

Madde Parametrelerinin Değişmezliğinin Çeşitli Boyutluluk Özelliği Gösteren Yapılarda Madde Tepki Kuramına Göre İncelenmesi*

Önder SÜNBÜL** Adnan ERKUŞ***

Özet: Bu çalışmada, çeşitli boyutluluk özelliklerine sahip yapılardaki madde parametrelerinin değişmezliği MTK çerçevesinde incelenmiştir. Bu amaç doğrultusunda, tamamlayıcı çok boyutlu madde tepki kuramı modeline dayalı iki boyutlu veri üretimi yapılmıştır. Veri üretiminde, birinci boyut maddelerinin sayısı (24) ve a_1 parametre ortalaması (1.00), bütün yapılarda aynı olacak şekilde korunmuştur. Değişimlemeler, ikinci boyutta yer alan maddeler üzerinde ve boyutlar arası korelasyonlarda yapılmıştır. İkinci boyut maddelerinin sayısı 4, 8, 12, 16, 20, 24; a_2 parametre ortalaması 0.25, 0.50, 0.75, 1.00 ve boyutlar arası korelasyon 0.00, 0.30, 0.60, 0.80 olacak şekilde değişimlenmiştir. Değişimleme ölçütlerinin tamamının çaprazlanması sonucu, 96 deneysel hücre elde edilmiştir ve her bir hücre için 90.000 bireyin madde tepkisini kapsayan veri üretimi yapılmıştır. Böylelikle 96 adet evren elde edilmiştir. Parametre değişmezliğinin örneklem büyüklüğünden nasıl etkilendiğini belirlemek amacıyla, her evrenden 500'lük 100 adet, 1000'lik 100 adet ve 3000'lik 100 adet örneklem, basit seçkisiz örnekleme yöntemine dayalı olarak çekilmiştir. Böylelikle 28 800(96 x 3 x 100) örneklem elde edilmiştir. Elde edilen örneklem MTK çerçevesinde analiz edilmiştir. Elde edilen analiz sonuçlarının değerlendirilmesinde, parametre değişmezliğinin göstergesi olarak standart hata ortalamalarından yararlanılmıştır. Veriler boyutlara göre gruplanmadan analiz edilse de, standart hata ortalamaları her boyut için ayrı ayrı hesaplanarak, her iki boyutun parametre değişmezliğinden nasıl etkilendiği belirlenmiştir. MTK için elde edilen sonuçlar incelendiğinde, hem a hem de b parametresi için parametre değişmezliğinin büyük oranda sağlandığı görülmekle beraber bazı yapılarda parametre değişmezliğinin sağlanamadığı görülmüştür. Bu yapılar; ikinci boyutun birinci boyuta kuvvetçe eşdeğer (madde sayısı ve a parametre ortalaması) ve boyutların birbirine dik konumlandığı deneysel hücrelerdir. Belirtilen deneysel hücreler için hesaplanan standart hata ortalamalarının, özellikle b parametresi için, kabul edilemeyecek kadar büyük olduğu görülmüştür. Belirtilen özel durumlar haricindeki deneysel hücrelerde; MTK tekboyutluluk sayılısının ihlal edilmesine rağmen, değişmez madde parametreleri üretilebilmiştir. Literatürde de belirtildiği üzere, MTK'nin çokboyutluluğa karşı oldukça dayanıklı olduğu görülmüştür.

Anahtar Kelimeler: Madde Tepki Kuramı, madde parametrelerinin değişmezliği, boyutluluk

ABSTRACT Examining item parameter invariance for several dimensionality types by using unidimensional item response theory. In this study, item parameter invariance in constructs which have characteristics of various dimensionality was examined within the framework of Unidimensional Item Response Theory (UIRT). For this purpose, two-dimensional data simulation was realized depending on compensatory MIRT model. For the simulation, the number of the items in the first dimension (24) and the mean of a_1 parameters (1.00) were kept constant in all constructs. Manipulations were applied to the items in the second dimension and correlation between dimensions. The numbers of the items in the second dimension were set to 4, 8, 12, 16, 20, 24; mean a_2 parameters were set to 0.25, 0.50, 0.75, 1.00, and correlations between dimensions were set to 0.00, 0.30, 0.60, 0.80. Simulation criteria were fully crossed and 96 experimental cells were obtained. Dichotomous responses of 90 000 subjects were simulated for each cell. By this way, 96 populations were established. In order to determine how parameter invariance is affected by sample size, 100 samples for 500 sample size, 100 samples for 1000 sample size and 100 samples for 3000 sample size were drawn from populations by using simple random sampling. Thus, 28800 samples were revealed. These samples were analyzed by using UIRT. In order to analyze the analysis results, average standard errors were used as the indicator of parameter invariance. Despite the fact that all datasets are treated as unidimensional for IRT analysis, average standard errors were computed separately for each dimension and how both dimensions have been affected by parameter invariance was determined in this manner. When the results for IRT were examined, parameter invariance were provided to a great extent for both a and b parameters but it cannot be provided for some constructs. Those constructs are empirical cells in which the first dimension is equivalent in terms of strength to the second dimension and in which the dimensions are orthogonal. Average standard errors which were computed for the particular empirical cells seem bigger than the acceptable level for especially b parameter. Empirical cells can produce invariance item parameters in spite of the violation of IRT's unidimensionality assumption. As shown in literature, UIRT is quite robust against multidimensionality.

Keywords: Item Response Theory, item parameter invariance, dimensionality

* Bu çalışma, Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı'nda Doktora Tezi olan çalışmanın bir bölümünden oluşmaktadır.

** Yrd. Doç. Dr., Mersin Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı. ondersunbul@gmail.com

*** Prof. Dr., Emekli Öğretim Üyesi adnanerkuspsi@gmail.com

GİRİŞ

Fen bilimlerinin ve sosyal bilimlerin bilim olarak kabul edilmesinde ve gelişmesinde ölçmenin yeri yadsınamaz. Bilim dalları, o bilim dalında yapılan ölçme işlemlerinin gelişmişliği ölçüsünde ilerleyebilmektedir (Baykul, 2000). Bilimin; gözlenebilirlik, ölçülebilirlik, iletilebilirlik, tekrarlanabilirlik, sağdanabilirlik gibi ölçütleri göz önüne alındığında (Karakaş, 1997), ölçmenin bu ölçütlerin merkezinde yer aldığı iddia edilebilir. Eğer ölçme olmazsa diğer ölçütler karşılanamaz, ölçmesiz bilim olamaz.

Fen bilimlerinde ölçme işleminin yapılması, sosyal bilimlere göre çok daha kolaydır; ilgilenilen değişkenler, çoğunlukla kararlı özelliklere sahip fiziksel değişkenler olduğundan, doğrudan ölçülebilmektedir. Sosyal bilimlerde ise daha çok psikolojik değişkenlerle ilgilenilmekte ve dolaylı ölçüm yapılmaktadır. Bireylerin psikolojik yapılarındaki değişkenlikten ve doğrudan ölçüm yapılamadığından dolayı, ölçme işlemi oldukça güçlüklerle karşı karşıyadır. Bireylerin sahip olduğu örtük özelliği en uygun şekilde ortaya koyabilmek için çeşitli kuramlar geliştirilmiştir. İlk geliştirilen kuram Klasik Test Kuramı'dır.

Klasik Test Kuramı (KTK)

Bireyleri tek bir psikolojik süreklilik üzerinde, ilgilenilen değişken açısından ölçekleyebilmeyi amaçlayan kuramlardan biri olan Klasik Test Kuramı, bireyin gözlenen özelliğinin içinde hem gerçek puan hem de hata puanı bulunduğunu kabul eder.

Bu kuramda, ölçülmek istenen özelliğe ilişkin puan bireyin ölçekten aldığı ham puana karşılık gelirken, bireyin yetenek ölçüsü, ham puanların standartlaştırılmış değerleri olmaktadır (Doğan, 2002). Her kuramın olduğu gibi Klasik Test Kuramı'nın da sayıltıları vardır. Klasik Test Kuramı'na ilişkin sayıltılar şu şekildedir: Hata puanlarının beklenen değeri, başka bir deyişle ortalaması sıfırdır. Gerçek puanlarla, hata puanları arasındaki korelasyon sıfırdır. İki paralel ölçmeye ilişkin hata puanları arasındaki korelasyon sıfırdır. İki farklı ölçmeye ait gerçek puanlar ve hata puanları arasındaki korelasyon sıfırdır (Gulliksen, 1950; Magnusson, 1968; Baykul, 2000). Kaynaklardaki sayıltılar yukarıda belirtildiği şekilde olmasına rağmen, Klasik Test Kuramı'nda ölçülmek istenen değişkenin tekboyutlu olması gibi bir örtük sayıltısı ve bireylere uygulanan testlerin paralel olması gibi başka sayıltıları da vardır. Klasik Test Kuramı uygulamalarının diğer ölçme kuramlarına göre daha kolay olmasından ve karışık olmayan az işlem yükü gerektirmesinden dolayı, yıllarca yaygın olarak kullanılmıştır ve hala kullanılmaktadır. Bu kuram, eğitimin ve psikolojinin bilim olarak kabul edilmesine değerli katkılarda bulunmuştur. Bununla beraber gerçeklik payı olan ve olmayan birçok eleştiriye uğramıştır. Yapılan eleştirilerin temelinde, Klasik Test Kuramı'nın çok sayıda sayıltısı olduğu; bu sayıltıların her zaman karşılanmadığı; Klasik Test Kuramı'ndan elde edilen madde ve test istatistiklerinin testin uygulandığı gruba ve teste bağımlı olması ve bütün yetenek ranjı için bir tek hata kestirimi yapılması gibi sınırlılıklar yatmaktadır (Lord ve Novick, 1968; Lord, 1980; Hambleton ve Swaminathan, 1985; Crocker ve Algina, 1986; Embretson ve Reise, 2000). Klasik Test Kuramı'nın temel sınırlılıklarını giderebileceği iddiasıyla, Örtük Özellikler Kuramı olarak da bilinen Madde Tepki Kuramı geliştirilmiştir.

Madde Tepki Kuramı (MTK)

Lord (1953), bireylerin Klasik Test Kuramı'nda kullanılan gözlenen puanlarının ve gerçek puanlarının, yetenek puanlarıyla aynı anlama gelmediğini belirtmiş, gözlenen puanlar ve gerçek puanlar teste bağımlı iken, yetenek puanlarının testten bağımsız olduğu ve bu nedenle de daha temel olduğunu vurgulamıştır (akt: Hambleton ve Jones, 1993). Bireyler zor testlerde daha düşük gerçek puana sahip olurken, kolay testlerde daha yüksek gerçek puana sahip olabilmektedir; ancak bireylerin yetenekleri sabit kalmaktadır (Hambleton ve Jones, 1993). Klasik Test Kuramı'nda gözlenen bu eksikliği giderebilmek amacıyla ölçme sürecine farklı bir bakış açısından yaklaşılmış ve Madde Tepki Kuramı geliştirilmiştir. Klasik Test Kuramı gözlenen puandan gerçek puan hakkında çıkarımlar yapmaya çalışırken, MTK bireyin yeteneğinden, bireyin madde üzerinde göstereceği performansın olasılığı üzerine yoğunlaşır.

MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE İNCELENMESİ

MTK'nin üzerine kurulu olduğu iki temel postula Hamleton ve Swaminathan (1985) tarafından bireyin bir test maddesindeki performansı; özellik, örtük özellik veya yetenek gibi faktörlerce yordanaabileceği ve bireyin madde performansı ile madde performansının altında yatan özellik kümesi arasındaki ilişki, madde karakteristik eğrisi (item characteristic curve) (MKE) veya fonksiyonu adı verilen monotonik artan bir fonksiyonla ifade edilebileceği şeklinde belirtilmiştir. Temel postulalar incelendiğinde, MTK'nin testten çok maddeye odaklandığını söylemek mümkündür. Bir başka deyişle MTK'nin özünde MKE yatmaktadır.

Madde Karakteristik Eğrisi

Madde puanının yetenek üzerindeki regresyonu, madde karakteristik eğrisi/fonksiyonu olarak adlandırılır. MKE, bireyin maddede başarılı olma olasılığını, testle ölçülen yeteneğine bağlayan matematiksel bir fonksiyondur. Bir başka deyişle, madde puanlarının test tarafından ölçülen yetenek üzerindeki doğrusal olmayan regresyonudur (Crocker ve Algina, 1986).

Belirli bir yetenek ranjındaki bireylerin, maddeyi doğru cevaplama olasılığını monotonik artan bir fonksiyon şeklinde temsil eden MKE'nin yatay eksenini bireylerin yetenek düzeylerini, dikey eksenini ise maddenin doğru cevaplanma olasılığını göstermektedir. Bireylerin sahip olduğu yetenek arttıkça maddenin doğru cevaplanma olasılığı da monotonik olarak artmaktadır. MKE asimptotik bir fonksiyondur. Madde karakteristik eğrileri, fonksiyonunda yer alan parametrelere bağlı olarak farklılaşmaktadır. Bu farklılaşma ise, çeşitli MTK modellerini doğurur.

Madde Tepki Kuramının Sayıtları

I. Tek Boyutluluk (Unidimensionality)

Tek boyutluluk, bireyin madde üzerindeki performansını etkileyen sadece ve sadece bir yetenek türü olması durumudur (Lord ve Novick, 1968; Lord, 1980; Hambleton ve Swaminathan, 1985; Crocker ve Algina, 1986; Embretson ve Reise, 2000). Bir başka deyişle, testi, daha geniş anlamda ölçme aracını oluşturan maddeleri bir arada tutan bir tek ortak özellik olmalıdır. Tekboyutlu yapılarda maddeler bütün popülasyon için istatistiksel olarak bağımlıdır ve paylaştıkları ortak bir varyans olduğu söylenebilir (Crocker ve Algina, 1986). Testi oluşturan maddelerin doğru cevap olasılıklarının belirlenmesinde iki veya daha fazla yetenek boyutu etkili olmamalıdır. Maddenin cevaplanma olasılığında eğer birden fazla özellik etkinse, tek boyutluluk sayıltısı karşılanamayacaktır. Tekboyutluluk sayıltısı MTK modellerini büyük ölçüde sınırlar, çünkü saf tekboyutluluğun pratik durumlarda sağlanması olanaksızdır. Bireylerin maddelere tepki verirken tek bir boyut doğrultusunda hareket etmeleri beklenemez; bununla beraber, ölçülmek istenen değişken kompleks bir yapıya sahip olabilir. Böyle bir yapı tekboyutluluk sayıltısına dayanan MTK modelleri ile ölçülemez. Bu gibi nedenlerden dolayı, tekboyutlu MTK modelleri, bu sayıltının pratikte karşılanamamasına ilişkin yoğun eleştiri almıştır (Embretson ve Reise, 2000; Kirisci, Hsu ve Yu, 2001).

Tekboyutluluk sayıltısının karşılanıp karşılanmadığını değerlendirmek için genellikle, elde edilmiş olan ham verilere faktör analizi yapılmakta ve baskın bir faktör arayışı içine girilmektedir. Eğer baskın bir faktör bulunabilirse tekboyutluluk sayıltısının karşılandığı iddia edilerek, madde ve yetenek parametrelerinin kestirimleri yapılmaktadır.

II. Yerel Bağımsızlık (Local Dependency)

MTK'nin ikinci sayıltısı olan yerel bağımsızlık, test performansını etkileyen yetenek *sabit tutulduğunda*, bireylerin maddelere vereceği tepkilerin istatistiksel olarak bağımsız veya ilişkisiz olmasıdır (Lord ve Novick, 1968; Lord, 1980; Hambleton ve Swaminathan, 1985; Reckase, 2009). Tekboyutluluk sayıltısının sağlanabilmesi için testte yer alan maddelerin ilişkili olması gerekirken, bu sayıltıda ise

maddelerin belirli bir yetenek düzeyi için bağımsız olması gerekmektedir. Bu durum ilk bakışta bir çelişki yaratıyor gibi görünmesine rağmen, gerçekte öyle değildir. Tekboyutluluk sayılısında maddelerin gösterdiği ilişkiler incelenirken, herhangi bir yetenek kısıtlaması veya koşulu yoktur. Yerel bağımsızlık sayılısında ise, maddeler arasındaki ilişki veya bağımsızlık belirli bir yetenek koşulu altında incelenmektedir. Bu durum, olasılıkta iki olayın bağımsız olabilmesi için, bu olayların birlikte olma olasılığının, her bir olayın olasılıkları çarpımına eşit olması gerektiği göz önüne alınarak, daha formel bir biçimde aşağıdaki gibi ifade edilebilir (Hambleton ve Swaminathan, 1985):

$$\begin{aligned} Prob(U_1, U_2, U_3, \dots, U_n | \theta) &= P(U_1 | \theta) \cdot P(U_2 | \theta) \cdot \dots \cdot P(U_n | \theta) \\ &= \prod_{i=1}^n P(U_i | \theta) \end{aligned}$$

Yerel bağımsızlık sayılısının karşılanabilmesi için tekboyutluluk sayılısının mutlaka karşılanmış olması gerekir. Eğer maddeye verilecek olan tepki birden fazla yetenekten kaynaklanıyorsa, maddeleri bağımsız yapmak için süreçte etkili olan yeteneklerden birini sabitlemek yeterli olmayacaktır ve sabitlenmemiş olan diğer yetenek, maddeler arasında bağımlılık veya ilişki oluşturacaktır. Yerel bağımsızlığı sağlanamaması, özellikle madde ve yetenek parametrelerinin kestirilmesinde kullanılan maksimum olabilirlik (maximum likelihood) fonksiyonunun temelini oluşturan, olabilirlik (likelihood) fonksiyonlarından hesaplanacak olan olasılıkları tehlikeye sokacaktır. Çünkü olabilirlik fonksiyonları, belirli bir yetenek koşulu altında, maddelere bağımsızlığı gibi davranarak olasılık sonuçları üretmektedir.

Yerel bağımsızlık sayılısının karşılanıp karşılanmadığını değerlendirmek için son yıllarda durumsal kovaryans hesaplamalarından yola çıkan çeşitli indeksler geliştirilmiştir. Bununla beraber; genel olarak tekboyutluluk sayılısının karşılanması, yerel bağımsızlık sayılısının da yerine getirilmesi için yeterli görülmektedir. Lord (1980), yerel bağımsızlığı ek bir sayılı olarak görmemekte, tekboyutluluğun bir getirisi olarak değerlendirmektedir. Tekboyutluluğun sağlanamadığı durumlarda, bir tek yetenek sabitlemesine bağlı yerel bağımsızlıktan söz edilemez. Bu nedenle MTK modelleri için tekboyutluluk hayati önem taşımaktadır.

Boyutluluk (Dimensionality)

Özellikle yapı geçerliğine ilişkin çalışmalarda sıkça karşılaşılan boyutluluk kavramı, ölçme ve sonrasındaki değerlendirme süreçleri için çok büyük öneme sahiptir. KTK açısından bakıldığında bireyler, ölçme aracından aldıkları ham puanlara veya standart puanlara göre ölçülmek istenen değişken bakımından bir psikolojik süreklilik üzerinde konumlandırılmaya çalışılmaktadır. Her ne kadar KTK'nin sayılıları arasında tekboyutluluk yer almasa da, bireylere uygulanan ölçme aracındaki maddelerden elde edilen madde puanlarının skaler olarak toplanabilmesi için tekboyutluluk bir gerekliliktir. MTK açısından bakıldığında ise tekboyutluluk, kuramın en temel sayılısı olarak karşımıza çıkmaktadır. Dolayısıyla her iki kuram için de boyutluluk araştırmalarının yapılması kaçınılmazdır. Bir nesne kümesine ait boyutluluk; “nesnelere arasındaki ayrı ve ilgi çekici değişkenlik kaynağı sayısı olarak tanımlanabilir” (Jacoby, 1991). Bir madde kümesinin boyut olarak ele alınabilmesi için, bu kümenin bireyler arasında değişkenlik kaynağı oluşturacak şekilde bir niteliğe sahip olması ve maddelerin ortak bir karakteristik özelliğe sahip olması gerektiği söylenebilir. Boyut, nesne kümelerinin karakteristik özelliklerini, analitik amaçlar için daha karşılaştırılabilir hale getirmek için sadeleştirir.

Steinberg, Thissen ve Wainer (2000), boyutluluk kaynaklarını, gruplar arası (between-groups) ve grup içi (within-group) olarak iki ana kategori altında toplamıştır. Buna göre gruplar arası boyutluluk, maddelerin gruplar arasında farklı karakteristik özelliklere sahip olmasından kaynaklanmaktadır. Başka bir deyişle maddenin farklı gruplar için farklı anlam taşıması, maddenin farklı gruplarda farklı şeyi ölçtüğü anlamına gelir. Bu durum MTK literatüründe DIF (differential item functioning) olarak adlandırılır. Grup içi boyutluluk, maddelerin her grup için aynı anlamı taşıması, dolayısıyla aynı şeyi ölçmesine rağmen, maddeye tepki verilebilmesi için birden fazla boyutun gerektiği durumlarda ortaya çıkmaktadır. Bununla

MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE İNCELENMESİ

beraber madde parametrelerinin değişmezliği yabancı literatürün büyük bir kısmında DIF olarak ele alınmaktadır (Rupp ve Zumbo, 2006).

Faktör Analizi ve Yapı Türleri

Faktör analizi başta sosyal bilimlerde olmak üzere, pek çok alanda ilişkili değişkenleri bir araya getirerek, az sayıda yeni ilişkisiz değişken bulmayı amaçlayan çok değişkenli bir analiz türüdür (Tatlıdil, 1992). Faktör analizinin temelleri Spearman'ın 1904 yılında zeka üzerine yaptığı çalışmalara dayanmaktadır (Bartholomew, 2007). Ölçme araçları açısından bakıldığında, faktör, ortak özelliğe sahip maddeleri bir arada tutan bir bağ gibi düşünülebilir. Faktör, tıpkı gerçek puan gibi doğrudan gözlenemeyen örtük bir değişkendir. Maddelerin faktörle göstermiş olduğu ilişki de faktör yükü olarak tanımlanabilir (Crocker ve Algina, 1986).

Yapı geçerliği çalışmalarında önemli bir yeri olan faktör analizi, açımlayıcı (exploratory) ve doğrulayıcı (confirmatory) olmak üzere ikiye ayrılmaktadır (McDonald, 1999). Eğer eldeki verilerin psikolojik uzaydaki yapısının nasıl olduğu bilinmiyorsa ve bu yapı keşfedilmek isteniyorsa, açımlayıcı faktör analizi uygulanır; maddelerin psikolojik uzaydaki yerleşimleri öngörülüyorsa ve bu öngörünün doğruluğu test edilmek isteniyorsa da doğrulayıcı faktör analizi uygulanır.

Maddelerin faktörlere verdiği yüklerin büyüklüğüne ve faktör yük dağılımına göre, basit (simple), yaklaşık basit (approximate simple) ve karışık (complex) olmak üzere 3 farklı yapı tanımlamak mümkündür. Eğer çokboyutlu bir testte her madde sadece ve sadece bir boyutu ölçüyorsa böyle bir çokboyutlu yapı basit yapı olarak adlandırılabilir (McDonald, 1999). Böyle bir yapıda maddelerin sadece bir boyuta yük vermesi gerekmektedir. Gerçek verilerde bu durumun çok ender ortaya çıkmasından dolayı “yaklaşık basit” adıyla başka bir yapı tanımlanmıştır (Walker, Azen ve Schmitt, 2006). Çokboyutlu bir yapıda madde birden fazla faktöre yük veriyor, ancak birincil olarak sadece bir faktörü ölçüyorsa, bu şekildeki yapılar yaklaşık basit yapı olarak adlandırılmaktadır. Çokboyutlu bir yapıda maddenin birden fazla faktöre yük vermesi ve maddenin ait olduğu faktör ayrımının yapılamadığı faktör yapıları ise “karışık yapı” olarak adlandırılabilir (McDonald, 1999).

Boyutluluk ve buna bağlı olarak tekboyutluluk incelemelerinde faktör analizi sıkça kullanılmasına rağmen bu incelemeler için başka yöntemler de bulunmaktadır. Özellikle tekboyutluluğun sınanması ölçme ve değerlendirme uygulamaları için önemli yer tutmaktadır. Hattie (1985), tekboyutluluk inceleme yöntemlerine ilişkin yaptığı bir derleme çalışmasında, yöntemleri; cevap yapısına bağlı yöntemler, güvenilirlik katsayısına bağlı indeksler, madde sayısını düzeltmeye yönelik indeksler, temel bileşenlere dayalı indeksler, faktör analizine dayalı indeksler, örtük özellik modellerine dayalı indeksler olarak gruplandırmıştır. Bu gruplandırmaya ek olarak Stout'un öncülük ettiği koşullu kovaryans teorisine dayalı çalışmalar son yıllarda popülerlik kazanmıştır (Zhang ve Stout, 1999).

Çokboyutlu verilerin tekboyutlu modellerle analiz edilmesinde yer alan önemli bir kavram ise referans bileşkesidir. Maddeye verilecek olan tepkinin belirlenebilmesi için birden fazla yetenek gerektiren bir veri kümesi birtek yetenekle açıklanmaya çalışılırsa, bir başka deyişle çokboyutlu bir yapının tekboyutluluk sayıltısı gerektiren MTK ile çözümlenmesine gidilirse, çokboyutlu yetenek uzayını, tekboyutlu şekilde temsil edecek bir yetenek doğrusuna ihtiyaç duyulur. Çokboyutlu uzayda yer alan yetenek noktalarının projeksiyonu, bu tekboyutlu doğrunun bir kestirimini verecektir (Reckase, 2009). Wang (1985, 1986) temsilci doğrunun Çokboyutlu Madde Tepki Kuramı parametrelerinden olan a parametreleri aracılığıyla bulunabileceğini belirterek bu doğruya *referans bileşkesi* (*reference composite*) adını vermiştir (akt: Reckase, 2009)

İlgili Araştırmalar

Drasgow ve Parsons (1983), faktör analitik yaklaşımdan hareketle ikinci dereceden bir faktörün ortak faktörler arası korelasyonu kontrol etmesine olanak sağlayacak şekilde tanımladıkları 5 ortak faktör barındıran veri simülasyonu yapmıştır. Simülasyonlar, 1000 birey ve 50 madde için yapılmıştır. Yapılan simülasyonda veri kümeleri gerçek tekboyutlu durumdan çokboyutlu duruma doğru değişimlenmiştir. Üretilen verilerden şans parametresinin olduğu ve olmadığı durum için, LOGIST kullanılarak, tekboyutlu MTK kestirimleri yapılmıştır. Faktör yüklerinden yola çıkılarak elde edilen madde parametreleriyle, MTK kestirimi sonucu elde edilen madde parametreleri arasındaki uyum RMSD kullanılarak incelenmiştir. Birey parametrelerinin uyumu için hesaplanan genel faktör puanlarıyla kestirilen θ yeteneği arasındaki ilişkiye bakılmıştır. Sonuç olarak MTK'nin orta düzeyde heterojenlik gösteren madde kümelerine de uygulanabileceğini belirtmişlerdir. Bununla beraber b parametresine ilişkin değerlerin çokboyutlu yapılardan, özellikle şans başarısının dahil edildiği koşullar için, tekboyutlu kestiriminde güçlüklerle karşılaşabileceğini belirtmişlerdir.

Asley ve Forsyth (1985), tamamlayıcı olmayan modele göre üretilmiş 2 boyutlu verilerden tekboyutlu kestirimler yapmıştır. Veri üretimi için gerçek bir teste ilişkin madde parametreleri kullanılmış, bununla beraber θ 'lar arasındaki korelasyon 0.00, 0.30, 0.60, 0.90 ve 0.95 olacak şekilde değişimlenmiştir. Buna göre kestirilen \hat{a} değerleri a_1 ve a_2 değerlerinin ortalamasına yakın; \hat{b} değerleri b_1 değerlerinden daha fazla olacak şekilde; $\hat{\theta}$ 'lar ise gerçek θ_1 ve θ_2 değerlerinin ortalamasıyla yüksek ilişki göstermiştir.

Harrison (1986), hiyerarşik faktör modeline göre veri üretmiş; simülasyon ölçütleri olarak 3 farklı test uzunluğu, 2 ortak faktör sayısı, 2 farklı faktör yükü dağılımı, 3 ortak faktörlerin genel faktöre bağlayan yük dağılımı kullanılmıştır. Bütün ölçütlere göre çaprazlama yapılmamakla beraber, 27 deneysel hücre oluşturulmuştur. Bu hücreler 1000 bireylik 5'er üretime tabii tutulmuştur. Test uzunlukları 30, 50, 70 madde olarak belirlenmiş, ortak faktör sayısı 4 ve 8 olarak ele alınmış, madde yük dağılımı ise uniform veya yüksek derecede çarpık olarak belirlenmekle beraber, basit yapı temel alarak faktör yüklenmesi sağlanmıştır. Ortak faktörleri genel faktörlere bağlayan bağlar ise, 3 farklı kuvvette ele alınmıştır. Faktör analizi ve MTK arası geçişten yararlanılarak elde edilen parametreleri, gerçek parametre değerleriyle karşılaştırmak amacıyla, RMSD indeksinden yararlanılmıştır. Tekboyutlu kestirim için kullanılan LOGIT programı, genel olarak bütün koşullar için madde ve yetenek parametrelerini etkili bir şekilde kestirebilmiştir. a parametresi açısından bakıldığında, ilk dikkat çeken bulgunun, diğer koşullardan bağımsız bir şekilde, test uzunluğu arttıkça a parametresinin daha etkili bir şekilde kestirilebildiğidir. a parametresi için elde edilen diğer bir bulgu, ortak faktör sayısı arttıkça a parametresine ilişkin RMSD değerlerinin düştüğü, bir başka deyişle a parametresinin daha etkili bir şekilde kestirildiği yönündedir. a parametresi için göze çarpan diğer bir bulgu ise, uniform faktör yük dağılımlarında a parametresinin daha etkili bir şekilde kestirilebilmesidir. Genel olarak a parametresinin başarılı bir şekilde kestirildiği belirtilmiştir. b parametresi açısından bakıldığında ise daha önceki çalışmalarda da olduğu gibi b parametresinin a parametresinden daha etkili bir şekilde kestirildiği belirtilmiştir. Test uzunluğunun b parametresi için de aynı etkiyi yarattığı ve uzun testlerin daha etkili b kestirimiyle sonuçlandığı bulgusuna ulaşılmıştır. İkinci dereceden faktörün gücü azaldıkça b parametresinin kestirimi de kötüleşmiştir. b parametresi için faktör yük dağılımının çarpıklığının önemli bir etkisi olmadığı gözlenmiştir. Genel olarak LOGIT'ten elde edilen p parametresi kestirimlerinin tekboyutluluk sayılına oldukça dayanıklı olduğu belirtilmiştir.

Way, Ansley ve Forsyth (1988), tamamlayıcı ve tamamlayıcı olmayan modele göre üretilen verilerden tekboyutlu kestirimler yapmışlardır. Tamamlayıcı olmayan modele göre yapılan kestirimler Asley ve Forsyth'in (1985), yaptığı çalışmayla aynı sonuçları vermiştir. Tamamlayıcı modele göre yapılan kestirimlerde ise; \hat{a} değerleri a_1 ve a_2 değerlerinin toplamına yakın sonuç verirken, \hat{b} değerleri b_1 ve b_2

*MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK
ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE
İNCELENMESİ*

değerlerinin aritmetik ortalamasına yakın, $\hat{\theta}$ 'lar ise gerçek θ_1 ve θ_2 değerlerinin ortalamasıyla yüksek ilişki göstermiştir.

Reckase, Ackerman ve Carlson (1988), iki veri kümesi kullanarak çalışmışlardır. Bunlardan birincisi; iki parametrelili çokboyutlu lojistik modele göre, iki değişkenli standart normal dağılımdan seçkisiz seçilmiş, 2000 yetenek vektörüne dayalı, simülasyon yoluyla üretilmiş veridir. Bu çokboyutlu yapıda toplam 80 maddeye yer verilmiş, ilk 20 maddenin açısı 90° , ikinci 20 maddenin 45° , üçüncü 20 maddenin 0° , dördüncü 20 maddenin açıları ise 0° ile 90° arasında eşit saçılma sahip olacak şekilde ayarlanmıştır. D değerleri ise ortalaması 0.00, standart sapması 0.60 olan normal dağılımdan seçkisiz seçilmiştir. İkinci veri kümesi ise, gerçek bir test uygulamasını içermektedir. 2738 bireyin katıldığı, ilk 40 maddesi matematik kullanımı alt testine, ikinci 28 maddesi sosyal çalışmalar okuma testine aittir. Her iki veri kümesi ÇBMTK analiz programlarından olan MIRTE ve tekboyutlu analiz programlarından olan LOGIST ile analiz edilmiştir. Programlardan elde edilen parametreler tekboyutluluk sayılısının ihlalinin göstergelerinden olan Q_3 istatistiğinin hesaplanması için kullanılmıştır. MIRTE sonuçlarına dayalı olarak maddeler kümelendikten sonra madde kümeleri LOGIST ile analiz edilmiştir. Elde edilen parametrelerden tekrar Q_3 istatistiği hesaplanmıştır. Elde edilen kümelerin tekboyutluluk sayılısını karşıladığı sonucuna ulaşılmıştır.

Ackerman (1989), 4 tamamlayıcı 4 de tamamlayıcı olmayan model için, boyutları farklı düzeylerde korelasyon gösteren 2 boyutlu yapılardan oluşan 8 veri kümesi üretmiştir. Her veri kümesinde güçlük, boyutla bağdaştırılmıştır. Her veri kümesi BILOG ve LOGIST kullanılarak analiz edilmiştir. Boyutluluğun güçlükle bağdaştırılması sonucu elde edilen yapılardaki parametre kestirimlerinin, BILOG programında LOGIST programından daha çok etkilendiği bulunmuştur.

Zeng (1989), tekboyutlu modellerin çokboyutluluğa karşı sağlamlığını belirlemek amacıyla yaptığı çalışmada boyutların olasılığa vereceği katkıyı değişimleyerek 1000'er bireylik ve 50'şer maddelik veri kümeleri üretmiştir. Birinci boyutun ölçülmek istenen özelliği düşünülerek karşılaştırmalar bu boyut çerçevesinde yapılmıştır. Karşılaştırmalar için hem parametre karşılaştırması hem de madde karakteristik eğrisi karşılaştırmalarına gidilmiştir. Sonuç olarak 1 parametrelili modelin, ana boyutun performansın %80'inden fazlasını açıkladığı durumlar için çokboyutluluğa karşı oldukça dayanıklı olduğunu; iki parametrelili modelin 1 parametrelili modelden daha az dayanıklı olduğunu belirtmiştir. Parametreler açısından bakıldığında ise; a parametresinin boyutluluğa duyarlı olduğu ve bazı koşullar için tekboyutluluğun ihlalinde kabul edilemeyecek düzeyde hatalar üretildiği bulunmuştur. Yetenek ve b parametresinin ana boyutun performansın en az %80' i açıkladığı durumlar için, daha küçük hatalar ürettiği belirtilmiştir.

Luecht ve Miller (1992), çokboyutlu testlerin tekboyutlu kalibrasyonu ve yorumlanması için yeni bir yöntem önermişlerdir. Bunun için her biri 2000 bireyden oluşan 50 maddelik 2 boyutlu 3 parametrelili çokboyutlu normal ogive modele göre veri üretimi yapılmıştır. 20° ve 70° olmak üzere 2 bileşke referans açısı belirlenmiş ve açıların standart sapması 0, 5 ve 10 olacak şekilde değişimlenmiştir. Çokboyutlu maddeler çokboyutlu uzaydaki, yönlerine göre kümelenecek ve elde edilen her ayrı küme için tekboyutlu kestirim yapılmıştır. Yetenek güçlük ve ayırt edicilik için yapılan tekboyutlu ve çokboyutlu kestirim sonuçları gerçek değerlerle beraber ele alınarak incelenmiştir. Sonuç olarak, iki basamaklı kestirimin uyum indekslerinden taviz vermeksizin kestirim dayanıklılığı ve yorumlanması açısından kazanç sağlayacağı belirtilmiştir.

Gelbal (1994), Madde Tepki Kuramı modellerinden olan Rash modeli ile Klasik Test Kuramı modelleri arasındaki benzerliği araştırmak amacıyla, 2072 bireyin tepkide bulunduğu Türkçe testini ve 2077 bireyin tepkide bulunduğu matematik testini kullanmıştır. Verilerin normal dağılmaması üzerine bireyler çıkarılarak veriler normalleştirilmiştir. Hem normal dağılan hem de normal dağılmayan veriler

Rasch modeline göre ve klasik kurama göre analiz edilmiştir. Madde parametreleri açısından bakıldığında, her iki teoriye ilişkin parametreler arasında, dağılımdan bağımsız olmak üzere, yüksek benzerlik bulunmuştur. Bununla beraber; normal dağılmayan veriden elde edilen madde istatistiklerine ilişkin standart hataların daha düşük olduğu görülmüş, bunun nedeni de örneklem büyüklüğüne bağlanmıştır. Bireylere ilişkin parametreler açısından yapılan incelemede klasik kurama göre elde edilen güçlük parametresinin kullanılması sonucu elde edilen yetenek kestirimleriyle, bg parametresi kullanılarak kestirilen yetenek parametreleri arasında çok yüksek korelasyon elde edilmiştir.

Kirisci, Hsu ve Yu (2001), test boyutluluğunun ve yetenek dağılımının etkilerini BILOG, MULTİLOG ve XCALİBRE programlarının maddelere ve bireylere ilişkin parametre kestirimlerinin uygunluğu açısından değerlendirmiştir. Yapılan çalışmada 2 farklı boyutluluk türü ele alınmıştır: Bunlardan ilki tekboyutlu yapıdır, ikincisi ise yetenekler arası korelasyonun 0.6 olduğu 3 boyutlu yapıdır. Yetenek dağılımları ise normal, çarpık ve sivri dağılımlar olarak ele alınmıştır. Çaprazlama sonucu oluşan hücrelere göre elde edilen yapılar BILOG, MULTİLOG ve XCALİBRE kullanılarak analiz edilmiştir. Ölçüt olarak gerçek ve kestirilen parametreler arası RMSE kullanılmıştır. Değerlendirmelerde elde edilen RMSE değerleri bağımlı değişken olarak kullanılmış ve varyans analizi yapılmıştır. Yetenek dağılımının temel etkisi ve diğer faktörlerle olan ortak etkisi anlamsız çıkmıştır. Boyutluluk türünün c_i için olan temel etkisi anlamlı çıkmıştır. Kestirim programların temel etkisi ise b_i , c_i , yetenek için anlamlı çıkmıştır. Boyutluluk ve program ortak etkisi ise c_i haricindeki bütün parametreler için anlamlı çıkmıştır. Buna göre kalibrasyon programı ve boyutluluk arasında bir ortak etki olduğu ve tekboyutluluğa dayanıklılığın, programın bir fonksiyonu olduğu sonucuna ulaşılmıştır. Bununla beraber, kestirim programlarının yetenek dağılımına duyarsız olduğu bulgusuna da ulaşılmıştır. BILOG'un en küçük RMSE değerine sahip olduğu, MULTİLOG ve XCALİBRE'nin ise tekboyutluluk ihlal edildiğinde parametre kestiriminde daha az değişim gösterdiği bulgusuyla beraber her 3 programında tek boyutluluk ihlaline karşı oldukça dayanıklı olduğu çıkarımını yapmışlardır.

Kelecioğlu (2001), 1994 yılında uygulanmış olan Anadolu Liseleri Giriş Sınavı'nın 30'ar maddelik Türkçe ve Matematik alt testlerine ilişkin 24701 bireyin tepkide bulunduğu verileri kullanarak KTK'ye ve MTK'ye ilişkin madde parametreleri arasındaki ilişkiyi incelemiştir. KTK'den elde edilen madde istatistikleri, geçiş formülleri kullanılarak MTK'deki denklemlerine dönüştürülmüş ve parametreler arası ilişki incelenmiştir. Bu inceleme, ayrıca puan dağılımlarının normal olup olmamasının parametre kestirimine yönelik etkilerini de barındırmaktadır. Buna göre, her iki puan dağılımı da normallik testleri sonuçlarına göre normal olmamasına rağmen, parametreler arasında oldukça yüksek korelasyon değerleri elde edilmekle beraber, Türkçe puanlarının Matematik puanlarından daha çok normal dağılım göstermesi nedeniyle, Türkçe testine ilişkin parametreler arasındaki korelasyonların daha yüksek çıktığı sonucuna ulaşılmıştır.

Doğan (2002), aynı evrene ait farklı yetenek dağılımı özellikleri gösteren örneklemelerden elde edilen KTK ve MTK madde istatistiklerinin birbirleriyle ve evren parametreleriyle olan ilişkilerini incelemiştir. MTK incelemesi 1, 2, 3 parametrelili modeller için ayrı ayrı yapılmıştır. İlişkilerin incelenmesinde Spearman Sıra Farkları Korelasyon Katsayısı kullanılmış ve korelasyon katsayıları arası farkın test edilmesinde Fisher'in z' dönüştürmesinden faydalanılmıştır. Buna göre hem KTK'den elde edilen hem de MTK'den elde edilen madde istatistiklerinin örnekleme bağılı olarak değişebildiği, her iki kuramın da puan dağılımından etkilendiği, her iki kurama ilişkin parametrelerin ilişkili olduğu ve farkın kuramsal anlamda kaldığı, MTK'nin parametre değişmezliğini sağlamadığı, bu konuda KTK'nin daha iyi sonuçlar verdiği rapor edilmiştir.

Kan (2004), tekboyutlu MTK'de madde parametrelerinin değişmezliğini test etmek için, 553108 öğrencinin katıldığı Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı'nın 25 maddelik Türkçe alt testini kullanmıştır. Bu çalışmada yüksek, orta ve düşük yetenek düzeyine sahip gruplardan kestirilen madde parametrelerinin evren parametre değerleriyle ve birbirleriyle olan tutarlılığı incelenmiştir. Buna göre düşük ve yüksek, orta ve yüksek, evren ve yüksek yetenek grubundan kestirilen a parametreleri arasında istatistiksel olarak anlamlı bir ilişki elde edilememiştir.

MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE İNCELENMESİ

Kan (2006), KTK'den ve MTK'den elde edilen madde parametrelerinin karşılaştırmasına yönelik yaptığı çalışmada, OKS Türkçe alt testine katılan 553108 öğrencinin verdiği tepkileri BILOG ve ITEMAN kullanarak analiz etmiştir. KTK'ye ve MTK'ye dayalı kestirilen parametreler arası ilişkiler, Spearman Sıra Farkları Korelasyon Katsayısı kullanılarak belirlenmiştir. Elde edilen sonuçlara göre ayırt edicilik parametrelerinin ve güçlük parametrelerinin kendi içlerinde oldukça yüksek ilişki gösterdiği bulgusuna ulaşılmıştır.

Levy (2006), Bayesian model kontrol prosedürlerini koşullu kovaryans teorisi ışığında kullanarak ve simülasyon yaparak MTK'de çokboyutluluğu incelemiştir. Bu çalışmayı yaparken boyutluluğu etkileyebilecek faktörler hipoteze edilmiş ve bu faktörler de incelemeye dahil edilmiştir. Boyutluluğu etkileyen 5 faktör olarak; veri üretim modeli (tamamlayıcı ve tamamlayıcı olmayan modeller), çokboyutlu madde sayısı oranı, boyuta bağlanma gücü, boyutlar arası korelasyon ve örneklem büyüklüğü olarak rapor edilmiştir. Boyuta bağlanma gücü, ÇBMTK parametrelerinden a parametreleri değiştirilerek kontrol edilmiştir.

Zhang (2008), çeşitli boyutluluk özellikleri içeren verileri MULTILOG kullanarak tekboyutlu MTK çerçevesinde analiz etmiştir. Simülasyon düzeneğinde ise 4 yetenek boyutunun, 15, 30 ve 60 olmak üzere 3 test uzunluğunun, %20, %40 ve %60 olmak üzere 3 farklı boyutlara duyarlı madde oranının ve 0.00, 0.40, 0.80 olmak üzere 3 farklı boyutlar arası korelasyon düzeyinin çaprazlanması sonucu oluşan deneysel hücreler kullanılmıştır. Herbir deneysel hücre için 1000'er bireyden oluşan 100 veri kümesi üretilmiştir. Verilerin analizinde korelasyondan ve RMSE uyum istatistiğinden yararlanılmıştır. Belirtilen analizlere ek olarak, temel boyutluluk incelemesi için DIMTEST kullanılarak inceleme yapılmıştır. DIMTEST sonuçlarına göre boyutluluk oluşumunda en önemli faktör olarak boyutlar arası korelasyon ön plana çıkmıştır. Korelasyon arttıkça veri tekboyutlu hale gelmiştir. Buna ek olarak uzun testlerin ve boyut sayısının tekboyutluluğa negatif etkisi görülmüştür. Kısa ve birden fazla boyuta sahip testlerin RMSE değerlerinin yüksek çıktığı, yetenek kestiriminin uygunluğunun test uzunluğu arttıkça iyileştiği belirtilmiştir. Madde parametreleri açısından yapılan incelemeler sonucunda ise; a parametresi için çokboyutlu yapılarda negatif yanlılık bulunduğu, bir başka deyişle, olduğundan daha az olacak şekilde kestirildiği (underestimated), boyutlar arası korelasyon arttıkça a parametresine ilişkin kestirimlerin iyileştiği, test uzunluğunun ve ikinci boyutu ölçen madde sayısının çok az etkisinin bulunduğu rapor edilmiştir. Güçlük parametresi olan b değeri için yapılan incelemelerde ise, elde edilen RMSE değerlerinin, a parametresi için elde edilen değerlerden daha düşük olduğu ve simülasyon koşulları için elde RMSE değerlerinin birbirine oldukça yakın olduğu belirtilmiştir. Genel olarak, tekboyutlu modellerin çokboyutlu yapılara karşı oldukça dayanıklı olduğu vurgulanmıştır.

Araştırmanın Amacı ve Önemi

Madde parametrelerinin değişmezliği, özellikle tekboyutlu MTK'nin uygulamaları için son derece önem taşımaktadır. Madde parametrelerinin değişmezliğine yönelik MTK çalışmalarında, madde parametrelerinin değişebildiği bulgusuna sıkça ulaşılmıştır. Lord (1980), parametre değişmezliğinin sağlanabilmesinin ancak tekboyutlu yapılarda ve standart bir yetenek ölçeklemesinin yapıldığı durumlarda oluşabileceğini çok açık bir şekilde belirtmiştir. MTK'nin tekboyutluluk sayılısının yeterince karşılanmaması sonucu yerel bağımsızlık sayılısı da beraberinde bozulacaktır. Yerel bağımsızlık sayılısının bozulması ise yanlı parametre kestirimleriyle sonuçlanacaktır.

Çokboyutluluğun, tekboyutlu MTK'de parametre değişmezliğini bozabileceği düşünülmektedir. Tekboyutlu MTK modellerinin çeşitli faktör yapılarına dayanıklılığını inceleyen çalışmalar olmakla beraber, yapılan incelemelerde parametre değişmezliğine odaklanan bir çalışmaya ulaşılamamıştır. Bu açıdan bakıldığında, yapılan çalışma bu boşluğu doldurma yolunda bir adım olacaktır. Yapılan çalışma ayrıca, MTK'nin hangi faktör yapıları için uygulanabileceği ve hangi faktör yapıları için uygulanamayacağı konusunda literatüre parametre değişmezliği bağlamında katkıda bulunacaktır.

Yukarıda belirtilen incelemeler herbir boyut için ayrı ayrı yapılacağından, hangi boyutun parametre değişmezliği bağlamında nasıl etkileneceği de görülebilecektir.

Belirtilen nedenlerden dolayı bu çalışmada çeşitli boyutluluk özelliklerine sahip yapılarda madde parametrelerinin değişmezliği MTK çerçevesinde araştırılmıştır.

Problem Cümlesi

I. Çeşitli boyutluluk özellikleri gösteren yapıların MTK'ye göre analiz edilmesi sonucu madde parametre değişmezliğinin göstergesi için elde edilen standart hata ortalamaları nasıl değişmektedir?

I.1. Birinci boyutta yer alan maddelerin MTK b parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

I.2. İkinci boyutta yer alan maddelerin MTK b parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

I.3. Birinci boyutta yer alan maddelerin MTK a parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

I.4. İkinci boyutta yer alan maddelerin MTK a parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

YÖNTEM

Araştırma Türü

Bu çalışma çeşitli boyutluluk özelliği gösteren yapılarda madde parametrelerinin değişmezliğini MTK çerçevesinde ele alması nedeniyle, temel araştırma olarak değerlendirilebilir.

Veri Üretim Çalışması

Veri üretim çalışması MIRTGEN 2.0 (Luecht, 2004) kullanılarak yapılmıştır. Üretim, 2 boyutlu ve iki kategorili (1-0) tamamlayıcı madde tepki kuramına dayalı olarak gerçekleştirilmiştir. Birinci faktörde yer alan 24 maddenin özellikleri bütün üretim koşulları için sabit tutulmuştur. Değişimlemeler ikinci faktördeki madde sayısı ve madde özellikleri ile boyutlar arası korelasyonlar üzerinde yapılmıştır. Üretilmiş olan bütün yapılarda basit yapı kullanılmıştır; dolayısıyla birinci faktör maddeleri sadece a_1 ve d değerine sahip olurken, ikinci faktör maddeleri sadece a_2 ve d değerlerine sahip olmaktadır. Birinci faktörde yer alan maddelerin a_1 değerleri ortalaması 1.00 ve standart sapması 0.05 olan normal dağılımdan elde edilmiştir. Birinci faktör maddelerine ilişkin d değerleri ise minimum değeri -3, maksimum değeri +3 olan uniform dağılımdan elde edildikten sonra, MDIFF değerlerinin gerçekçi değerler almasını sağlamak amacıyla çokboyutlu ayırtedicilik parametresi olan MDISC kullanılarak ölçeklenmiştir. İkinci faktör maddelerinin üretimi içinde aynı yol izlenmekle beraber, ikinci faktörde yer alan maddelerin sayısı ve a_2 parametre ortalaması değişimlenmiştir. Bu amaç doğrultusunda, ikinci faktörde yer alan madde sayıları 4, 8, 12, 16, 20, 24 olarak alınırken (Grafik üzerinde sırasıyla N4, N8, N12, N16, N20, N24 olarak gösterilmiştir.), bu maddelerin a_2 parametre ortalaması ise 0.25, 0.50, 0.75 ve 1.00 olarak (Grafik üzerinde sırasıyla A25, A50, A75 ve A100 olarak gösterilmiştir.) değişimlenmiştir. Bu durumda $6 \times 4 = 24$ farklı deneysel hücre

MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE İNCELENMESİ

oluşturmuştur. Madde vektörlerine açı verebilmek amacıyla, yukarıda belirtilen koşullara boyutlar arası korelasyonlar da eklenmiştir. Boyutlar arası korelasyonlar 0.00, 0.30, 0.60 ve 0.80 olarak (Grafik üzerinde sırasıyla C0, C3, C6, C8 olarak gösterilmiştir.) değişimlenmiştir. Böylelikle $6 \times 4 \times 4 = 96$ farklı deneysel yapı ortaya çıkmıştır. Üretim düzeneği Tablo 1'de sunulmuştur.

Yapılan üretimin sağlıklı olup olmadığını denetlemek için, üretilmiş olan evren verileri NOHARM III ile analiz edilerek, herbir yapıya ilişkin a_1 , a_2 ve d katsayıları kestirilmiştir. Kestirilen bu değerlerle tasarlanan değerler arasındaki korelasyon değerleri hesaplanmıştır. a_1 parametresi için hesaplanan korelasyon katsayısının 0.97 ile 0.99 arasında değiştiği, a_2 parametresi için hesaplanan korelasyon katsayılarının 0.99 olduğu ve d için hesaplanan korelasyon katsayılarının 0.99 olduğu görülmüştür. Bu bulgulardan sonra evren üretimlerinin ve kestirimlerinin sağlıklı bir şekilde gerçekleştirildiği sonucuna ulaşılmıştır. Bu işlem, parametreleri teker teker kontrol etmeye yaradığından her üç parametrenin kestirilen ve tasarlanan değerleri 3 boyutlu grafik üzerinde örtüşmeleri incelenmiş ve tasarlanan ve kestirilen parametrelerin uyduğu gözlenmiştir.

Örneklerin Belirlenmesi

Örneklem büyüklüğünün parametre kestirim sürecine etkisini inceleyebilmek amacıyla, 3 farklı örneklem büyüklüğü kullanılmıştır. Bu amaç doğrultusunda, her bir evrenden 500'lük 100 adet, 1000'lik 100 adet ve 3000'lik 100 adet örneklem çekilmiştir (Grafik üzerinde sırasıyla S500, S1000 ve S3000 olarak gösterilmiştir.). Örneklem yöntemi olarak basit seçkisiz örneklem ve yerine koyma metodu kullanılmıştır. Böylelikle $96 \times 3 \times 100 = 28\ 800$ adet örneklem elde edilmiştir. Örneklem işlemi, araştırmacı tarafından R programlama dilinde bir program yazılarak gerçekleştirilmiştir.

İşlem

MTK kapsamındaki a ve b parametreleri BILOG MG 3 kullanılarak kestirilmiştir. Bu işlem birinci boyut ve ikinci boyut maddeleri beraber analiz edilerek yapılmıştır. MTK için veri kestirimi 2 parametrelili logistik modele dayalı olarak gerçekleştirilmiştir. BILOG MG 3 kestirimlerinin "batch" modunda yapılabilmesi için gereken komut dosyaları, program yazılarak elde edilmiştir. Elde edilmiş olan 28 800 çıktı dosyası her bir parametre, örneklem büyüklüğü ve değişimleme ölçütlerine göre düzenlenmiştir. Bu düzenleme sonucu elde edilen matrislerin satır sayısı madde sayısına (birinci boyut madde sayısı+ikinci boyut madde sayısı) sütun sayısı ise, her bir koşul ve örneklem büyüklüğü için 100'er örneklem çekildiğinden dolayı 100'e eşit olmaktadır. Bu şekilde bir düzenleme ile her bir parametre, örneklem büyüklüğü ve değişimleme ölçütü için 288 matris oluşturulmuştur (a_2 parametresi yük ortalaması \bar{x} ikinci boyuttaki madde sayısı x boyutlar arası korelasyon ρ örneklem büyüklüğü).

Verilerin Analizi

Elde edilen matrislerin satırlarındaki değişkenlik, madde parametresinin örneklemden örneklem değişme ölçütü olarak ele alınmıştır. Bu doğrultudan hareketle, matris satırlarına ilişkin standart sapmalar hesaplanmış ve bu standart sapmalar "parametre kestiriminin standart hatası" olarak değerlendirilmiştir. Değişmezliği test düzeyinde değerlendirebilmek amacıyla da, her testte yer alan madde standart hatalarının ortalaması (SHO) alınmıştır.

Çok boyutluluğun her iki boyut üzerindeki etkisini ayrı ayrı görebilmek amacıyla her boyuta ilişkin maddelerin standart hataları ayrı ayrı değerlendirilmiştir. Bu değerlendirmede her teori, her parametre, her koşul ve örneklem büyüklüğü için elde edilmiş olan standart hataların aritmetik ortalamaları her boyut için ayrı ayrı hesaplanmıştır.

SÜNBÜL

Elde edilen standart hata ortalamalarının her koşul ve örneklem büyüklüğü açısından görselleştirilerek yorumlanabilmesi için çokdeğişkenli grafiksel yöntemlerden yararlanılmıştır.

Tablo 1: Evrenlerin Oluşturulması için İkinci Boyut Üzerinde Yapılan Değişimlere İlişkin Üretim Düzenegi

Madde No	İkinci Boyutta Yer Alan Madde Sayısı					
	4	8	12	16	20	24
25	X	X	X	X	X	X
26	X	X	X	X	X	X
27	X	X	X	X	X	X
28	X	X	X	X	X	X
29		X	X	X	X	X
30		X	X	X	X	X
31		X	X	X	X	X
32		X	X	X	X	X
33			X	X	X	X
34			X	X	X	X
35			X	X	X	X
36			X	X	X	X
37				X	X	X
38				X	X	X
39				X	X	X
40				X	X	X
41					X	X
42					X	X
43					X	X
44					X	X
45						X
46						X
47						X
48						X
Boyutlar Arası Korelasyon				0.00, 0.30, 0.60, 0.80		
a ₂ Parametre Ortalaması				0.25, 0.50, 0.75, 1.00		

BULGULAR

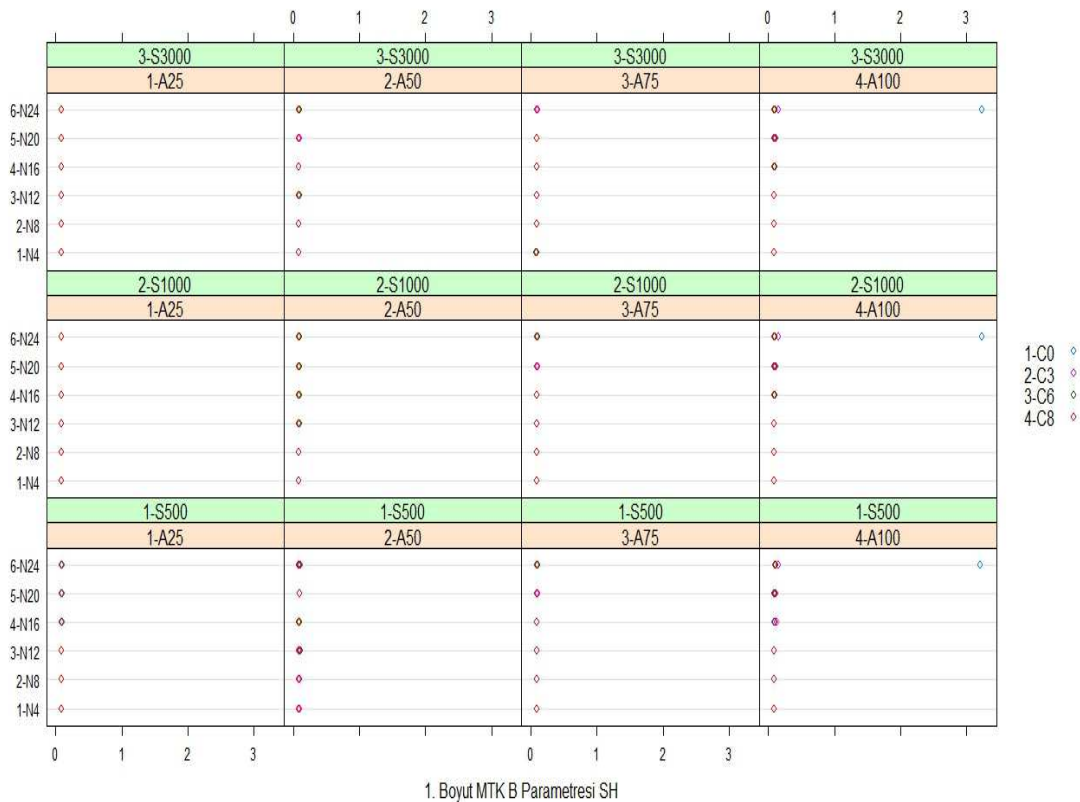
**MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK
ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE
İNCELENMESİ**

Bu bölümde, alt problemlerin sırası dikkate alınarak, alt problemlere ilişkin bulgulara yer verilmiştir. Bulguların tamamında çokboyutlu veri grafiklenmesinde kullanılan grafiklerden yararlanılmıştır. Elde edilen değerlerin tamamını kapsayabilmesini sağlayabilmek amacıyla, grafik ölçekleri normatif olarak belirlenmiştir; bu nedenle, grafiklerin karşılaştırılmasında grafik ölçeklerine dikkat edilmesi gerekmektedir.

Çeşitli boyutluluk özellikleri gösteren yapıların MTK'ye göre analiz edilmesi sonucu madde parametre değişmezliğinin göstergesi için elde edilen standart hata ortalamaları nasıl değişmektedir?

I. Birinci boyutta yer alan maddelerin MTK b parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

Birinci boyutta yer alan maddelerin MTK kestiriminden elde edilen b parametrelerinin standart hata ortalamalarının çeşitli koşullar ve örneklem büyüklükleri için dağılımı aşağıdaki grafikte sunulmuştur.



Şekil 1: 1. Boyut MTK b Parametresi SHO Saçılımı

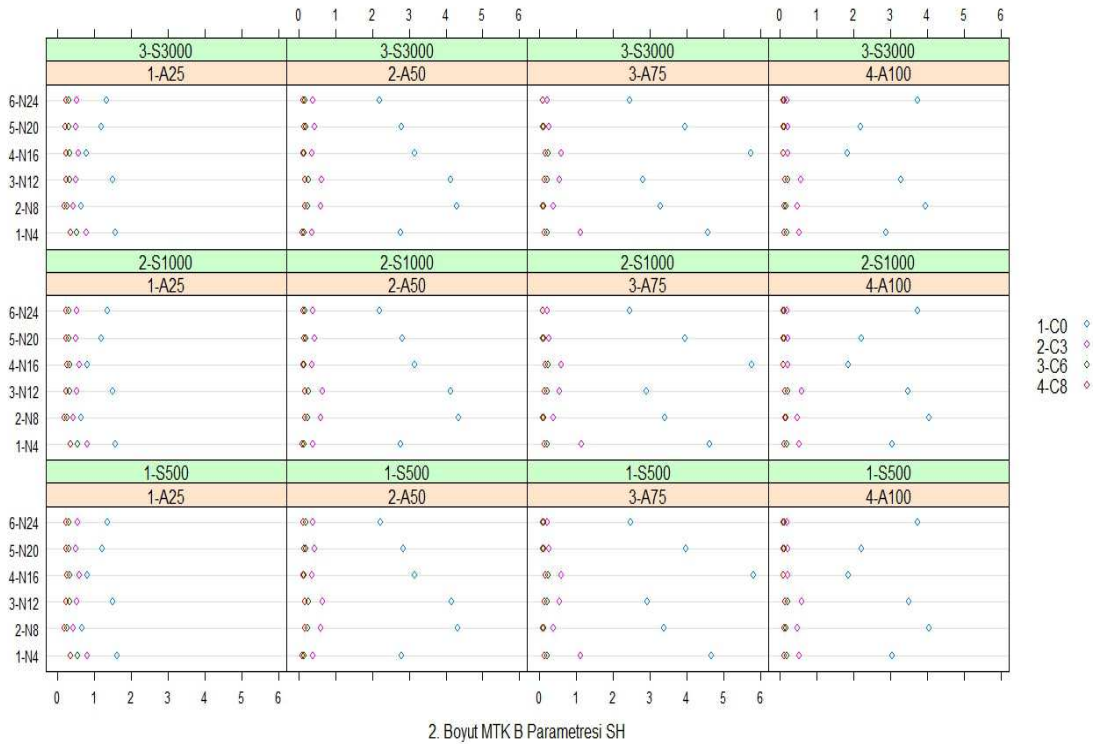
Yukarıdaki grafik incelendiğinde, birinci boyutta yer alan maddelerin b parametresi için elde edilen standart hata ortalamalarının 0.083 ile 3.233 arasında yer aldığı görülebilir. SHO dağılımlarının minimum ve maksimum değerleri dikkate alındığında, bazı durumlar için parametre değişmezliğinin korunamadığı söylenebilir. SHO dağılımlarının genel olarak birbirine benzer ve 0.00 etrafında homojen olduğu, ancak üç koşul için bu genellemenin bozulduğu görülmektedir. Bu koşulların ortak özelliği ikinci

SÜNBÜL

boyutta yer alan madde sayısının 24, ikinci faktör maddelerinin a_2 parametre ortalamasının 1.00 olduğu ve boyutlar arası korelasyonun 0.00 olduğu durumlardır. Belirtilen durumlar için SHO değerleri kabul edilemeyecek kadar büyüktür ve bu durumlarda parametre değişmezliğinden söz etmek mümkün değildir. Dikkat edilmesi gereken diğer bir nokta, örneklem büyüklüğünün b parametresi kestirimindeki önemsizliğidir; bütün örneklem büyüklüklerinde hemen hemen aynı sonuç elde edilmiştir.

II. İkinci boyutta yer alan maddelerin MTK b parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

İkinci boyutta yer alan maddelerin MTK kestiriminden elde edilen b parametrelerinin standart hata ortalamalarının çeşitli koşullar ve örneklem büyüklükleri için dağılımı aşağıdaki grafikte sunulmuştur.



Şekil 2: 2. Boyut MTK b Parametresi SHO Saçılımı

Yukarıda yer alan SHO değerleri incelendiğinde, elde edilen değerlerin 0.076 ile 5.812 arasında yer aldığı, bazı durumlar için oldukça yüksek SHO değerleri elde edildiği ve parametre değişmezliği sayılısının yerine getirilemediği söylenebilir. Yukarıdaki dağılıma bakılarak, parametre değişmezliğini tehdit eden en önemli faktörün, boyutlar arasındaki korelasyon olduğunu söylemek mümkündür. Korelasyon katsayısının değeri azaldıkça, SHO değerlerinin artma eğilimine girdiği özellikle boyutlar arası korelasyon 0.00 olduğu zaman kabul edilemeyecek kadar büyük SHO değerlerinin oluştuğu görülebilir. Bu artışı tetikleyen diğer bir faktör ise, ikinci faktörde yer alan maddelerin a_2 parametre ortalamalarıdır. a_2 parametre ortalaması arttıkça 0.00 korelasyon düzeyleri için SHO değerlerindeki büyüme de belirgin bir

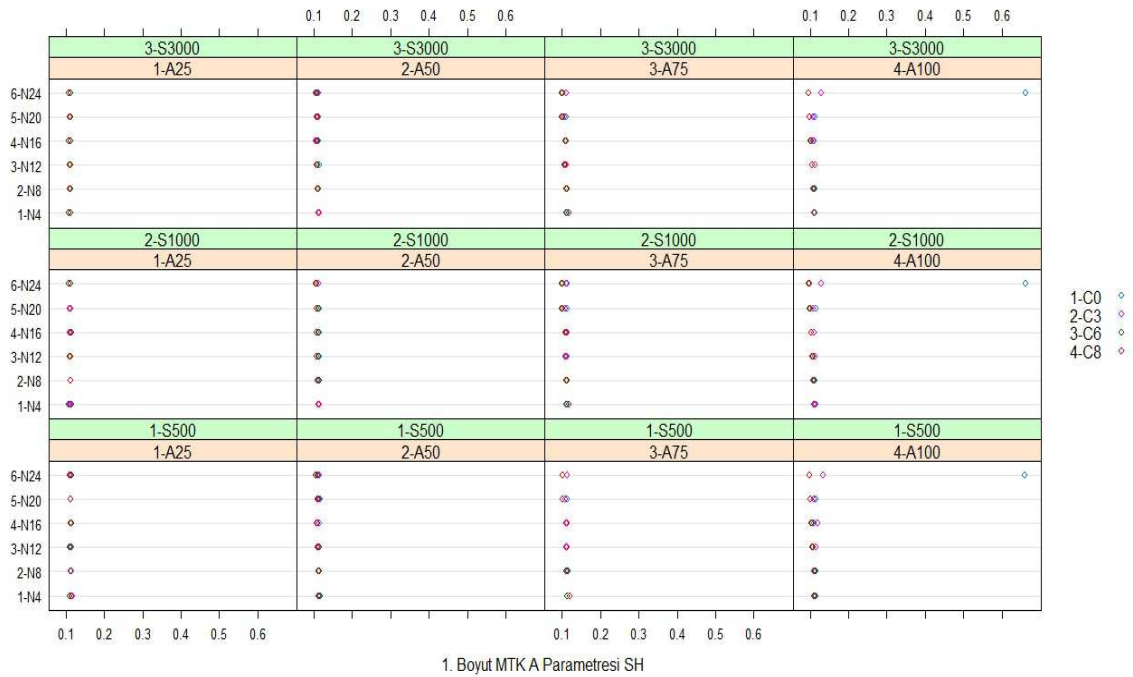
**MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK
ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE
İNCELENMESİ**

şekilde artmaktadır. İkinci boyuttaki madde sayısının ve örneklem büyüklüğünün ikinci boyutta yer alan maddelerin b parametresine ilişkin SHO değerleri açısından önemli etkisi gözlenmemiştir.

Dikkat edilmesi gereken diğer bir nokta ise, b parametresi için, birinci boyuttaki maddelerden elde edilen SHO ve ikinci boyuttaki maddelerden elde edilen SHO dağılımları arasındaki farklılıklardır. Birinci boyuttaki maddelerin b parametresine ilişkin SHO değerleri, sadece ikinci boyutun birinci boyuta kuvvet açısından eşdeğer olduğu ve boyutların korelasyon göstermediği durumlar için sapkın sonuçlar verirken, ikinci boyuttaki maddelerin b parametresine ilişkin SHO değerleri boyutların kuvvetçe eşit olmadığı durumlarda dahil olmak üzere boyutların korelasyon göstermemesinden çokça etkilenmiştir.

III. Birinci boyutta yer alan maddelerin MTK a parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

Birinci boyutta yer alan maddelerin MTK kestiriminden elde edilen a parametrelerinin standart hata ortalamalarının çeşitli koşullar ve örneklem büyüklükleri için dağılımı aşağıdaki grafikte sunulmuştur.



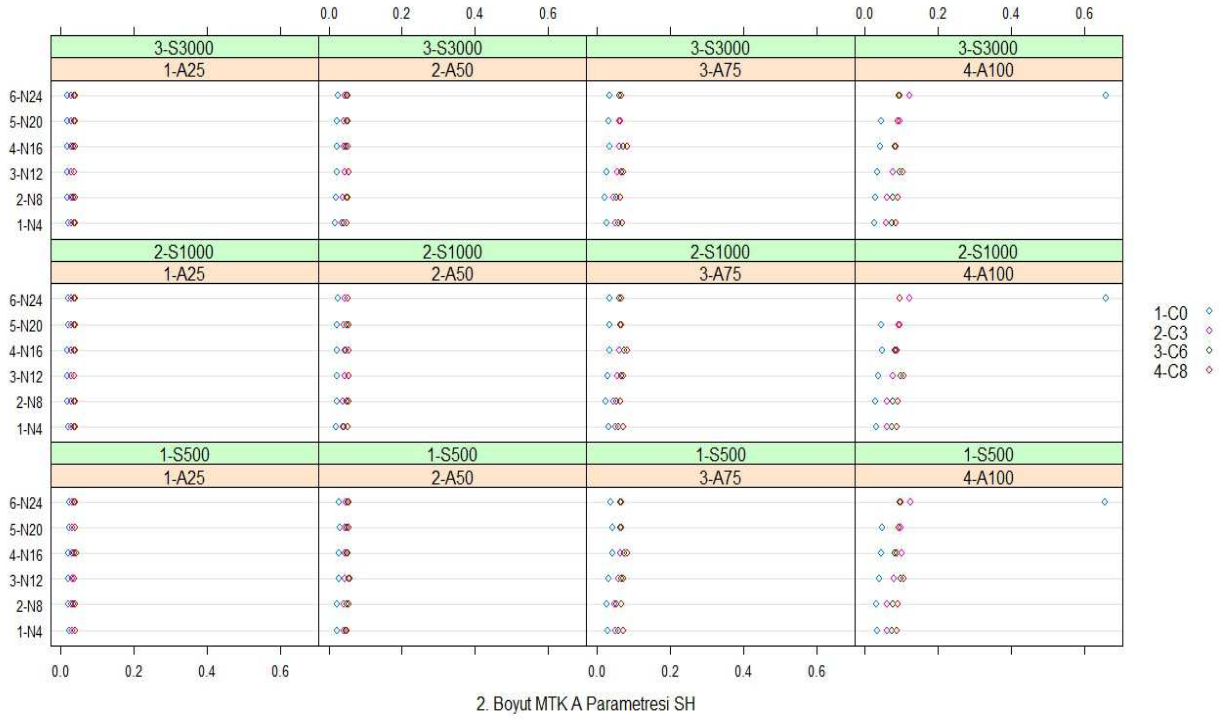
Şekil 3: 1. Boyut MTK a Parametresi SHO Saçılımı

Yukarıdaki grafik incelendiğinde, birinci boyutta yer alan maddelerin a parametresi için elde edilen standart hata ortalamalarının 0.095 ile 0.662 arasında yer aldığı görülebilir. SHO dağılımlarının minimum ve maksimum değerleri dikkate alındığında, bazı durumlar için parametre değişmezliğinin korunamadığı söylenebilir. SHO dağılımlarının genel olarak birbirine benzer, 0.10 etrafında homojen olduğu, ancak üç koşul için bu genellemenin bozulduğu görülmektedir. Bu koşulların ortak özelliği ikinci boyutta yer alan madde sayısının 24, ikinci faktör maddelerinin a_2 parametre ortalamasının 1.00 olduğu ve boyutlar arası korelasyonun 0.00 olduğu durumlardır. Belirtilen durumlar için SHO değerleri diğer durumlardan büyük kopuş göstermiştir ve bu durumlarda parametre değişmezliğinden söz etmek mümkün değildir. Dikkat

edilmesi gereken diğer bir nokta ise örneklem büyüklüğünün a parametresi kestirimindeki önemsizliğidir; bütün örneklem büyüklüklerinde hemen hemen aynı sonuç elde edilmiştir.

IV. İkinci boyutta yer alan maddelerin MTK a parametresi kestirimleri için elde edilen standart hata ortalamaları, çeşitli faktör yapıları ve örneklem büyüklükleri için nasıldır?

İkinci boyutta yer alan maddelerin MTK kestiriminden elde edilen a parametrelerinin standart hata ortalamalarının, çeşitli koşullar ve örneklem büyüklükleri için dağılımı aşağıdaki grafikte sunulmuştur.



Şekil 4: 2. Boyut MTK a Parametresi SHO Saçılımı

Yukarıda yer alan SHO değerleri incelendiğinde, elde edilen değerlerin 0.017 ile 0.659 arasında yer aldığı görülmektedir. Üç durum dışındaki SHO değerleri ise 0.00 ile 0.2 arasında yer almaktadır. Bu üç durumun ortak özelliği ikinci boyuttaki madde sayısının 24 ve boyutlar arası korelasyonun 0.00 olduğu koşullardır. Bu üç koşul için parametre değişmezliğinin korunamadığını söylemek mümkündür.

Grafikte yer alan SHO dağılımlarına ilişkin dikkat edilmesi gereken diğer bir nokta, ikinci faktör maddelerine ait a_2 parametre ortalamasının SHO dağılımları üzerindeki etkisidir. Grafikten de görülebileceği gibi a_2 parametre ortalaması arttıkça SHO değerleri artmaya ve boyutlar arası korelasyon düzeyleri için heterojenleşmeye başlamış; böylelikle boyutlar arası korelasyon düzeyleri açısından SHO'lar daha belirgin bir hale gelmiştir. Daha önce belirtilen üç durum dışındaki bütün durumlarda, beklenenin tersine, boyutlar arası korelasyon arttıkça SHO değeri de artmaya başlamış ve en küçük SHO değerleri boyutlar arası korelasyon 0.00 olduğu zaman ortaya çıkmıştır.

Genel olarak birinci boyut MTK a parametresine ilişkin SHO değerleri, ikinci boyut MTK a parametresi SHO değerlerinden daha düşük olmakla beraber ikinci boyutun birinci boyuta kuvvetçe eşit olduğu ancak boyutlar arası korelasyonun 0.00 olduğu üç koşul haricindeki bütün durumlar için parametre değişmezliğinin sağlandığı söylenebilir.

**MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK
ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE
İNCELENMESİ**

TARTIŞMA ve YORUM

SHO değerleri incelendiğinde, hem birinci boyut hem de ikinci boyut için elde edilen değerlerin genel olarak 0.00 etrafında homojen yığılım gösterdiği, bir başka deyişle parametrelerin çokboyutluluk olmasına rağmen değişme olmadan kestirilebildiği söylenebilir. Bu bulgu MTK modellerinin çokboyutluluğa karşı oldukça dayanıklı olduğunu belirten çalışmalarını destekler niteliktedir (Drasgow ve Parsons, 1983; Harrison, 1986; Zeng, 1989, Kirisci, Hsu ve Yu, 2001; Zhang, 2008). Ancak, ikinci boyut maddeleri hem madde sayısı bakımından hem de a_2 parametresi ortalamasının birinci boyut maddelerine eşit olduğunda, bir başka deyişle birinci boyutun ikinci boyuta kuvvetçe eşit olduğunda ve boyutlar arası korelasyon 0.00 olduğunda, çok yüksek SHO değerlerine ulaşılmaktadır. Bu durum, b parametre kestirimi açısından çok büyük farklılıklara yol açarken, a parametre kestirimi için daha düşük farklılıklara yol açmaktadır. Bu bulgu, Drasgow ve Parsons'un (1983), "b parametresinin a parametresine göre daha kararlı kestirilebileceği" görüşüyle tutarlıdır; ancak, Kirisci, Hsu ve Yu'nun (2001), bulgularıyla çelişmektedir. Bu durumun, referans bileşkesinin konumlanmasından kaynaklı olduğu söylenebilir. Bir başka deyişle, boyutların eşit kuvvete sahip ve dik olduğu durumda referans bileşkesinin konumlandırılması zorlaşmaktadır.

Referans bileşkesi bireylere ait bir yetenek eksenini olmasına rağmen, b parametresinin referans bileşkesi ile aynı ölçeği paylaşmasından kaynaklı olarak, referans bileşkesinin yerinin belirlenmesindeki güçlüğün b parametresi için yapılan kestirimleri zorladığı düşünülmektedir.

Tüm bunlardan da önemlisi, elde edilen bulgular MTK'nin temel sayılısını oluşturan tekboyutluluk sayılısının, parametre değişmezliği çerçevesindeki varlığını kuşku hale getirmektedir. Bulgular; gerçekte karşılaşılma olasılığı düşük uç durum olan madde sayısının 24, boyutlara ilişkin a parametre ortalamalarının eşit ve korelasyon 0.00 durumu dışında, MTK'nin her iki madde parametresi için, gayet kararlı parametre kestirimleri yaptığını göstermektedir. Bu bulgular, ayrıca, faktör analizi ile faktöriyel yapıların döndürülmesinde kullanılan yöntemlerin de tartışılması gerektiğini ortaya koymaktadır. Aşağıda sırasıyla örneklem büyüklüğüne göre, ikinci boyutta yer alan madde sayısına göre, a_2 parametre ortalamasına göre ve boyutlar arası korelasyona göre değerlendirmelerde bulunulmuştur.

Örneklem büyüklüğüne göre: MTK çerçevesinde incelenen madde parametreleri araştırmada değişimlenen örneklem büyüklüklerinden etkilenmemektedir. Buna ek olarak, bu değişmezliğin örneklem büyüklüğünün 500 olduğu durumlarda bile, daha büyük örneklem gerektireniki parametrelili modelle elde edilmiş olması da önemli bir bulgudur.

İkinci boyutta yer alan madde sayısına göre: MTK çerçevesinde incelenen madde parametreleri araştırmada değişimlenen madde sayısından sadece ikinci boyut madde sayısı birinci boyut madde sayısına eşit olduğunda etkilenmektedir.

a_2 parametre ortalamasına göre: MTK çerçevesinde incelenen b parametreleri; birinci boyutta, ikinci boyuttaki madde sayısı 24 ve boyutlar arası korelasyonun 0.00 olduğu durumda, araştırmada değişimlenen a_2 parametre ortalamalarından etkilenmektedir. İkinci boyutta ise; hangi a_2 ortalaması olursa olsun, boyutlar arası korelasyonun 0.00 olduğu koşulda, a_2 parametre ortalamalarından etkilenmektedir. a parametreleri ise, birinci ve ikinci boyutta boyutlar arası korelasyonun 0.00, ikinci boyuttaki madde sayısının 24 ve $a_2=1.00$ olduğunda etkilenmektedir.

Boyutlar arası korelasyona göre: MTK çerçevesinde incelenen b parametresi araştırmada değişimlenen boyutlar arası korelasyon büyüklüğünden, hem birinci hem de ikinci boyutta boyutlar arası korelasyonun 0.00 olduğunda etkilenmektedir. a parametresi ikinci boyut madde sayısı 24, a_2 ortalaması 1.00 ve boyutlar arası korelasyonun 0.00 olduğunda, yani ikinci boyut birinci boyuta kuvvetçe eşit olduğunda değişmektedir.

SONUÇLAR

MTK çerçevesinde yapılan madde parametre deęişmezlięi incelemelerine göre, hem güçlük hem de ayırt edicilik parametresi için elde edilen SHO deęerlerinin genel olarak küçük bir ranjda yer aldığı görülmüştür. Bu genellenenin bozulduğu çokboyutlu yapıların ortak özelliğinin ise; boyutların birbirine dik, her iki boyutta yer alan madde sayılarının 24 ve a_2 parametre ortalamalarının 1.00'e eşit olduğu faktör yapılarının olmasıdır. Belirtilen koşul için deęişmez madde parametreleri elde edilememiştir. MTK'nin çokboyutlu yapılara oldukça dayanıklı olduğu ve yukarıda belirtilen koşul haricindeki faktör yapılarında deęişmez madde parametreleri ürettięi sonucuna ulaşılmıştır..

ÖNERİLER

Yapılan çalışma iki boyutlu ve tamamlayıcı modele ilişkin yapılarla sınırlıdır. Aynı çalışma boyut sayısı artırılarak boyut sayısının parametre deęişmezlięine etkisi farklı kuramlar açısından araştırılabilir. Yapılan çalışma madde parametreleriyle sınırlıdır. Aynı çalışma madde parametreleri yerine yetenek parametreleri çerçevesinde yapılabilir. Yapılan çalışmada kullanılan simülasyon teknięi ÇBMTK'ye dayalı olarak yapılmış ve maddelerin faktörlere bağlanması a parametreleri deęişimlenerek sağlanmıştır. Aynı çalışma ikinci dereceden faktörleri barındıracak ve deęişimlenmesine olanak sağlayacak şekilde veri üretilerek yapılabilir. Aynı çalışma çeşitli açılara sahip referans bileşkeleri tanımlanarak ve maddelerin belirlenen referans bileşkesinden belirli ölçütlere göre ayrıklaşması deęişimlenerek yapılabilir. Çalışmada yer alan simülasyon ölçütleri, özellikle 0.00 ile 0.30 arasındaki boyutlar arası yetenek korelasyonu düzeyleri ve a_2 parametre ortalamasının 0.75 ile 1.00 arası, duyarlı hale getirilerek yapılabilir. Çalışmada kullanılan yapıların hepsi kendi içinde tekboyutlu doğrulayıcı faktör analizine tabii tutularak hangi yapının tekboyutluluęa uyum vereceęi incelenebilir. Aynı verilerin MTK çerçevesinde analizindeki öteleme (iteration) sayıları ve kavuşma (convergence) deęerleri incelenerek, MTK'nin hangi faktör yapısında ne derece zorlandığı ortaya çıkarılabilir. Çıkan sonuçlar yapılmış olan çalışmanın bulgularıyla beraber deęerlendirilebilir. Yapılan çalışmada MTK analizleri BILOG MG3 kullanılarak yapılmıştır. Aynı çalışma başka analiz programları kullanılarak tekrarlanabilir. Farklı analiz programlarının gücü deęerlendirilebilir. MTK çerçevesinde şans başarısı dahil edilmemiştir. Her iki teori içinde şans başarısının dahil edildięi durumlar, parametre deęişmezlięi bağlamında incelenebilir. MTK çerçevesinde, aynı faktör yapıları kullanılarak, birey ve madde uyum istatistiklerinin hangi faktör yapısında nasıl sonuç verdięi deęerlendirilebilir. MTK çerçevesinde aynı faktör yapıları kullanılarak test bilgi fonksiyonunun çeşitli faktör yapılarındaki davranışı incelenebilir. Özellikle sorunlu deneysel hücrelerde SHO ortalaması alınmak yerine her koşul için, standart hataların dağılımının incelenmesi üzerine ek çalışmalar yapılabilir.

KAYNAKÇA

- Ackerman, T. A. (1989). Unidimensional IRT calibration of compensatory and noncompensatory multidimensional items. *Applied Psychological Measurement*, 13,113-127.
- Ansley, T. N. ve Forsyth R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two dimensional data. *Applied Psychological Measurement*, 9 (1), 37-48.
- Bartholomew, D. J. (2007). Three faces of factors analysis. R. Cudeck, R. C. MacCallum, (Ed), *Factor analysis at 100:historical developments and future directions* içinde (9-23). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Baykul, Y. (2000). *Eęitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Crocker, L. ve Algina, J. (1986). *Introduction to classical and modern test theory*. USA: Rinehart and Winston Inc.
- Doęan, N. (2002). *Klasik test kuramı ve örtük özellikler kuramının örneklem bağlamında karşılaştırılması*, Yayınlanmamış doktora tezi, Hacettepe Üniversitesi, Ankara.
- Drasgow, F. ve Parsons, C. K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, 7, 189-199.
- Embretson, S. E. ve Reise, S. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates.
- Ferrando, P. J. (2009). Difficulty, discrimination and information indices in the linear factor analysis model for continuous item responses. *Applied Psychological Measurement*, 33, 9-24.

**MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK
ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE
İNCELENMESİ**

- Gelbal, S. (1994). *p madde güçlük indeksi ile rasch modelinin b parametresi ve bunlara dayalı yetenek ölçüleri üzerine bir karşılaştırma*. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi, Ankara.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- Hambleton, R. K. ve Swaminathan, H. (1985). *Item response theory: principles and applications*. Boston: Academic Publishers Group.
- Harrison, D. A. (2009). Robustness of IRT parameter estimation to violations of the unidimensionality assumption. *Journal of Educational Statistics*, 11(2), 91-115.
- Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9(2), 139-164.
- Jacoby, W. G. (1991). *Data theory and dimensional analysis*. London: Sage Publications.
- Kan, A. (2004). *Farklı yetenek ve örneklem gruplarından kestirilen madde parametrelerinin karşılaştırılması*. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi, Ankara.
- Kan, A. (2006). Klasik test teorisine ve örtük özellikler teorisine göre kestirilen madde parametrelerinin karşılaştırılması üzerine ampirik bir çalışma. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 2(2), 227-235.
- Karakaş, S. (1997). *Bilimsel psikoloji: temel ilkeler*. Ankara: Türk Psikologlar Derneği Yayınları.
- Kelecioğlu, H. (2001). Örtük özellikler teorisindeki b ve a parametreleri ile klasik test teorisindeki p ve r parametreleri arasındaki ilişki. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 20, 104-110.
- Krisci, L., Hsu, T. ve Yu, L. (2001). Robustness of item parameter estimation programs to assumptions of unidimensionality and normality. *Applied Psychological Measurement*, 25, 146-162.
- Levy, R. (2006). *Posterior predictive model checking for multidimensionality in item response theory and bayesian network*. Yayınlanmamış doktora tezi, Maryland Üniversitesi, Maryland.
- Lord, F. M. ve Novick M. R. (1968). *Statistical theories of mental test scores*. New York: Addison- Wesley Publishing Company.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Luecht, R. M. ve Miller, T. R. (1992). Unidimensional calibrations and interpretations of composite traits for multidimensional tests. *Applied Psychological Measurement*, 16, 279-292.
- Luecht, R. (2004) MIRTGEN 2.0 [Bilgisayar Yazılımı]. Yazar.
- Magnusson, D. (1968). *Test theory*. Massachusetts. Addison-Wesley.
- McDonald, R. P. (1999). *Test theory: a unified treatment*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Reckase, M. D. Ackerman, T. A. ve Carlson, J. E. (1988). Building a unidimensional test using multidimensional items. *Journal of Educational Measurement*, 25, 193-203.
- Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer Dordrecht Heidelberg.
- Rupp, A. A. ve Zumbo, B. D. (2006). Understanding parameter invariance in unidimensional IRT models. *Educational and Psychological Measurement*, 66, 63-84.
- Steinberg, L., Thissen, D. ve Wainer, H. (2000). Validity . H. Wainer. (Ed), *Computerized adaptive testing: a primer* içinde (185-230). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Tatlıdil, H. (1992). *Uygulamalı çok değişkenli istatistiksel analiz*. Ankara: Akademi Matbaası.
- Walker, C. M., Azen, R. ve Schmitt, T. (2006). Statistical versus substantive dimensionality: the effect of distributional differences on dimensionality assesment using DIMTEST. *Educational and Psychological Measurement*, 66, 721-738.
- Way, W. D., Ansley., T. N. ve Forsyth, R. A. (1988). The comparative effects of compensatory and noncompensatory two-dimensional data on unidimensional IRT estimates. *Applied Psychological Measurement*, 12, 239-252.
- Zeng, L. (1989). *Robustness of unidimensional latent trait models when applied to multidimensional data*. Yayınlanmamış doktora tezi, Georgia Üniversitesi, Athens.
- Zhang, B. (2008). Application of unidimensional item response models to tests with items sensitive to secondary dimensions. *The Journal of Experimental Education*, 77 (2),147-166.
- Zhang, J. ve Stout, W. F. (1996a). Conditional covariance structure of generalized compensatory multidimensional items. *Psychometrika*, 64, 129-15.

Extended Abstract

In educational and psychological measurement several test theories have been used. One of the widely and primarily used test theories is Classical Test Theory. Classical test theory has been used for many decades by psychometrician. There are so many critiques related to test and sample dependency of Classical test theory parameter estimations. Item Response Theory claims to fill the lacknesses of classical test theory about test and sample dependent parameter estimation. Item response theory claims to estimate invariant item and person parameters independent from sample and test. There are 2 important and vital assumptions of item response theory. First assumption is unidimensionality and the second is local independency. Unidimensionality requires to predict the performance of an individual on an item only and only from a single latent trait. Local dependency requires the independence of responses that were given to an item for a fixed ability level. Local dependency assumption is vital for estimation procedures such as maximum likelihood estimation. Likelihood functions play role for maximum likelihood estimation and likelihood functions needs independent response probabilities. The violation of unidimensionality assumption will cause the violation of local dependency assumption. Unidimensionality is a prerequisite for local dependency assumption. The violation of local dependency will be a risk for independent parameter estimation. From this perspective; it will be claimed that multidimensionality will be a risk for invariant parameter estimation. Despite the fact that there are several studies about the calibration of multidimensional data by unidimensional models such as Drasgow and Parsons (1983), Asley and Forsyth (1985), Harrison (1986), Way, Ansley and Forsyth (1988), Reckase, Ackerman and Carlson (1988), Ackerman (1989), Zeng (1989), Luecht and Miller (1992), Kirisci, Hsu and Yu (2001), Zhang (2008), Levy (2006); no study seems available which focused on item parameter invariance for the multidimensional cases. There are several studies such as Gelbal (1994), Kelecioğlu (2001), Doğan (2002), Kan (2006), Kan (2004). which focused on parameter invariance for unidimensional data sets. This study will be a combination of two types of studies above. This study will also serve to detect convenience of unidimensional item response theory estimation for several factorial structures. For this purpose; this study investigates unidimensional item response theory item parameter mean standard errors which were obtained as an indicator of item parameter invariance for several factorial structures.

In this study, item parameter invariance in constructs which have characteristics of various dimensionality was examined within the framework of Unidimensional Item Response Theory (UIRT). For this purpose, two-dimensional data simulation was realized depending on compensatory MIRT model. For the simulation, the number of the items in the first dimension (24) and the mean of a_1 parameters (1.00) were kept constant in all constructs. Manipulations were applied to the items in the second dimension and correlation between dimensions. The numbers of the items in the second dimension were set to 4, 8, 12, 16, 20, 24; mean a_2 parameters were set to 0.25, 0.50, 0.75, 1.00, and correlations between dimensions were set to 0.00, 0.30, 0.60, 0.80. Simulation criteria were fully crossed and 96 experimental cells were obtained. Dichotomous responses of 90 000 subjects were simulated for each cell. By this way, 96 populations were established. In order to determine how parameter invariance is affected by sample size, 100 samples for 500 sample size, 100 samples for 1000 sample size and 100 samples for 3000 sample size were drawn from populations by using simple random sampling. Thus, 28800 samples were revealed. These samples were analyzed by using UIRT. In order to analyze the analysis results, average standard errors were used as the indicator of parameter invariance. Despite the fact that all datasets are treated as unidimensional for IRT analysis, average standard errors were computed separately for each dimension and how both dimensions have been affected by parameter invariance was determined in this manner. When the results for IRT were examined, parameter invariance was provided to a great extent for both a and b parameters but it cannot be provided for some constructs. Those constructs are empirical cells in which the first dimension is equivalent in terms of strength to the second dimension and in which the dimensions are orthogonal. Average standard errors which were computed for the particular empirical cells seem higher than the acceptable level for especially b parameter. Empirical cells can produce invariance item parameters in spite

*MADDE PARAMETRELERİNİN DEĞİŞMEZLİĞİNİN ÇEŞİTLİ BOYUTLULUK
ÖZELLİĞİ GÖSTEREN YAPILARDA MADDE TEPKİ KURAMINANA GÖRE
İNCELENMESİ*

of the violation of IRT's unidimensionality assumption. As shown in literature, UIRT is quite robust against multidimensionality.

To summarize the results for simulation criteria; item parameter invariance wasn't effected by the sample sizes which were used in the study. Item parameter invariance is effected from the numbers of items in the second dimension only when the number of items of second dimension is equal to the numbers of items in the first dimension. For the mean a2 parameters simulation criteria, it can be said that b parameters for the items in the first dimension were effected when the number of items in the second dimension is equal to 24 and the correlation between dimensions is equal to 0.00. In the second dimension, b parameter invariance is effected from each value of mean a2 parameter when the correlation between dimensions is 0.00. a parameter invariance is effected when the correlation between dimension is 0.00, mean a2 is equal to 1.00 and the number of items in the second dimension is equal to 24. For the correlation between dimensions criteria, b parameter invariance cannot be hold for 0.00 correlation. Paramater invariance for a parameter cannot be hold when the correlation between dimensions is 0.00, mean a2 is equal to 1.00 and the number of items in the second dimension is equal to 24.