*, 17, 297-334.
- Ozbek, O. Y. (2004). *An empirical investigation of the validity of Secondary School Institutions Student Selection and Placement Test (SSISSPT) in Turkey*. Unpublished doctoral dissertation, Department of Education Psychology, University of Illinois, Urbana-Champaign.
- Roussos, L. A., & Stout, W. F. (1996). A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement*, 20 (4), 355-371.
- Ryan, K. E., & Fan, M. (1996). Examining gender DIF on a multiple-choice test of mathematics: a confirmatory approach. *Educational Measurement: Issues & Practice*, 15 (4), 15-20.
- Scheuneman, J. (1979). A new method for assessing bias in test items. *Journal of Educational Measurement*, 16, 143-152.
- Shealy, R. T. (1989). *An item response theory-based statistical procedure for detecting concurrent internal bias in ability tests*. Unpublished doctoral dissertation, Department of Statistics, University of Illinois, Urbana-Champaign.
- Shealy, R. T., & Stout, W. F. (1991). *A procedure to detect test bias present simultaneously in several items* (Tech. Rep. 1991 #3). Washington DC: Office of Naval Research.
- Shealy, R. T., & Stout W.F. (1993). An item response theory model for test bias. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (197-239). Hillsdale NJ: Erlbaum.
- Swaminathan, H. & Rogers, J. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27, 361-370.
- Walker, C. M., & Beretvas, S. N. (2001). An empirical investigation demonstrating the multidimensional DIF paradigm: A cognitive explanation for DIF. *Journal of Educational Measurement*, 38 (2), 147-163.

İLETİSİM ADRESİ

Dr. Özlem Yeşim ÖZBEK
Gaziosmanpaşa Üniversitesi
Eğitim Fakültesi
e-mail:ozlem.ozbek@gmail.com

RASYONEL SAYILARLA BÖLME İŞLEMİNİ İLKÖĞRETİM ÖĞRENCİLERİNİN ALGILAYIŞLARI

Yard. Doç. Dr Soner DURMUŞ

Abant İzzet Baysal Üniversitesi Eğitim Fakültesi
İlköğretim Bölümü

ÖZET

İlköğretim öğrencilerinin rasyonel sayılarla bölme işleminde gerek kavramsal ve gerekse işlemsel olarak zorluklar yaşadıkları bir çok araştırma ile ortaya konmuştur. Öğrencilerin rasyonel sayılardaki bölme kavramını işlemsel olarak yapmalarına rağmen kavramsal olarak anlamlandırıp anlamlandırmadıkları açık değildir. Bu çalışma ile öğrencilerin: i) rasyonel sayılarla bölme işlemi yaparken kullandıkları yöntemler ve kullanma gerekçelerini belirlemek, ii) kavram yanlışlarını ortaya çıkarmak ve iii) kullandıkları yaklaşımların sınıflara göre farklılık gösterip göstermediğini saptamak amaçlanmıştır. Çalışmanın sonuçları, literatürdeki bulgular doğrultusunda eleştirel bir yaklaşımla tartışılmıştır.

Anahtar Kelimeler: Rasyonel sayılar, bölme

ABSTRACT

PRIMARY SCHOOL STUDENTS' CONCEPTION ABOUT RATIONAL NUMBER DIVISIONS.

Research findings confirm that primary school students have learning difficulties on division operations with rational numbers. It is difficult to derive a conclusion about students' conceptual understanding of rational number division based on formal operations with rational numbers. This study aims i) to determine what kind of strategies that students use while dealing with fraction divisions, ii) to reveal (if any) misconceptions and iii) to determine strategies that students use taking the grade levels into consideration. Findings and educational implications will be discussed in detail.

Key Words: Rational numbers, division