Ankara University, Journal of Faculty of Educational Sciences, year: 2006, vol: 39, no: 1, 15-37

The Comparison of Reliability Coefficients in Parallel, Tau-Equivalent, and Congeneric Measurements

Halil Yurdugül*

ABSTRACT: The recent studies on reliability coefficients in composite tests and specially the biased of Cronbach α have become intense. If the items in composite tests have parallel measurements or congeneric measurements, then reliability coefficients are yielded differently. In the composite tests, consist of binary items, when the items have congeneric measurements KR20 is applied, otherwise KR21 is applied as reliability coefficient. Similarly, for the polythomous items with congeneric measurements McDonald's ω,, otherwise Cronbach's α is suggested to be used as reliability coefficient. In this study, first basic measurement types and reliability coefficients are defined, then Cronbach's α, Armor's θ, Heise and Bohrnstedt's Ω, Revelle's β, and McDonald's ω coefficient, in relation to different measurements are compared.

Key Words: Reliability Coefficients, Parallel Measure, Tau-Equivalent Measure, Essentially Tau- Equivalent Measure, Congeneric Measure.

SUMMARY

Introduction

The purpose of this study was to compare the coefficients of reliability in parallel, tau-equivalent, essentially tau-equivalent and congeneric

* Dr. Hacettepe University, Faculty of Education. yurdugul@hacettepe.edu.tr

measurements.. The coefficients of reliability were consisting of Cronbach's α, Armor's θ, Heise and Bohrnstedt's Ω, McDonald's ω and Revelle's β.

In classical testing theory, the item score, X_i can be represented as $X_i=T_i+E_i$, where T_i and E_i denote the true score and error score. The items are called *parallel* if any two items i and j have equal true scores, $T_i=T_j$, and equal error variances, $Var(E_i)=Var(E_j)$. With the assumption of error variances dropped, the items are called (a) *tau-equivalent*, if $T_i=T_j$ and $Var(E_i)\neq Var(E_j)$ (b) *essentially tau-equivalent*, if $T_i=T_j+a_{ij}$, where $a_{ij}\neq 0$ is a constant; and (c) *congeneric* if $T_i=b_{ij}T_j+a_{ij}$, where b_{ij} is a constant ($b_{ij}\neq 0$, $b_{ij}\neq 1$).

The Cronbach's α will be equal the real reliability of the test only if the items are parallel or tau-equivalent or essentially tau-equivalent, otherwise it gives a number that can be interpreted as a lower bound to real reliability (Traub, 1994).

$$\alpha = \frac{k}{k-1}\frac{\sum_{i\neq j}\sigma_{X_i X_j}}{\sum_{i,j}\sigma_{X_i X_j}}$$

Where k denotes the number of items in test and σ_{XiXj} denotes the covariance between items.

The Armor's theta (Armor, 1974), θ, is the alpha coefficient of weighted scale of k items. Theta can be considered as a maximized alpha. The θ is explained in terms of principal components analysis:

$$\theta = \frac{k}{k-1}\left[1-\frac{1}{\delta}\right]$$

where δ is the largest eigenvalue in principle component analysis (Carmines & Zeller, 1979).

Omega, Ω, introduced by Heise and Bohrnstedt (1970), estimates the reliability of all the common factors in items. Omega is based on the explaratory common factor analysis:

$$\Omega = 1-\frac{\sum\sigma^2{}_{X_i} - \sum\sigma^2{}_{X_i} h^2{}_{X_i}}{\sigma_{X_i X_j}}$$

where Ω is variance of the ith item; h^2 is communality of the ith item in common factor analysis.

McDonald's ω is also known as "congneric reliability" (Lucke, 2005a, 2005b; Zinbarg, Revelle, Yovel & Li, 2005) and is explained in terms of confirmatory factor analysis:

$$\omega = \frac{\left(\sum_{i=1}^{k}\lambda_i\right)^2}{\left(\sum_{i=1}^{k}\lambda_i\right)^2 + \sum_{i=1}^{k}\psi_i}$$

where λ_i is factor loadings of the i^{th} item and ψ_i is unique variance of the i^{th} item (McDonald, 1985, 1999).

Revelle proposed an index labeled coefficient beta (β) that he showed equals the proportion of variance in scale scores accounted for by a general factor under more general conditions than α.

$$\beta = k^2 \frac{\overline{\sigma}_{X_i X_j}}{\sigma^2{}_Y}$$

where $\overline{\sigma}_{X_i X_j}$ is the average covariance between items and $\sigma^2{}_Y$ is variance of total test (Zinbarg, Revelle, Yovel & Li, 2005).

Methods and Results

In this study, to compare the reliability coefficients, the data were used in Traub (1994). The α and the β were obtained from variance-covariance matrix and the θ was obtained from principal factor analysis, and the Ω and the β were obtained from the principal axis factoring method of common factor analysis and the ω was obtained from confirmatory factor analysis.

According to results, in parallel measurement, values of the reliability coefficients are equal ($\alpha=\theta=\beta=\Omega=\omega$). In tau-equivalent and essentially measurements, α, Ω and ω are equal and higher than θ and β ($\alpha=\Omega=\omega>\theta=\beta$). Finally, the results in congeneric measurements is $\alpha=\Omega=\theta=\beta<\omega$.

Discussion and Conclusion

The purpose of this study's is to make a simple comparison of reliability coefficients in parallel, tau-equivalent, essentially tau-equivalent and congeneric measurements. The data used for this study was taken from Traub (1994). Carmines and Zeller (1979) stated that α, θ and Ω are equal in parallel measurements. However, in this study, the three coefficients are also found to be equal in congeneric measurements as well. Finally, in congeneric measurement McDonald's ω has got the highest value.

Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, yıl: 2006, cilt: 39, sayı: 1, 15-37

Paralel, Eşdeğer ve Konjenerik Ölçmelerde Güvenirlik Katsayılarının Karşılaştırılması

Halil Yurdugül*

ÖZ: Son zamanlarda çoklu derecelenmiş testlerde kullanılan güvenirlik katsayıları üzerine ve özellikle Cronbach'ın α katsayısının yanlılığına yönelik çalışmalar yoğunluk kazanmıştır. Birleşik testlerde yer alan maddelerin paralel, eşdeğer ya da konjenerik olması durumuna göre güvenirlik katsayıları farklı sonuçlara ulaşmaktadır. İkili derecelenmiş testlerde güvenirlik katsayısı olarak, maddelerin konjenerik ölçmeler olduğu durumlarda KR20, maddelerin eşdeğer ölçmeler olduğu durumlarda ise KR21 katsayıları kullanılmaktadır. Benzer şekilde çoklu derecelenmiş testlerde maddeler konjenerik ölçme ise McDonald'ın ω, eşdeğer ölçme ise Cronbach'ın α katsayısı önerilmektedir. Bu çalışmada; ölçme yapıları ve güvenirlik için temel kavramlar ele alınmış ve değişik ölçme kümelerinde Cronbach'ın α, Armor'un θ, Heise ve Bohrnstedt'ın Ω, Revelle'nin β ve McDonald'ın ω güvenirlik katsayıları karşılaştırılmıştır.

Anahtar Kelimeler: Güvenirlik Katsayısı, Paralel Ölçme, Eşdeğer Ölçme, Eşbiçimli Ölçme, Konjenerik Ölçme

GİRİŞ

Eğitim alanında yapılan ölçmeler, öğrencilerin eğitim sürecinde ele alınan bilişsel, duyuşsal ve devinişsel özelliklere sahip olma düzeylerini belirlemeye yöneliktir. Bu özellikler gizil bir yapıda olduğu için doğrudan gözlenemezler ve öğrencilerin ölçme araçlarında yer alan maddelere verdikleri gözlenebilir yanıtlar/performanslar ile dolaylı olarak kestirilmeye

* Dr. Hacettepe Üniversitesi, Eğitim Fakültesi. yurdugul@hacettepe.edu.tr

çalışılır (Ghiselli, 1964). Ancak ölçme esnasında, ölçme sonuçlarına daima çeşitli hatalar karışır. Bu hatalardan özellikle rastgele (random) hatalar ölçme sonuçlarının güvenilirliği ile doğrudan ilişkilidir (Baykul, 2000).

Bir ölçme aracındaki herhangi bir maddeye verilen yanıtlar (X_i); bu maddenin ölçmeye çalıştığı özelliğe öğrencinin sahip olma düzeyi (T_i) ve hata terimleri (E_i) ile açıklanır. Bu bağıntı klasik test kuramının modeliyle ifade edilir.

$$X_i=T_i+E_i$$

Eğer i. madde için rasgele hatalardan arınık bir ölçme yapıldıysa bu durumda ölçme aracından elde edilen puan aynı zamanda öğrencinin ölçülmek istenilen özelliğe sahip olma düzeyine eşit ($X_i=T_i$) olacaktır. Bu durumda ölçme sonuçları tam güvenilir (perfect reliability) olarak tanımlanır.

Güvenirlik kavramı[1], ölçmelerdeki hatasızlığı tanımlamak için kullanılır. *Güvenirlik indeksi* ise güvenirliğin nicel büyüklüğünü ifade eder ve klasik test modelindeki terimlerle ifade edilir. Buna göre güvenirlik indeksi, gerçek puanlar varyansının gözlenen puanlar varyansına oranı olarak formüle edilir. Ancak gerçek puanlar (T_i) doğrudan gözlenemediğinden dolayı klasik test kuramının ve güvenirlik indeksinin "platonik" bir yapısı vardır (Lord ve Novick, 1968). Bu nedenden dolayı ölçme sonuçlarının güvenirliklerini elde edebilmek için aynı özelliği ölçmeye yönelmiş ek ölçmelere/maddelere ihtiyaç duyulur. Eğitim alanında ek ölçmeler ise genellikle birleşik testler (composite test) şeklindedir. Birleşik testlerde, ölçme aracının güvenirlik indeksi, *güvenirlik katsayıları* olarak adlandırılan bir takım formüller sayesinde kestirilebilmektedir (Baykul, 2000).

Günümüze kadar ölçme araçlarının güvenirliğini kestirmeye yönelik çok sayıda güvenirlik katsayısı önerilmiştir. Güvenirlik katsayılarının sayıca çok olmasının temel nedeni; bu katsayıların farklı madde yapılarında farklı değer üretmeleridir (Osborn, 2000). Buna rağmen, eğitim ve psikoloji alanında yapılan çalışmalar incelendiğinde; çoklu derecelenmiş (polythomous) testlerin güvenirliğinin elde edilmesinde yaygın olarak Guttman ve Cronbach tarafından geliştirilen alfa (α) katsayısının, ikili derecelenmiş (dichothomous) testlerde ise Kuder ve Richarson tarafından geliştirilen KR-20 ve KR-21 katsayılarının kullanıldığı gözlenmektedir.

[1] Bu çalışmada güvenirlik kavramı olarak "iç tutarlılık" anlamındaki güvenirlik ele alınmıştır.

Diğer taraftan; α katsayısı, birleşik testlerde yer alan maddelerin paralel, eşdeğer (tau-equavalent) ya da eşbiçimli (essentially tau-equavalent) ölçmeler olduğu durumlarda gerçek güvenirliği yansız olarak kestirmektedir. Test maddelerinin konjenerik (congeneric) ölçmeler olduğu durumlarda ise α katsayısı yanlı sonuçlar vermektedir (Alwin, 1976; Bacon vd., 1995, DeVellis, 2003; Feldt & Qualls, 1996; Jöreskog, 1971; Lucke, 2005a, 2005b; Miller, 1995; Raykov, 1977a, 1977b, 2001; Traub, 1994; Zinbarg vd., 2005). Daha açık bir ifadeyle; ölçme sonuçlarının çözümlemesinde kullanılan faktör analizinde, maddelere ilişkin faktör yükleri eşit ise bu tür maddeler paralel, eşdeğer ve/veya eşbiçimli madde olarak adlandırılmaktadır. Bu maddeler üzerinden elde edilen α güvenirlik katsayısı ise gerçek güvenirliği vermektedir. Ancak maddelere ilişkin faktör yükleri eşit değilse bu tür maddeler konjenerik madde olarak adlandırılmaktadır ve bu durumda α güvenirlik katsayısı gerçek güvenirliğin altında değerler üretmektedir (Lucke, 2005a).

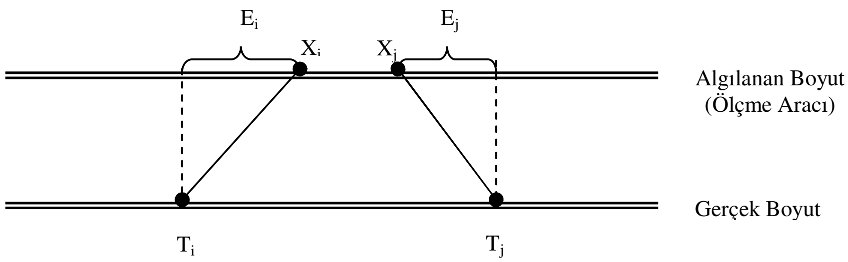
Bu çalışmanın amacı; çoklu derecelendirilmiş ölçme araçlarında kullanılmak üzere geliştirilmiş güvenirlik katsayılarının paralel, eşdeğer, eşbiçimli ya da konjenerik ölçmelerdeki davranışlarını incelemeye yöneliktir. Karşılaştırmalarda kullanılmak üzere ele alınan güvenirlik katsayıları Cronbach'ın α, Revelle'nin β, Armor'un θ, Heise ve Bohrnstedt'ın Ω ve McDonald'ın ω güvenirlik katsayılarıdır.

Klasik Test Kuramı ve Ölçmelerin Yapısı

Eğitim ve psikoloji alanında yapılan ölçmeler öğrencilerin doğrudan gözlenemeyen özelliklerini kapsadığından dolayı psikofizik uzayda tanımlıdır. Psikofizik uzayda yapılan ölçmeler ise iki boyutta ele alınır. Bunlardan ilki; ölçülmeye çalışılan özelliğin "gerçek" boyutu[2], diğeri ise öğrencinin gözlenebilen davranışlarının ölçmeci tarafından "algılanan" boyutudur. Örneğin i ve j gibi iki ölçme maddesi ele alındığında; öğrencinin ölçülmek istenilen özelliğe ilişkin gerçek puanı ile gözlenen performansı arasındaki farklılık ($E_i=X_i-T_i$) hata olarak adlandırılır.

---

[2] Ölçülmeye çalışılan özelliğin öğrencide bulunuş düzeyi "gerçek puan" (true score) olarak adlandırılmıştır.



Çizim 1: Psikofizik ölçmelerde tanımlı boyutlar ve ölçme bileşenleri

Çizim 1'e göre; ölçme aracında yer alan her bir ölçme/madde için klasik test modeli (Lord ve Novick, 1968);

$$X_i=T_i+E_i \tag{1}$$

$$X_j=T_j+E_j \tag{2}$$

şeklinde ifade edilebilir. DeVellis (2003)'in ifade ettiği gibi; klasik test kuramı ve bu kurama ilişkin model paralel ölçmeler üzerine kuruludur. Buna göre, Eşitlik 1 ve Eşitlik 2 ile verilen durumlarda $T_i=T_j$ ve $E_i=E_j$ ise bu iki ölçme için *paralel ölçme* ifadesi kullanılır. Bu ölçmelere ilişkin ilk çalışma Spearman (1904) tarafından ele alınmıştır. Buna göre her iki madde; aynı özelliği eşit büyüklükte ve eşit duyarlıkta (hata terimleri varyansının eşit olma durumu) ölçmektedirler. Ancak Novick ve Lewis (1967) $T_i=T_j$ ve $E_i \neq E_j$ olduğu durumları ise *eşdeğer ölçmeler* (tau-equivalent measure) olarak adlandırmışlardır. Eşdeğer ölçmelerin paralel ölçmelerde olduğu gibi en önemli özellikleri ise gözlenen puanlar arasındaki kovaryansların gerçek puanlar varyansına eşit olmasıdır (McDonald, 1999).

$$\sigma_{X_i X_j} = \sigma^2_{T_i} = \sigma^2_{T_j}$$

Eşdeğer ölçmelerde her bir madde aynı özelliği eşit büyüklükte ölçmektedir. Eğer ölçülmek istenilen gerçek puanlar arasındaki sabit bir bağıntı var ise;

$$T_i - T_j = a_{ij} \tag{3}$$

bu durumda her iki ölçme için *eşbiçimli ölçme* (essentially tau-equivalent) adlandırılması yapılmaktadır (Traub, 1994).

Birleşik testlerdeki eşbiçimli ölçmelerde söz konusu edilen gerçek puanlar arasındaki nicel farklılık sabit bir farklılığı ötesinde, değişken bir farklılık ise bu durumda gerçek puanlar arasında fonksiyonel bir bağıntıdan söz edilebilir.

$T_i=a_{ij}+b_{ij}T_j$ (4)

Jöreskog (1971), bu bağıntıya dayalı olarak i. ve j. maddeler için konjenerik ölçme tanımlamasını yapmıştır. Ancak konjenerik ölçmeler ile paralel, eşdeğer ve eşbiçimli ölçmeler arasındaki en önemli farklılık gözlenen puanlar arasındaki kovaryanslardan kaynaklanmaktadır. Şöyle ki; paralel, eşdeğer ve eşbiçimli ölçmelerde maddeler arasındaki kovaryanslar eşit iken konjenerik ölçmelerde bu eşitlik bozulmaktadır.

Paralel, eşdeğer, eşbiçimli ve konjenerik ölçmeler arasındaki ilişkiler 3 maddeden oluşan birleşik ölçmedeki simgesel gösterimleri Tablo 1'de verilmiştir.

Tablo 1: Birleşik testi oluşturan ölçmelerin/maddelerin yapısı

Ölçmeler	**Varsayımlar**	**Eşitlikler**	**Açıklamalar**
Paralel	$T_1=T_2=T_3$ $\sigma^2_{E_1}=\sigma^2_{E_2}=\sigma^2_{E_3}$	$\overline{X}_1=\overline{X}_2=\overline{X}_3$ $\sigma_{X_1}=\sigma_{X_2}=\sigma_{X3}$ $\sigma_{X_1X_2}=\sigma_{X_2X_3}$	Her üç madenin ortalamaları, standart sapmaları ve kovaryansları eşit. (Faktör yükleri eşit.)
Eşdeğer	$T_1=T_2=T_3$ $\sigma^2_{E_1}\neq\sigma^2_{E_2}\neq\sigma^2_{E_3}$	$\overline{X}_1=\overline{X}_2=\overline{X}_3$ $\sigma_{X_1}\neq\sigma_{X_2}\neq\sigma_{X3}$ $\sigma_{X_1X_2}=\sigma_{X_2X_3}$	Maddelere ilişkin ortalamalar eşit ve standart sapmalar farklı, kovaryanslar eşit. (Faktör yükleri eşit.)
Eşbiçimli	$T_1=T_2+a_{12}$ $T_1=T_3+a_{13}$ $T_2=T_3+a_{23}$ $\sigma^2_{E_1}\neq\sigma^2_{E_2}\neq\sigma^2_{E_3}$	$\overline{X}_1\neq\overline{X}_2\neq\overline{X}_3$ $\sigma_{X_1}\neq\sigma_{X_2}\neq\sigma_{X3}$ $\sigma_{X_1X_2}=\sigma_{X_2X_3}$	Maddelere ilişkin ortalamalar ve standart sapmalar farklı, kovaryanslar eşit. (Faktör yükleri eşit.)
Konjenerik	$T_1=b_{12}T_2+a_{12}$ $T_1= b_{13}T_3+a_{13}$ $T_2= b_{23}T_3+a_{23}$ $\sigma^2_{E_1}\neq\sigma^2_{E_2}\neq\sigma^2_{E_3}$	$\overline{X}_1\neq\overline{X}_2\neq\overline{X}_3$ $\sigma_{X_1}\neq\sigma_{X_2}\neq\sigma_{X3}$ $\sigma_{X_1X_2}\neq\sigma_{X_2X_3}$	Maddelere ilişkin ortalamalar, standart sapmalar ve kovaryanslar farklı. (Faktör yükleri eşit değil.)

Tablo 1'den de görülebileceği gibi birleşik testi oluşturan maddelerin kovaryansları farklı ise bu durumda konjenerik ölçmelerden söz edilebilmektedir. Bu farklılığı aynı zamanda faktör analizi sonuçlarında görmekte olanaklıdır. Faktör analizinde standartlaştırılmamış faktör yükleri aynı zamanda maddeler arası kovaryansın bir fonksiyonu olduğu için, faktör yükleri eşit olmayan maddeler aynı zamanda konjenerik ölçme olarak nitelendirilmektedir (Lucke, 2005a, McDonald, 1999).

Faktör Analizi ve Ölçmelerin Yapısı

Klasik test kuramının çoklu biçimi faktör analizi ile ifade edilebilir (Lord & Novick, 1968). Buna göre klasik test kuramına karşılık gelen faktör analitik modeli;

$$X_i = \mu_i + \lambda_i F + E_i \qquad (5)$$

$$X_i - \mu_i = \lambda_i F + E_i$$

$$Z_i = \lambda_i F + E_i \qquad (6)$$

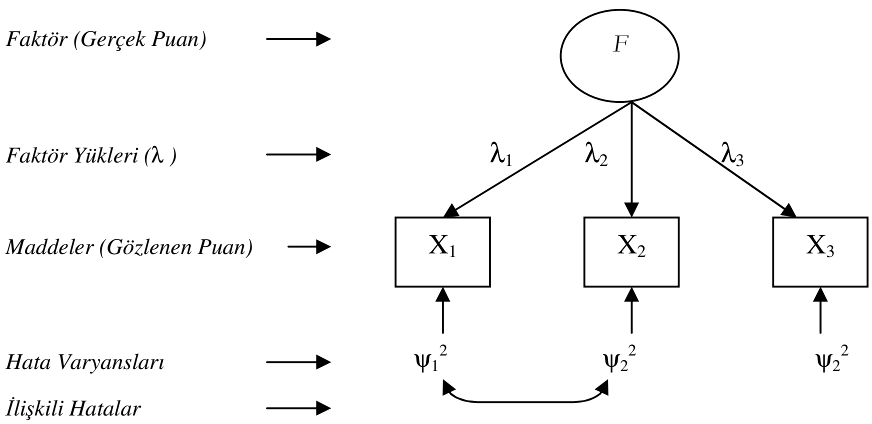
şeklindedir. Burada μ_i i. madde ortalamasını, λ_i ise i. maddenin faktör yükünü ve F ise ölçülmek istenilen gerçek puanlara ilişkin faktör puanlarını göstermektedir (McDonald, 1999; Nunnally & Bernstein 1994).

Eşitlik 6, standartlaştırılmış faktör analitik modelini göstermektedir. Bu model temel bileşenler analizi (principal components analysis) ya da ortak faktör analizi (common factor analysis) yöntemleri kullanılarak kestirilebilir (Nunnally & Bernstein, 1994). Eşitlik 5 ise standartlaştırılmamış faktör çözümleri üretir ve genellikle doğrulayıcı faktör analizi (confirmatory factor analysis) yöntemleri kullanılarak elde edilir (Comrey ve Lee, 1992; McDonald, 1985; Nunnally & Bernstein, 1994). Eşitlik 1 ile verilen klasik test modelini ve Eşitlik 5 ve Eşitlik 6 ile verilen faktör analitik modelleri arasındaki bağıntılar Lord ve Novick (1968) tarafından ayrıntılı bir şekilde verildiği için burada yer verilmeyecektir. Ancak bazı önemli bağıntılar aşağıda verilmiştir. Özellikle kovaryans terimlerini standartlaştırılmamış faktör çözümlerinden elde edebilmek için;

$$\sigma_{X_i X_j} = \lambda_i \lambda_j \qquad (7)$$

$$\sigma^2_{X_i} = \lambda_i^{\,2} + \psi_i^{\,2} \qquad (8)$$

bağıntısı kullanılabilir (McDonald, 1985, 1999). Burada ψ_i^2 tekil varyansı (unique variance) göstermektedir.



Çizim 2: Faktör analitik modeli

Maddelerin yapısına ilişkin özellikler faktör analitik modelinin terimleri ile de ifade edilebilir. Faktör analitik modeline göre birleşik testlerde yer alan k adet madde için $\lambda_1=\lambda_2=...=\lambda_k$ ve $\psi_1^2=\psi_2^2=...=\psi_k^2$ eşitlikleri söz konusu olduğunda bu maddeler *paralel ölçme*, olarak adlandırılmaktadır. Benzer şekilde $\lambda_1=\lambda_2=...=\lambda_k$ ve $\psi_1^2\neq\psi_2^2\neq...\neq\psi_k^2$ ise k madde için eşdeğer ya da eşbiçimli ölçme[3] ve $\lambda_1\neq\lambda_2\neq...\neq\lambda_k$ ve $\psi_1^2\neq\psi_2^2\neq...\neq\psi_k^2$ olduğunda ise bu tür maddeler için konjenerik ölçme olduğu ifade edilir (Jöreskog, 1971, Lucke, 2005a; McDonald, 1999). Bu açıklamalara ilişkin faktör çözümlemelerinin şematik gösterimleri Ek 1'de verilmiştir.

Klasik Güvenirlik Katsayıları

Farklı madde yapılarına dayalı ölçme kümelerinden elde edilen güvenirlik katsayıları da farklı sonuçlar vermektedir. Bu durum, ölçme literatüründe güvenirlik katsayılarının bu denli çok olmasına da neden olmuştur (Osborn, 2000). Çalışmanın bu kesiminde birleşik ölçmelerde güvenirlik indeksinin kestiricisi niteliğinde olan ve bu çalışmanın kapsamında yer alan bazı güvenirlik katsayıları verilmiştir.

[3] Faktör modelindeki eşdeğer ve eşbiçimli ölçme ayırımı için özvektörlere göre karar verilmektedir (Lucke, 2005a). Çalışmada bu konuda ayrıntıya girilmeyecektir.

Guttman ve Cronbach α Katsayısı

Klasik test kuramı paralel ölçmeler üzerine kurulu (DeVellis, 2003) olduğu için bu tür ölçme araçlarından elde edilen güvenirlik katsayıları gerçek güvenirliği vermektedir. Guttman (1945) paralel ölçmelerin dışındaki tüm ölçme kümeleri için güvenirlik katsayıları gerçek güvenirliğin altında değer üreteceğinden dolayı "güvenirliğin alt sınırı" olarak altı adet (Γ_1, Γ_2, Γ_3, Γ_4, Γ_5, Γ_6) katsayı önermiştir. Guttman (1945) tarafından önerilen katsayısılar içerisinden Γ_3 katsayısı aynı zamanda Cronbach'ın α güvenirlik katsayısı (Cronbach, 1951) olarak bilinir. Cronbach (1951) ise, paralel maddelerin yanı sıra eşdeğer maddeler için de α güvenirlik katsayısının gerçek güvenirliği vereceğini ifade etmiştir. Novick ve Lewis (1967), α güvenirlik katsayısının, eşbiçimli ölçme yapısına sahip maddeler içinde gerçek güvenirliği vereceğini rapor etmişlerdir.

Cronbach α güvenirlik katsayısı, Eşitlik 9'da verildiği gibi maddelere ilişkin kovaryans terimleri üzerine kuruludur.

$$\alpha = \frac{k}{k-1} \frac{\sum_{i \neq j} \sigma_{X_i X_j}}{\sum_{i,j} \sigma_{X_i X_j}} \tag{9}$$

Burada k madde sayısını, $\sum_{i \neq j} \sigma_{X_i X_j}$ ifadesi kovaryans matrisinin köşegen dışı elemanlarının toplamını ve $\sum_{i,j} \sigma_{X_i X_j}$ ifadesi ise kovaryans matrisinin tüm elemanlarının toplamını göstermektedir (Christmann & Aelts, 2005). Cronbach'ın α katsayısı aynı zamanda doğrulayıcı faktör analizi terimleri ile de ifade edilebilir (McDonald, 1985).

$$\alpha = \frac{k}{k-1} \left[\frac{k(\overline{\lambda})^2 - (\overline{\lambda^2})}{k(\overline{\lambda})^2 + \overline{\psi}^2} \right]$$

Burada; $(\overline{\lambda})^2$, faktör yüklerinin ortalamasının karesini, $(\overline{\lambda^2})$ ise faktör yüklerinin kareler ortalamasını göstermektedir. Cronbach'ın α katsayısı maddelere ilişkin kovaryansların eşit olduğu (paralel, eşdeğer ya da eşbiçimli) durumlarda güvenirlik indeksini vermektedir. Ancak α katsayısı konjenerik ölçmelerde gerçek güvenirliğin altında bir değer üretmektedir (Alwin, 1976, Feldt & Qualls, 1996; Lucke, 2005a, 2005b; Raykov, 1997a, 1997b, 2001; Traub, 1994; Zinbarg, 2005). Ancak Komaroff (1997), Raykov (1997b, 2001) Rae (2006) ve Zimmerman vd. (1993) ise maddelere ilişkin hata terimlerinin negatif korelasyona sahip olduğunda & katsayısının gerçek güvenirliğin altında değerler ürettiğini; hata terimlerinin pozitif korelasyona sahip olduğunda ise gerçek güvenirliğin üzerine değerler ürettiğini ifade etmişlerdir.

Armor'un θ Katsayısı

Armor (1974) tarafından önerilen θ katsayısı temel bileşenler analizinden elde edilen en yüksek özdeğer (eigenvalue) ile elde edilmektedir (Armor, 1974; Carmines & Zeller, 1979).

$$\theta = \frac{k}{k-1}\left[1-\frac{1}{\delta}\right] \tag{10}$$

Burada δ terimi, birleşik test maddelerinden temel bileşenler analizi ile elde edilen en büyük özdeğeri göstermektedir. Armor (1974), Comrey ve Lee (1992) ve Cortina (1993), güvenirliğin aynı zamanda Spearman tarafından geliştirilen faktör modelindeki "genel faktörü" ifade ettiğini ve bu nedenle güvenirliğin en büyük özdeğerin bir fonksiyonu olduğunu ifade etmişlerdir.

Heise ve Bohrnstedt'ın Ω Katsayısı

Birleşik ölçmelerde güvenirlik indeksinin bir diğer kestiricisi ise Heise ve Bohrnstedt (1970) tarafından önerilen Ω katsayısıdır. Ω katsayısı ortak faktör analizi sonuçlarından elde edilir.

$$\Omega = 1 - \frac{\sum \sigma^2{}_{X_i} - \sum \sigma^2{}_{X_i} h^2{}_{X_i}}{\sigma_{X_i X_j}} \tag{11}$$

Burada, h_i^2, i. maddenin ortak varyansını (communality) göstermektedir. Ω katsayısı, ortak faktör analiz (common factor analysis) yöntemleri ile hesaplanabilir (Carmines ve Zeller, 1979; Heise & Bohrnstedt, 1970). Bu çalışmada Ω katsayısındaki ortak varyansları elde edebilmek için temel faktör analizi (principal factoring analysis) kullanılmıştır. Heise, güvenirlik katsayısının ortak varyansları kullanması nedeniyle elde edilen Ω değerlerinin aynı zamanda ölçme aracının geçerliğininde bir göstergesi olduğunu ifade etmektedir (David Heise ile kişisel iletişim).

Konjenerik Güvenirlik Katsayıları

Klasik güvenirlik katsayılarına, özellikle α katsayısına yönelik eleştirilerin temelinde, konjenerik ölçmelerin varlığı durumunda α katsayısının yanlı sonuçlar üretmesi yatmaktadır. Bu yaklaşım ile klasik test kuramının bir alt kuramı olan konjenerik test kuramı (congeneric testing theory) geliştirilmiştir (Jöreskog, 1971; Lucke, 2005a; McDonald, 1999; Yurdugül, 2005).

Konjenerik test kuramı; daha öncede bahsedildiği gibi, birleşik testte ($Y=X_1+...+X_k$) yer alan maddelerin;

$$X_1=T_1+E_1$$

$$X_2=T_2+E_2$$

$$: \quad : \quad :$$

$$X_k=T_k+E_k$$

ölçmeye yöneldiği özelliği aynı nicel büyüklükte ölçmediği durumlarda, gerçek puanlar arasındaki fonksiyonel bir bağıntının varlığı ile açıklanır (Jöreskog, 1971). Bu model standartlaştırılmamış faktör analitik modelinin (Eşitlik 5) terimleri ile ifade edilirse;

$$\tau = \nu + \lambda\eta \tag{12}$$

şeklindedir. Burada τ gerçek puanı, ν, madde kolaylığını (madde güçlüğünün tersi), λ, madde ayırıcılık gücünü (aynı zamanda faktör yüklerini) ve η ise maddenin ölçmeye yöneldiği özelliği göstermektedir (Lucke, 2005a; McDonald, 1999).

Tek boyutlu testlerde maddelere ilişkin gerçek puanların (T_i) tek bir özelliği ölçmeye yöneldiği durumlarda, konjenerik test modeline ilişkin ifade Eşitlik 13'te verildiği gibidir.

$$\begin{pmatrix} T_1 \\ T_2 \\ : \\ T_k \end{pmatrix} = \nu + \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ : \\ \lambda_k \end{pmatrix} \eta \tag{13}$$

Tüm maddelere ilişkin ayırıcılık güçleri (faktör yükleri) eşit olduğunda ($\lambda_1=\lambda_2=...=\lambda_k$) bu durum maddelerarası kovaryansın eşitliği ile açıklanmaktadır. Diğer taraftan madde ayırıcılık güçleri eşit olmadığında ($\lambda_1\neq\lambda_2\neq...\neq\lambda_k$) ise maddelerin konjenerik ölçmeler olduğu ifade edilebilir. Buna göre faktör yükleri eşit ise bu tür maddeler eşbiçimli aksi halde konjenerik madde olarak nitelendirilir (Lucke, 2005a; McDonald, 1999).

Çok boyutlu testlerde ise konjenerik ölçmelere ilişkin model Eşitlik 14'te verilmiştir.

$$\begin{pmatrix} T_1 \\ T_2 \\ : \\ T_k \end{pmatrix} = \nu + \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ : \\ \lambda_k \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \\ : \\ \eta_m \end{pmatrix} \quad (14)$$

Eşitlik 14'te verilen *m* ifadesi, ölçme aracında yer alan alt boyutları göstermektedir. Bu çalışmada çok boyutlu ölçme kümeleri kapsam dışı bırakılmış, yalnızca tek boyutlu ölçmeler çalışma kapsamına alınmıştır.

McDonald'ın ω Katsayısı

McDonald tarafından geliştirilen ω katsayısı özellikle konjenerik ölçmeler için tasarlanmıştır ve standartlaştırılmamış faktör analizi terimleri ile ifade edilmektedir (McDonald, 1985).

$$\omega = \frac{\left(\sum_{i=1}^{k} \lambda_i\right)^2}{\left(\sum_{i=1}^{k} \lambda_i\right)^2 + \sum_{i=1}^{k} \psi_i} \quad (15)$$

McDonald'ın ω katsayısı aynı zamanda "yapısal güvenirlik" (construct reliability) olarak adlandırılır (Nunnally & Bernstein, 1994) ve doğrulayıcı faktör analizi yöntemiyle elde edilir. Diğer taraftan ω tüm ölçmelerde α'ya eşit ya da daha büyük çıkmaktadır (Bacon vd, 1995). Bununla birlikte ω katsayısının elde edilmesinde hiyerarşik faktör analizi de kullanılabilmektedir (Zinbarg vd, 2005).

Revelle'nin β Katsayısı

Revelle tarafından önerilen β katsayısı maddelerin kovaryanslarının ortalamaları üzerine kuruludur.

$$\beta = k^2 \frac{\overline{\sigma}_{X_i X_j}}{{\sigma_Y}^2} \quad (16)$$

Burada $\overline{\sigma}_{X_i X_j}$ kovaryans matrisinin ortalamasını ve ${\sigma_Y}^2$ ise toplam test varyansını göstermektedir (Zinbarg vd., 2005).

Bu çalışmada, inceleme kapsamında ele alınan güvenirlik katsayılarının farklı ölçme yapılarında aldığı değerlerin karşılaştırılması amaçlanmıştır.

YÖNTEM

Bu çalışmanın uygulama kısmında, Traub (1994) tarafından ele alınan paralel, eşdeğer, eşbiçimli ve konjenerik veriler üzerinden, çalışma kapsamına alınan güvenirlik katsayıları karşılaştırılmıştır. Bu katsayılar; Cronbach'ın α, Armor'un θ, Heise ve Bohrnstedt'ın Ω, McDonald'ın ω ve Revelle'nin β katsayılarıdır.

Tablo 2: Birleşik testteki ölçmelere ilişkin sayısal veriler*

	Ölçmeler	Madde Ortalamaları	Madde Varyansları	Madde Güvenirlik**	Kovaryans/Korelasyon Matrisi***		
Paralel	X_1	100	225	0,92	-	207	207
	X_2	100	225	0,92	0,92	-	207
	X_3	100	225	0,92	0,92	0,92	-
Eşdeğer	X_1	100	225	0,92	-	207	207
	X_2	100	256	0,81	0,86	-	207
	X_3	100	289	0,72	0,81	0,76	-
Eşbiçimli	X_1	100	225	0,92	-	207	207
	X_2	95	256	0,81	0,86	-	207
	X_3	105	289	0,72	0,81	0,76	-
Konjenerik	X_1	100	225	0,92	-	172,50	230,00
	X_2	79,17	192,75	0,75	0,83	-	191,67
	X_3	116,67	336,56	0,76	0,84	0,75	-

* Veriler Traub, R. E. (1994), "*Reliability for the social sciences: Theory and Applications*" kaynağından alınmıştır.

** Güvenirlikler maddenin ölçmeye çalıştığı gerçek puan varyansı (207) ile gözlenen varyans (225) oranında elde edilmiştir (Traub, 1994).

*** Kovaryans-Korelasyon matrisinin alt üçgeni korelasyon değerlerinden (köşegen değerleri 1) ve üst üçgeni ise kovaryans değerlerinden (köşegen değerleri maddelerin varyansından) oluşmaktadır.

Çalışmada kullanılan veriler Tablo 2'de verilmiştir. Bu verilerin faktör çözümlemeleri için LISREL ve SPSS paket programlarından yararlanılmıştır[4].

BULGULAR

Verilere ilişkin standartlaştırılmış ve standartlaştırılmamış faktör çözümleri Tablo 3'te verilmiştir. Faktör çözümlerinden görüleceği gibi; standartlaştırılmamış faktör çözümleri için maddeler arası kovaryansların eşit olmasından dolayı paralel, eşdeğer ve eşbiçimli ölçmelerde faktör yükleri aynı elde edilirken konjenerik ölçmelerde faktör yükleri farklıdır. Bu faktör çözümlerine ilişkin elde edilen güvenirlik katsayıları Tablo 4'te verilmiştir.

Bu sonuçlara göre; paralel ölçmelerde tüm katsayılar eşit olarak elde edilmiştir ($\alpha=\theta=\beta=\Omega=\omega$). Ancak eşdeğer ve eşbiçimli ölçmelerde α, Ω ve ω katsayıları eşit iken θ ve β katsayıları diğer güvenirlik katsayılarına göre

[4] Çalışmada kullanılan LISREL ve SPSS betikleri (script) ve kullanılan faktör modelleri http://yunus.hacettepe.edu.tr/~yurdugul/AU_EBF_guvenirlik.html adresinde sunulmuştur.

daha düşük elde edilmiştir ($\alpha=\Omega=\omega>\theta=\beta$). Konjenerik ölçmelerde ise en yüksek değer olarak ω katsayısı elde edilirken bu ölçme kümesinde diğer katsayılar eşit olarak elde edilmiştir ($\alpha=\Omega=\theta=\beta<\omega$).

Tablo 3: Ölçmelere ilişkin faktör çözümleri

	Standartlaştırılmış Faktör Çözümleri (Açıklayıcı Faktör Çözümleri**)				Standartlaştırılmamış Faktör Çözümleri (Doğrulayıcı Faktör Çözümleri)			
	Paralel λ_i* (ψ_i^2) [$\rho_i=\lambda_i^2$]	**Eşdeğer** λ_i (ψ_i^2) [$\rho_i=\lambda_i^2$]	**Eşbiçimli** λ_i (ψ_i^2) [$\rho_i=\lambda_i^2$]	**Konjenerik** λ_i (ψ_i^2) [$\rho_i=\lambda_i^2$]	**Paralel** λ_i (ψ_i^2) [$\rho_i=\lambda_i^2$]	**Eşdeğer** λ_i (ψ_i^2) [$\rho_i=\lambda_i^2$]	**Eşbiçimli** λ_i (ψ_i^2) [$\rho_i=\lambda_i^2$]	**Konjenerik** λ_i (ψ_i^2) [$\rho_i=\lambda_i^2$]
$\mathbf{X_1}$	**0,96** (0,08) [0,92]	**0,96** (0,08) [0,92]	**0,96** (0,08) [0,92]	**0,96** (0,08) [0,92]	**14,39** (18,00) [0,92]	**14,39** (18,00) [0,92]	**14,39** (18,00) [0,92]	**14,39** (18,00) [0,92]
$\mathbf{X_2}$	**0,96** (0,08) [0,92]	**0,90** (0,19) [0,81]	**0,90** (0,19) [0,81]	**0,86** (0,26) [0,72]	**14,39** (18,00) [0,92]	**14,39** (49,00) [0,81]	**14,39** (49,00) [0,81]	**11,99** (50,00) [0,75]
$\mathbf{X_3}$	**0,96** (0,08) [0,92]	**0,85** (0,08) [0,72]	**0,85** (0,08) [0,72]	**0,87** (0,24) [0,76]	**14,39** (18,00) [0,92]	**14,39** (82,00) [0,72]	**14,39** (82,00) [0,72]	**15,99** (81,00) [0,76]
δ^+	2,840	2,621	2,621	2,607				

* Burada λ_i, i. maddenin faktör yükünü, ψ_i^2, i. maddenin tekil varyansını ve ρ_i ise i. maddenin güvenirliğini göstermektedir. Madde güvenirlikleri, maddenin faktör yüklerinin karesinden elde edilmiştir.

** Açıklayıcı faktör çözümleri için temel eksen faktör (principal axis factoring) yöntemi kullanılmıştır.

+ δ; temel bileşenler analizi ile elde edilen en büyük özdeğerleri göstermektedir.

Tablo 3'teki Traub (1994) tarafından verilen yapay veri kümeleri üzerinden gerçekleştirilen çözümler göz önüne alındığında; standartlaştırılmamış (kovaryans matrisinin çözümlenmesiyle elde edilen) faktör yükleri paralel, eşdeğer ve eşbiçimli madde yapılarını içeren veri kümelerinde eşit değerler üretmiştir. Ancak konjenerik madde yapısına sahip maddelerin faktör yükleri farklı çıkmıştır. Tablo 4'te ise bu yapay veri kümelerinden elde edilen güvenirlik katsayıları verilmiştir.

Tablo 4: Ölçmelerde elde edilen güvenirlik katsayıları

	Paralel	**Eşdeğer**	**Eşbiçimli**	**Konjenerik**
α	0,97	0,97	0,97	0,92
θ	0,97	0,93	0,93	0,92
Ω	0,97	0,97	0,97	0,92
ω	0,97	0,97	0,97	0,93
β	0,97	0,93	0,93	0,92

Güvenirlik katsayıları elde edilirken α ve β için varyans-kovaryans matrisinden, ω katsayısı için doğrulayıcı faktör analizi sonuçlarından, Ω için temel faktör analizinden ve θ katsayısı içinde temel bileşenler analizinden yararlanılmıştır.

Çalışma kapsamına alının tüm güvenirlik katsayıları, paralel maddeler için eşit değerler almıştır. Eşdeğer ve eşbiçimli maddeler için θ ve β güvenirlik katsayıları diğer güvenirlik katsayılarına göre daha düşük değerlere sahiptir. Güvenirlik literatüründe bu iki katsayı α güvenirlik katsayısından daha düşük elde edilemeyeceği yönünde bilgiler bulunmaktadır (Carmines & Zeller. 1979; Zinbarg vd., 2005). Bu durumun çalışmada kullanılan verilerin çok az sayıda madde içermesinden kaynaklandığı düşünülmektedir. Faktör başına düşen madde sayısı MacCallum, Widaman, Zhang ve Hong (1999) tarafından "aşırı belirlenme" (overdetermination) ifadesi ile kavramlaştırmışlardır ve madde sayısının az olması ortak varyansı gereğinden düşük kestirdiği ifade etmişlerdir.

Konjenerik ölçmeler için en yüksek değeri ω katsayısı üretmiştir. Konjenerik ölçmelerde hata terimleri ilişkili olmasından dolayı maddelere ilişkin standartlaştırılmamış faktör yükleri farklı değerler almaktadır ve bu nedenle klasik güvenirlik katsayıları gerçek güvenirlikten daha düşük değerler üretmektedir. Bu durum güvenirlik konusundaki diğer araştırmalar tarafından desteklenmektedir (Green & Hershberger, 2000; Komaroff, 1997; Lucke, 2005a, 2005b; Miller, 1995; Osburn, 2000; Raykov, 1997a, 1997b, 2001).

SONUÇ VE YORUMLAR

Ölçme literatüründe; ikili derecelenmiş (binary/dichotomous) testlerde maddeler paralel, eşdeğer ya da eşbiçimli yapıya sahip; bir diğer ifade ile maddeler arası varyans-kovaryanslar terimleri birbirine eşit ise KR-21, aksi taktirde KR-20 güvenirlik katsayısı önerilir (McDonald, 1999). Benzer bir durum çoklu derecelenmiş (polythomous) testler içinde söz konusudur. Testteki maddeler paralel ölçmeler olduğu sürece güvenirlik katsayıları aynı değerler üretirken ($\alpha=\theta=\Omega=\omega=\beta$) konjenerik ölçmeler de ($\alpha=\theta=\Omega=\beta<\omega$) eşitsizliği ortaya çıkmaktadır. Bu durum, ele alınan güvenirlik katsayılarının yanlığından kaynaklamaktadır. Özellikle α ile ω arasındaki yanlılık miktarı;

$$\omega-\alpha=\frac{k}{k-1}\left(\frac{\sigma_{\lambda}{}^{2}}{\sigma_{Y}{}^{2}}\right)$$

olarak elde edilmektedir. ω-α=0 olma durumu ancak faktör yükleri eşit olduğunda olanaklıdır (McDonald, 1999). Konjenerik ölçmelerde McDonald'ın ω güvenirlik katsayısı da yanlıdır ancak bu tür ölçmelerde en küçük yanlılık ω katsayısına ait olduğundan (Zinbarg vd., 2005) aynı zamanda en yüksek değere sahip katsayı olarak elde edilmektedir.

Bu sonuçlara göre ikili derecelenmiş testlerde olduğu gibi, çoklu derecelenmiş testlerde maddelere ilişkin kovaryans terimleri ya da faktör yükleri eşit olduğunda α, β, Ω ya da ω aksi taktirde ise yalnızca ω katsayısının kullanımı önerilebilir.

Bu çalışma, farklı yapıdaki ölçmeleri ve bu ölçmeler için güvenirlik katsayılarının kullanımı ve bu katsayılara ilişkin basit karşılaştırmalar üzerine kurulmuştur. Elde edilen sonuçlar yalnızca Tablo 2'te verilen veri kümeleri için geçerlidir. Örneğin; Carmines ve Zeller (1979) ancak paralel ölçmelerde α=θ=Ω olabileceğini belirtirken bu çalışma da paralel ve konjenerik ölçmeler için de her üç katsayı eşit bulunmuştur. Literatürde α ve ω katsayıları için kuramsal düzeyde karşılaştırmalar yapılmıştır (Lucke, 2005a, 2005b; Raykov, 1997a, 1997b; Zinbarg vd, 2005). Benzer karşılaştırmalar daha çok sayıdaki maddeler için tekrar ele alınabilir.

Güvenirlik çalışmalarını içeren araştırmalarda α güvenirlik katsayısı, bazı çalışmalarda bir kavram yanılgısı olarak ölçme aracının gerçek güvenirliği olarak rapor edildiği görülmektedir. Oysaki bu durum ancak ölçme aracını oluşturan maddelerin paralel, eşdeğer ya da eşbiçimli ölçmeler olduğu durumlar için söz konusudur. Bir diğer ifade ile; ölçme aracındaki maddelere ilişkin standartlaştırılmamış faktör yükleri eşit olmadığı durumlarda elde edilen güvenirlik katsayılarının değerleri, gerçek güvenirlik değerinin alt sınırını vermektedir. Araştırmacıların güvenirlik katsayılarını "gerçek güvenirlik değeri" olarak rapor etmeden önce maddelerin yapısının analiz edilmesi önerilir.

KAYNAKLAR

Adamson, G., Shevlin, M., Lloyd, N.S.V. & Lewis, C. A. (2000). An integrated approach for assessing reliability and validity: an application of structural equation modeling to the measurements of religiosity. *Personality and Individual Differences*, 29, 971-979.

Alwin, D. F. (1976). Attitude scales as congeneric tests: A re-examination of an attitude-behavior model. *Sociometry, 39,* 377-383.

Armor, D. J. (1974). *Theta reliability and factor scaling.* In Costner, H. L. (ed.), Sociological Methodology. Jossey-Bass, San Francisco. 17−50.

Bacon, D. R., Sauer, P. L. & Young M. (1995). Composite Reliability in Structural Equations Modeling. *Educational and Psychological Measurement*, 55, 394-406.

Baykul, Y. (2000). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması.* Ankara: ÖSYM Yayınları.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*, Sage, Beverly Hills, Calif.

Christmann, A. & Aelst, S. V. (2005). Robust estimation of Cronbach's Alpha. *Journal of Multivariate Analysis*. (Baskıda).

Comrey, A. L. & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: Erlbaum.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78,* 98–104.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika,16,* 297-334.

DeVellis, R. F. (2003). *Scale development: Theory and application*, Sage Publications, California.

Feldt, L. S. & Qualls, A. L. (1996). Bias in coefficient alpha arising from heterogeneity of test content. *Applied Measurement in Education. 9(3),* 277-286.

Ghiselli, E. E. (1964). *Theory of Psychological Measurement*. McGrawHill, New York.

Green, S. B., & Hershberger, S. L. (2000). Correlated errors in true score models and their effect on coefficient alpha. *Structural Equation Modeling*, *7*, 251–270.

Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, *10*, 255-282.

Heise, D. R., & Bohrnstedt, G. W. (1970). *Validity, invalidity and reliability.* In Borgatta, E. F. and Bohrnstedt, G. W. (eds.), *Sociological Methodology*. Jossey-Bass, San Francisco. 104–129.

Jöreskog, K. G. (1971). Statistical analysis of congeneric tests. *Psychometrika*, 36, 109-133.

Komaroff, E. (1997). Effect of simultaneous violations of essential tau-equivalence and correlated errors on coefficient alpha. *Applied Psychological Measurement*, 21, 337–348.

Lord F. M. & Novick, R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.

Lucke, J. F. (2005a). The α and ω of congeneric test theory: An extension of reliability and internal consistency to heterogeneous tests. *Applied Psychological Measurements*. *29(1)*, 65-81.

Lucke, J. F. (2005b). "Rassling the hog": the influence of correlated item error on internal consistency, classical reliability, and congeneric reliability. *Applied Psychological Measurements. 29(2)*, 106-125.

MacCallum, R. C., Widaman, K. F., Zhang, S., and Hong, S. (1999). Sample size in factor analysis. *Psychological Methods* 4: 84 –99.

McDonald, R. (1985). *Factor analysis and related methods.* Hillsdale, N J: Erlbaum.

McDonald R. P. (1999). *Test theory: A unified treatment.* Mahwah, NJ: LEA Publisher.

Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling*, 2, 255–273.

Novick, M. R. & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika, 32*, 1-13.

Nunnally, J. C. & Bernstein, I. H. (1994): *Psychometric theory*. 3rd Edition. McGraw-Hill: New York.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods, 5,* 343–355.

Rae. G. (2006). Correcting Coefficient Alpha for Correlated Errors: Is α_K a Lower Bound to Reliability? *Applied Psychological Measuremet,* 30(1), 56-59.

Raykov, T. (1997a). Estimation of composite reliability for congeneric measures, *Applied Psychological Measurements*. *21(2)*, 173-184.

Raykov, T. (1997b). Bias of coefficient α for fixed congeneric measures with correlated errors. *Applied Psychological Measurements*. *25(1)*, 69-76.

Raykov, T. (2001). Bias of coefficient α for fixed congeneric measures with correlated errors. *Applied Psychological Measurements*. *25(1)*, 69-76.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15,* 72–101.

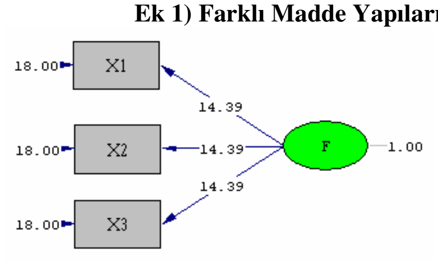
Traub, E. R. (1994). *Reliability for the social sciences: Theory and Applications*. Measurement methods for the social sciences. Sage Publications, 1994.

Yurdugül, H. (2005). Konjenerik test kuramı ve konjenerik madde analizi: Tek boyutlu çoktan seçmeli testler üzerine bir uygulama. *A.Ü. Eğitim Bilimleri Fakültesi Dergisi 38(2),* Baskıda.

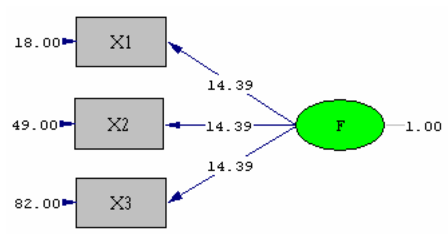
Zinbarg, R. E., Revelle, W., Yovel, I. & Li, W. (2005). Cronbach's α, Revelle's, β and McDonalds ω: their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70(1)*, 1-11.

Zimmerman, D. W., Zumbo, B.D. & Lalonde, C. (1993). Coefficient Alpha as an estimate of test reliability under violation of two assumptions. *Etucational and Psychological Measurement, 53* (1), 33-49.

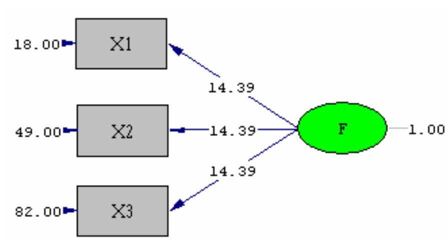
Ek 1) Farklı Madde Yapılarına Göre Faktör Çözümleri



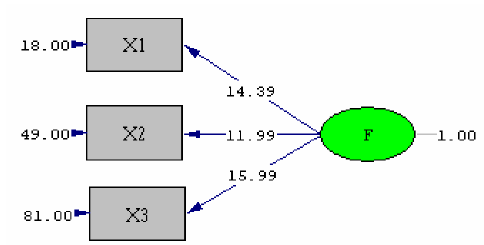
a) Paralel maddelere dayalı standartlaştırılmamış faktör çözümlemesi



b) Eşdeğer maddelere dayalı standartlaştırılmamış faktör çözümlemesi



c) Eşbiçimli maddelere dayalı standartlaştırılmamış faktör çözümlemesi



d) Konjenerik maddelere dayalı standartlaştırılmamış faktör çözümlemesi