

## DÖRT FARKLI ALFA KATSAYISININ PSİKOMETRİK AÇIDAN KARŞILAŞTIRILMASI

### THE COMPARISON OF FOUR DIFFERENT COEFFICIENT ALPHAS FROM A PSYCHOMETRIC POINT OF VIEW

Halil YURDUGÜL\*

**ÖZET:** Bu çalışmada ölçme araçlarının içtutarlılığını belirlerken kullanılan dört farklı alfa katsayısının psikometrik özellikleri karşılaştırılmıştır. Bu katsayılar sırasıyla; a) ölçmelerin varyans-kovaryans terimlerinden elde edilen standartlaştırılmamış alfa ( $\alpha_{VK}$ ) katsayısı, b) ölçmeler arasındaki korelasyon terimlerinden elde edilen standartlaştırılmış alfa ( $\alpha_{STD}$ ) katsayısı, c) temel bileşenler analizinden elde edilen enbüyük özdeğer üzerine kurulu olan ve Armor (1974) tarafından önerilen  $\alpha_{Max}$  katsayısı ve d) doğrulayıcı faktör analizinden elde edilen faktör yükleri üzerine kurulu olan ve McDonald (1985) tarafından tanımlanan  $\alpha_M$  katsayısının özellikleri Monte-Carlo simülasyonu ile karşılaştırılmıştır. Karşılaştırmalar 3 aşamada gerçekleştirilmiştir. Bunlar sırasıyla; a) bu katsayıların örneklem karşılıklarının hata kareler ortalaması karekökü (RMSE) değerlerine göre sağlam kestirici özellikleri, b) tekboyutlu ve çokboyutlu veri kümelerindeki değerleri ve c) paralel, eşdeğer ve konjenetik veri kümelerindeki davranışları karşılaştırılmıştır. Araştırma sonucunda;  $\alpha_{Max}$  katsayısı her üç aşamada da diğer katsayılarla göre üstün özelliklere sahiptir. Yine her üç aşamada en olumsuz özellik ise  $\alpha_M$  katsayısında gözlenmiştir.

**Anahtar sözcükler:** güvenilirlik, içtutarlılık, güvenilirlik katsayısı, alfa katsayısı.

**ABSTRACT:** In this study, the properties of four different coefficient alphas in reliability were compared. Those were a) unstandardized alpha ( $\alpha_{VK}$ ) based on variance-covariance matrix, b) standardized alpha ( $\alpha_{VK}$ ) based on correlation matrix, c) maximum alpha ( $\alpha_{Max}$ ) based on first eigenvalues by Armor (1974) and d)  $\alpha_M$  defined by McDonald (1985) and based on confirmatory factor analysis. The comparisons made to determine the best coefficient conducted on Monte-Carlo simulations and received priority consideration were carried out in three stages: a) investigation of robust estimator properties according to root mean square error (RMSE) for four coefficient alphas, b) comparison of coefficient alphas' values in unidimensional data sets and multidimensional data sets, and c) comparison of coefficient alphas' behaviors in parallel, tau-equivalent, and congeneric data sets. As a result, the coefficient  $\alpha_{Max}$  is found to be the optimum coefficient among those three stages. The coefficient  $\alpha_M$  is found to be non-optimum in all the stages.

**Keywords** reliability, internal consistency, reliability coefficient, coefficient alpha.

## 1. GİRİŞ

Eğitim araştırmalarında ölçme araçlarının güvenilirliği önemli konulardan birisidir. Özellikle çoklu derecelendirilmiş (polytomous) ölçme araçlarının içtutarlılığının kestiriminde Guttman (1945) ve Cronbach (1951) tarafından geliştirilen alfa katsayısının yaygın bir şekilde kullanıldığı görülmektedir. Alfa katsayısının, tanımlanmış bazı olumsuzluklarına (ileride açıklanmıştır) karşın hesaplanmasındaki kolaylık ve paket programlarda yer alması nedeniyle halen popülerliğini korumaktadır.

Diğer taraftan ilgili literatürde alfa katsayısının farklı versiyonları tanımlanmıştır. Bunlardan yaygın olarak ele alınanlardan dört tanesi; a) varyans-kovaryans matrisi üzerinden elde edilen *standartlaştırılmamış alfa katsayısı* ( $\alpha_{VK}$ ), b) korelasyon matrisi üzerinden elde edilen *standartlaştırılmış alfa katsayısı* ( $\alpha_{STD}$ ), c) Armor (1974) tarafından tanımlanan ve temel bileşenler analizi terimleriyle açıklanan *maksimum alfa katsayısı* ( $\alpha_{Max}$ ) ve d) McDonald (1985) tarafından tanımlanan ve doğrulayıcı faktör analizi terimleriyle açıklanan alfa katsayısıdır ( $\alpha_M$ ).

Bu çalışmada yukarıda tanımlanan dört farklı alfa katsayısının karşılaştırması yapılarak optimallikleri araştırılmıştır. Bu amaca yönelik olarak üç karşılaştırma ölçütü belirlenmiştir. Bunlardan ilki; alfa katsayılarının örneklem karşılıklarının sağlam kestirici (robust estimator) özelliğinin araştırılması, ikincisi bu katsayıların farklı boyutlardaki veri kümelerinden elde edilen değerlerinin

\* Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, yurdugul@hacettepe.edu.tr

karşılaştırılması ve son olarak bu katsayıların konjenerik veri kümelerindeki performanslarının karşılaştırılmasıdır. Bu ikinci ve üçüncü ölçütler, özellikle alfa katsayısının çokboyutlu ve/veya konjenerik veri kümelerinde gerçek güvenilirliğin altında değer ürettiğine ilişkin eleştirilere karşılık ele alınmıştır. Araştırmada veri kümesi olarak; Schwarzer ve Jerusalem (1995) tarafından geliştirilen 10 maddelik ‘genel özyeterlik ölçeği’ yardımıyla 25 ülkeden 19120 bireyden elde edilen ve Scholz vd. (2002) tarafından çalışmalarında rapor edilen veriler kullanılmıştır. Araştırmanın amacına uygun olarak; bu veri kümesi çalışmada evren olarak ele alınmış bootstrap yöntemine dayalı Monte-Carlo simülasyonu tekniğinden yararlanılmıştır.

### 1.1. Güvenirlik ve Alfa Katsayıları

Güvenirlik ölçmelerin hatasızlığına ilişkin bir kavramdır ve güvenilirliğin nicel değeri; ölçülmek istenilen gerçek puanlar varyansının  $[\text{Var}(T)]$ , gözlenen toplam puanlar varyansına  $[\text{Var}(X)]$  oranı ile elde edilir. “Gerçek güvenilirlik” (true reliability) olarak adlandırılan bu *güvenirlik indeksi* gerçek puanlar doğrudan elde edilemediğinden dolayı istatistiksel yöntemler ile kestirilir ve kestirim değeri ise *güvenirlik katsayısı* olarak adlandırılır. Ölçmelerin içtutarlılığı anlamında güvenilirliğine yönelik ilk çalışma Spearman (1904) ile başlamıştır. Ancak Spearman (1904) güvenilirliğin kestirimini paralel (ortalama ve varyansları eşit) ölçmeler üzerinde tanımlamıştır. Guttman (1945) ise davranış bilimlerinde paralel ölçmeleri elde etmenin uygulamalarda olanaklı olmadığını ifade etmiş ve Spearman’ın önerdiği katsayıya alternatif olarak 6 adet güvenilirlik katsayısı önermiştir. Guttman (1945), bu katsayıların her zaman gerçek güvenilirliği değil, gerçek güvenilirliğin alt sınırını vereceğini ifade ederek bu katsayıları güvenilirlik katsayısı yerine “güvenirliğin alt sınırı” olarak tanımlamıştır. Guttman’ın sıralaması ve orijinal sembolleri ile bu katsayılar  $L_1, \dots, L_6$  katsayılarıdır. Bu katsayılardan özellikle  $L_3$  katsayısı ölçme literatüründe diğerlerine göre daha çok ön plana çıkmaktadır.

$$L_3 = \frac{k}{k-1} \left[ 1 - \frac{\sum_{i=1}^k \text{Var}(X_i)}{\text{Var}(X)} \right] \quad (1)$$

Burada  $k$ , ölçmelerin (madde puanlarının) sayısını,  $\text{Var}(X_i)$  i. ölçmenin varyansını ve  $\text{Var}(X)$  ise tüm ölçme kümesinin varyansını göstermektedir. Cronbach (1951) ise  $L_3$  katsayısının ölçmelerin paralel olmasa bile eşdeğer olması durumunda dahi gerçek güvenilirliği vereceğini ifade ederek bu katsayıyı  $\alpha$  gösterimiyle yeniden önermiştir. Bu nedenle  $\alpha$  katsayısı bazı kaynaklarda Guttman-Cronbach katsayısı olarak verilmektedir. Varyans-kovaryans terimleri (matrisi) üzerine kurulu olan  $\alpha_{VK}$  katsayısının matrisler ile gösterimi Eşitlik 2’de verilmiştir.

$$\alpha_{VK} = \frac{k}{k-1} \left[ \frac{iz(\Sigma)}{\mathbf{1}'\Sigma\mathbf{1}} \right] \quad (2)$$

Burada  $\Sigma$ ,  $k \times k$  boyutlu varyans-kovaryans matrisini,  $\mathbf{1}$  ise tüm elemanları 1 olan  $k \times 1$  boyutlu vektörü ve  $iz(\Sigma)$  ise  $\Sigma$  matrisinin iz’ini (köşegen elemanlarının toplamı) göstermektedir. Evrene ilişkin  $\alpha_{VK}$  katsayısının örneklem karşılığı ise;

$$\hat{\alpha}_{VK} = \frac{k}{k-1} \left[ \frac{iz(\mathbf{S})}{\mathbf{1}'\mathbf{S}\mathbf{1}} \right], \quad (3)$$

şeklinde. Burada  $\mathbf{S}$ , örneklemden elde edilen varyans-kovaryans matrisidir. Çokdeğişkenli normal dağılım varsayımı altında; örneklem alfa katsayısı ( $\hat{\alpha}_{VK}$ ) aynı zamanda evren alfa katsayısının  $\alpha_{VK}$  ençok olabilirlik kestiricisidir (maximum likelihood estimator) (van Zyl vd., 2000; Yuan & Bentler, 2002).

Bununla birlikte; çoğu uygulamada ve özellikle bilgisayar paket programlarında alfa katsayısı standartlaştırılmış veriler (ya da korelasyon matrisi) üzerinden elde edilmektedir. Bu yaklaşımda Eşitlik 1 ile verilen ifade;

$$\alpha_{STD} = \frac{k\bar{\rho}}{1 + \bar{\rho}(k-1)} \quad (4)$$

şekline dönüşmektedir. Burada  $\bar{\rho}$  ölçmeler arası korelasyonların ortalamasını göstermektedir. Diğer taraftan standartlaştırılmış alfa katsayısının ( $\alpha_{STD}$ ) matris notasyonu ile gösterimi ise;

$$\alpha_{VK} = \frac{k}{k-1} \left[ \frac{k}{\mathbf{1}'\mathbf{R}\mathbf{1}} \right] \quad (5)$$

şeklinde ve  $\mathbf{R}$ , kxk boyutlu ölçmeler arasındaki korelasyon matrisini göstermektedir.

Güvenirlilik katsayılarına faktör analizi perspektifinden bakıldığında ise; birleşik testlerde kullanılan güvenirlik katsayıları aynı zamanda tüm ölçmelerin birinci faktörünün büyüklüğünü ölçmektedir (Cortina, 1993; Crano & Brewer, 1973). Bu nedenle  $\alpha_{VK}$  katsayısı temel bileşen analizindeki ilk temel bileşenin özdeğeri ile doğrudan ilişkilidir (Cortina, 1993). Armor (1974) bu bağıntıdan yola çıkarak  $\alpha_{VK}$  katsayısının yeni bir versiyonu olan  $\alpha_{Max}$  katsayısını önermiştir.

$$\alpha_{Max} = \frac{k}{k-1} \left[ 1 - \frac{1}{\delta_1} \right] \quad (6)$$

Burada  $\delta_1$ , temel bileşenler analizinden elde edilen en büyük özdeğerdir. Bazı kaynaklarda  $\alpha_{Max}$  katsayısı *enbüyüklenmiş  $\alpha$  katsayısı* (maximized alfa) olarak da adlandırılmaktadır. Bunun nedenlerinden birisi ise  $\alpha_{VK}$  ile  $\alpha_{Max}$  arasındaki bağıntının  $\alpha_{VK} \leq \alpha_{Max}$  olmasından kaynaklanmaktadır (Carmines & Zeller, 1979). Diğer taraftan McDonald (1985),  $\alpha_{VK}$  katsayısının doğrulayıcı faktör analizinden elde edilen faktör yükleri üzerine kurulu versiyonunu ( $\alpha_M$ ) tanımlamıştır:

$$\alpha_M = \frac{k}{k-1} \left[ \frac{k(\bar{\lambda})^2 - (\bar{\lambda}^2)}{k(\bar{\lambda})^2 - \bar{\theta}^2} \right] \quad (7)$$

Burada,  $\bar{\lambda}$  doğrulayıcı faktör analizinden elde edilmiş standartlaştırılmış faktör yükleri ortalamasını ve  $\bar{\theta}$  ise ölçmelerin hata terimleri ortalamasını göstermektedir.

Aslında Guttman'ın (1945) çalışması güvenirlik literatüründe bir lokomotif görevi görmüş ve 1945 yılından sonra, gerçek güvenirliği en yüksek kestirdiği iddia edilen çok sayıda güvenirlik katsayısı önerilmiş ve/veya en popüler kullanıma sahip olan alfa katsayısını enbüyüklemek için çeşitli katsayılar ya da yöntemler geliştirilmiştir. Ancak günümüzde hesaplama kolaylığı ve paket programlarda yer alması nedeniyle en yaygın kullanıma sahip olan güvenirlik katsayısının  $\alpha$  ya da  $\alpha_{STD}$  katsayısı olduğu gözlenmektedir. Öyle ki 2004 yılına kadar 5590'dan fazla araştırmada Cronbach'ın (1951) makalesine referans verilmiş (Cronbach & Shavelson, 2004) ve yılda ortalama 325 adet Social Sciences Citation Index-SSCI kapsamındaki dergilerden bu makaleye referans verilmiştir (Liu & Zumbo 2007).

Alfa katsayısının bu denli yaygın kullanılmasına karşın bazı olumsuzlukları da söz konusudur. Bu olumsuzlukların en önemlilerini iki kategoride ele almak olanaklıdır: a) veri kümesi tekboyutluluktan uzaklaştıkça alfa katsayısı gerçek güvenirliğe göre düşük değerler üretmektedir (Cortina, 1993; Cronbach, Schonemann & McKie, 1965; Green, Lissitz & Mulaik, 1977; Schmitt, 1996; Zimmerman, Zumbo, & Lalonde, 1993), b) özellikle konjenerik ölçmelerde alfa katsayısı gerçek güvenirliği yanlış bir şekilde kestirmektedir (Feldt & Qualls, 1996; Jackson & Agunwamba, 1977; Komaroff, 1997; Lucke, 2005; Rae, 2006; Raykov, 1997, 1998, 2001; Traub, 1994; Zimmerman, Zumbo, & Lalonde, 1993; Zinbarg vd, 2005).

Bu çalışmanın kapsamına dört farklı alfa katsayısı ( $\alpha_{VK}$ ,  $\alpha_{STD}$ ,  $\alpha_{Max}$  ve  $\alpha_M$ ) ele alınmış ve bu katsayıların optimallikleri araştırılmıştır. Optimallik araştırması, parametre ile kestirici tutarlılığı, alfa katsayılarının değişik boyut sayısına sahip veri kümelerindeki ve farklı ölçme yapısına (paralel, eşdeğer ve konjenerik) sahip veri kümelerindeki davranışları üzerine yapılandırılmıştır.

## 2. YÖNTEM

### 2.1. Veri Kümesi

Araştırmanın veri kümesi olarak Matthias Jerusalem ve Ralf Schwarzer tarafından geliştirilen “genel özyeterlik algı” ölçeğine ilişkin 25 ülkeden 19120 kişiden toplanan veriler<sup>1</sup> kullanılması amaçlanmıştır. Ancak, veri kümesinde bazı boş gözlemlerin (çalışma sonuçlarından *kayıp gözlem* etkilerini arındırmak için) veri kümesinden çıkartılması nedeniyle veri kümesi 18178 gözleme inmiştir. Bu veri kümesi araştırmada evren olarak kabul edilmiş ve bootstrap yöntemi ile Monte-Carlo simülasyonuna dayalı olarak farklı genişlikteki örneklemde elde edilerek araştırmada kullanılmıştır.

### 2.2. Araştırma ve Simülasyon Tasarımı

#### 2.2.1. Parametre-kestirici tutarlılığı:

Ölçme araçlarının bir özelliği olan güvenilirlik kavramı “güvenirlik indeksi” ile nicelenir. Güvenirlik katsayıları ise ölçme sonuçları ile ilişkilidir ve bu katsayılar güvenilirlik indeksinin bir kestiricisidir. Güvenirlik indeksleri doğrudan gözlenemediği için güvenilirlik katsayıları yardımı ile kestirilir. Bu kestirimde bir bilgi kaybı söz konusudur. Buna ek olarak bir de güvenilirlik katsayıları örneklemde elde edildiği zaman buna [Yan=parametre-kestirici] hataları da eklenmektedir. Bu nedenle dört farklı alfa katsayısının optimallik karşılaştırmasında ilk aşamayı parametre ve kestirici tutarlılığı oluşturmaktadır.

Bu aşamada; evrene ilişkin veri kümesinden elde edilen parametreler ( $\alpha_{VK}$ ,  $\alpha_{STD}$ ,  $\alpha_{Max}$  ve  $\alpha_M$ ) ile Monte-Carlo simülasyonu ile elde edilen örneklem değerleri ( $\hat{\alpha}_{VK}$ ,  $\hat{\alpha}_{STD}$ ,  $\hat{\alpha}_{Max}$ ,  $\hat{\alpha}_M$ ) karşılaştırılmıştır. Bu karşılaştırma yapılırken, belirli örneklem genişliklerinin (N=30, 100, 300 ve 500) her biri için M=1000 adet basit seçkisiz (random) örneklem yöntemi ile örneklem çekilmiş ve bu örneklemde alfa katsayılarının örneklem değerleri elde edilerek bu katsayılarla ilişkin hata indeksi hesaplanmıştır. Hata indeksi olarak hata kareler ortalaması karekökü (root mean square error, RMSE) kullanılmıştır.

$$RMSE = \sqrt{\frac{1}{M} \sum_{j=1}^M (Kestirici_j - Parametre)^2} \quad (8)$$

Burada  $Kestirici_j$ , j. örneklemde elde edilen kestirilen alfa değerini göstermektedir. M ise simülasyondaki tekrarlanan örneklem sayısını göstermektedir. RMSE değerleri 0'a yaklaştıkça ilgili kestiricinin o derecede sağlam kestirici (robust estimator) ve aynı zamanda tutarlı kestirici olduğu ifade edilir (Poon, Leung, & Lee, 2002). Buna göre;  $\alpha_{VK}$ ,  $\alpha_{STD}$ ,  $\alpha_{Max}$  ve  $\alpha_M$  katsayılarının farklı örneklem genişliklerinde RMSE değerlerine göre tutarlılıkları incelenmiştir.

#### 2.2.2. Veri Kümesinin Boyutluluğu ve Alfa Katsayıları:

Alfa katsayısına yönelik eleştirilerin başında; tek boyutlu veri kümelerinde güvenilirlik indeksini daha yansız kestirdiği, çok boyutlu veri kümelerinde ise güvenilirlik indeksinden daha düşük değerler ürettiği yönündeki olumsuzluk gelmektedir. Bu araştırmada dört alfa katsayısının bir önceki aşamada elde edilen örneklem veri kümelerindeki faktör sayılarına göre davranışı incelenmiştir. Böylelikle farklı boyutlardaki veri kümeleri için  $\alpha_{VK}$ ,  $\alpha_{STD}$ ,  $\alpha_{Max}$  ve  $\alpha_M$  katsayılarının aldığı değerlerin büyüklüklerine göre sıralanması amaçlanmıştır. Bu aşamada faktör sayıları Kaiser-Guttman kuralı

<sup>1</sup> Ölçek verilerinin bu çalışmada kullanımı için Dr. Schwarzer tarafından izin alınmıştır. Bu ölçme sonuçlarına ilişkin genel değerlendirme Scholz, Gutiérrez-Doña, Sud & Schwarzer (2002) tarafından rapor edilmiştir.



(Gorsuch, 1983) gereği; temel bileşenler analizinden elde edilen özdeğerlerin 1 değerinden büyük olma durumuna göre belirlenmiştir.

### 2.2.3. Ölçmelerin Yapısı ve Alfa Katsayıları:

Araştırmalarda elde edilen ölçmelerin paralel veya eşdeğer olması (standartlaştırılmamış faktör yüklerinin ya da diğer ifade ile ölçmelerin varyanslarının eşit olması) istenilen bir durumdur. Ancak eğitim ve psikoloji alanındaki araştırmalarda paralel veya eşdeğer ölçmeleri elde etmek neredeyse olanaksızdır. Oysaki alfa katsayısı ancak paralel ve eşdeğer ölçmelerde güvenilirlik indeksini (gerçek güvenilirlik) yansız kestirebilmektedir (Jöreskog, 1971; Miller, 1995; Traub, 1994; Yurdugül, 2005). Konjenerik ölçmelerde (faktör yüklerinin eşit olmama durumu) ise alfa katsayısı yanlış kestirimlerde bulunmakta, gerçek güvenirliliğin altında değerler üretmektedir. Bu nedenle dört farklı alfa katsayısına ilişkin karşılaştırmanın son aşamasında doğrulayıcı faktör analizinden elde edilen yüklerin ortalamaya göre sapmalarına (faktör yüklerinin varyansları- $\phi$ ) ilişkin alfa katsayılarının farkının davranışı incelenmiştir<sup>2</sup>.

$$\phi = \frac{\sum_{i=1}^{k=10} (\beta_i - \bar{\beta})^2}{k-1} \quad (9)$$

Burada  $\beta_i$ , i. ölçmeye ilişkin faktör yükünü göstermektedir. Bilindiği gibi; standartlaştırılmış faktör yüklerinin eşit olması ölçmelerin paralel ya da eşdeğer olduğunu göstermektedir. Buna göre tüm yükler eşit olduğunda  $\phi=0$  olacaktır. Diğer taraftan  $\phi$  değeri 0'dan uzaklaştıkça ( $\phi \rightarrow +\infty$ ) ölçmelerde o denli paralel ya da eşdeğer ölçme özelliğinden uzaklaşıp konjenerik ölçmeye yöneldiğini göstermektedir<sup>3</sup>.

Araştırmanın yöntem bölümünde belirlenen karşılaştırmalardaki simülasyonlar için SIMREL adı verilen bir program (Yurdugul, baskıda) kullanılmıştır.

## 3. BULGULAR

### 3.1. Alfa Katsayılarının Sağlam Kestirici Özellikleri

Araştırmanın evrene ilişkin veri kümesinden elde edilen dört adet alfa katsayısının *parametre değerleri* sırasıyla;  $\alpha_{VK}=0.856$ ,  $\alpha_{STD}=0.857$ ,  $\alpha_{Max}=0.859$  ve  $\alpha_M=0.856$  bulunmuştur. Daha sonra evren kümesinden örneklem genişliği sırasıyla N=30, 100, 300 ve 500 olacak şekilde (her bir örneklem genişliği için) basit seçkisiz örnekleme yöntemi ile 1000 adet örneklem çekilerek dört adet alfa katsayısının örneklem değerleri elde edilmiştir. Her bir örneklem değeri için  $RMSE(\alpha_{VK})$ ,  $RMSE(\alpha_{STD})$ ,  $RMSE(\alpha_{max})$  ve  $RMSE(\alpha_M)$  değerleri hesaplanmıştır. Bu simülasyon sonuçlarına ilişkin alfa katsayılarının farklı örneklem genişliklerindeki RMSE değerleri ve grafiği aşağıda verilmiştir.

**Tablo 1: Simülasyon sonuçlarına göre N=30, 100, 300, 500 örneklem genişliklerindeki alfa katsayılarının RMSE değerleri**

N	RMSE ( $\alpha_{VK}$ )	RMSE ( $\alpha_{STD}$ )	RMSE ( $\alpha_{Max}$ )	RMSE ( $\alpha_M$ )
30	0.058	0.055	0.046	0.060
100	0.030	0.030	0.028	0.030
300	0.015	0.015	0.015	0.015
500	0.011	0.011	0.011	0.011

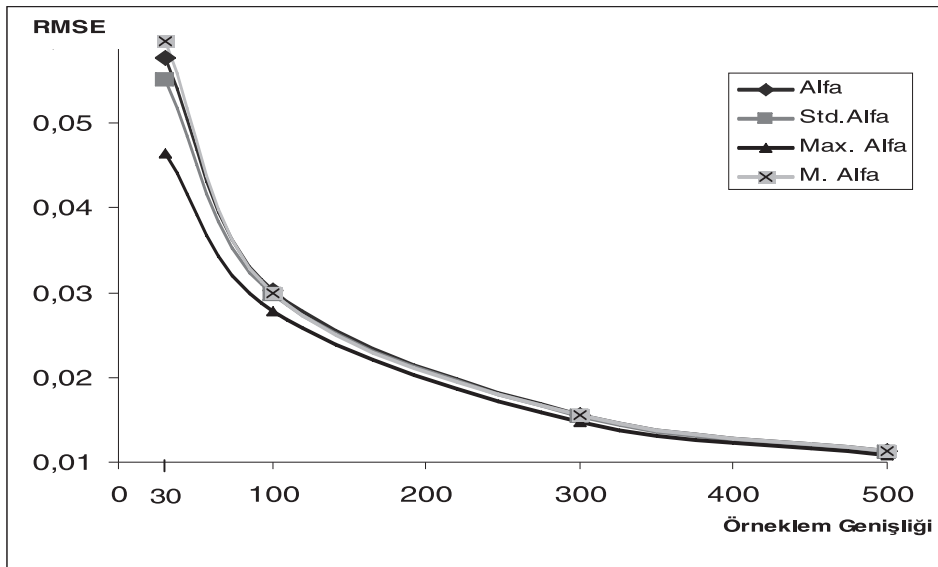
<sup>2</sup> Konjenerik ölçmelerin önemli nedenlerinden birisi (2. aşama karşılaştırmalarında konu edilen) çokboyutlu yapılardır. Ancak konjenerik ölçmelere sahip tekboyutlu veri kümelerinin varlığı nedeniyle (Lucke, 2005; Zinbarg vd., 2005) ölçmelerin yapılarına göre alfa katsayılarının karşılaştırmaları ayrı olarak ele alınmıştır.

<sup>3</sup> Bu çalışmada  $\phi=0$  değeri ölçme varyanslarının eşit olduğunu göstermektedir. Bu durumda ölçmeler paralel ya da eşdeğer olabilir (Lucke, 2005). Ölçmelerin paralel mi yoksa eşdeğer mi olduğunu anlamak için ortalamalarındaki sapmaları araştırmak gerekmektedir. Ancak bu farklılık ölçme çalışmaları ile doğrudan ilişkili olmadığı için  $\phi=0$  değeri taşıyan ölçmeler "eşdeğer ölçme" olarak adlandırılmıştır.

Daha önce ifade edildiği gibi; RMSE değerleri incelemeye konu olan istatistiklerin sağlam kestirici özelliklerini ortaya koyan bir değerdir ve bu değer 0'a yakın ise incelenen istatistik parametrenin o denli sağlam bir kestiricisi olarak nitelendirilir. RMSE değerleri örneklem genişliklerinden fazlaca etkilenir. Tablo 1'e göre; N=300 ve N=500 olduğunda tüm alfa değerlerinin RMSE değerleri eşit olarak elde edilmiştir. Ancak bu çalışmada ele alınan en düşük örneklem genişliğinde (N=30), en küçük RMSE değeri  $\alpha_{Max}$  ve en büyük RMSE değeri ise  $\alpha_M$  için gözlenmiştir.

Alfa katsayılarına ilişkin RMSE değerlerinin farklı örneklem genişliklerindeki davranışı Çizim 1'de verilmiştir. Buna göre N=30 olduğunda  $RMSE(\alpha_{Max}) < RMSE(\alpha_{STD}) < RMSE(\alpha_{VK}) < RMSE(\alpha_M)$  ilişkisi gözlenmiştir. N=100 için ise  $RMSE(\alpha_{Max}) < RMSE(\alpha_{VK}) = RMSE(\alpha_{STD}) = RMSE(\alpha_M)$  sıralaması söz konusudur. Son olarak N=300 ve N=500 olduğunda ise her dört alfa katsayısının RMSE değerleri eşit olarak elde edilmiştir. Buna göre dört farklı alfa katsayılarının örneklem karşılıklarından küçük örneklem için (N=30, N=100) en sağlam kestiricinin  $\alpha_{Max}$  olduğu söylenebilir.

Genel olarak RMSE değerleri ele alındığında  $RMSE(\alpha_{Max})$  değerinin tüm örneklem genişliklerinde diğer alfa katsayılarının RMSE değerlerinde düşük ya da eşit olduğu Çizim 1'de görülmektedir. Buna göre; sağlam kestirici özelliğinden dolayı en iyi alfa katsayısının Armor (1974) tarafından önerilen  $\alpha_{Max}$  katsayısı olduğu ifade edilebilir.



**Çizim 1: Farklı örneklem genişliklerinde alfa katsayılarının RMSE değerlerinin davranışı**

### 3.2. Veri Kümesinin Boyutluluğu ve Alfa Katsayıları:

Güvenirlik konusuna ilişkin çalışmalarda; üzerinde çalışılan veri kümesinin boyut sayısı arttıkça alfa katsayısının düşük değerler ürettiği ifade edilir. Bu nedenle; bu araştırma kapsamındaki alfa katsayılarının temel bileşenler analizinden elde edilen faktör sayısına göre aldığı değerlerin büyüklükleri incelenmiştir. Faktör sayıları, veri kümesinin özdeğerlerinin 1'den büyük olma durumlarına göre belirlenmiştir. Diğer taraftan simülasyon sonucu elde edilen veri kümelerini 1, 2, 3 ve 4 faktörlü yapılara sahip olduğu görülmüştür<sup>4</sup>. Tablo 2'de bu faktörlere sahip veri kümelerinden elde edilen dört farklı alfa katsayılarının örneklem ortalamaları verilmiştir.

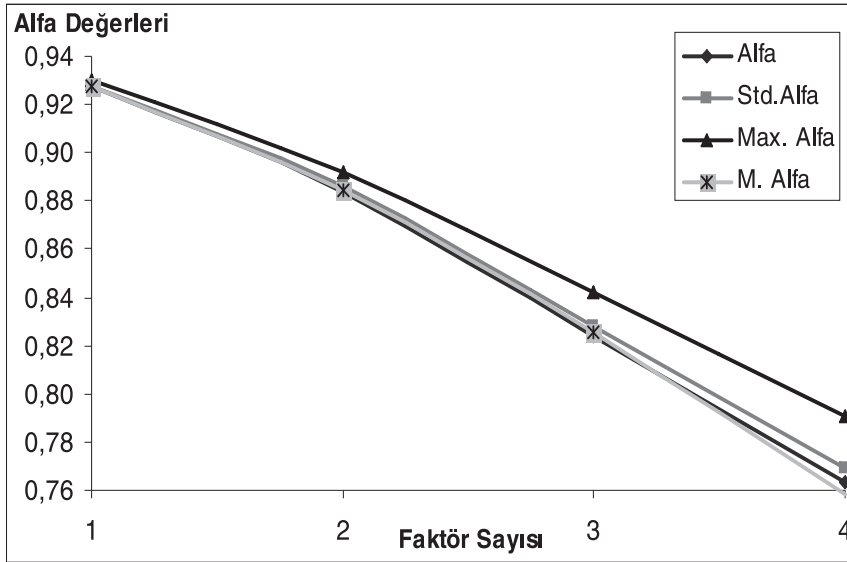
<sup>4</sup> Evren veri kümesi tekboyutlu olduğu için büyük örneklemelerden elde edilen örneklem veri kümelerinde faktör sayıları da 1 olmaktadır. Bu nedenle faktör sayılarındaki çeşitliliği artırmak için örneklem genişliği 200 olarak belirlenmiştir.

**Tablo 2: Simülasyon sonuçlarına göre farklı boyutlardaki örneklemelerden elde edilen alfa katsayılarının ortalama değerleri**

Faktör	$\alpha_{VK}$	$\alpha_{STD}$	$\alpha_{Max}$	$\alpha_M$
4	0.763	0.769	0.790	0.759
3	0.824	0.828	0.842	0.825
2	0.884	0.886	0.892	0.885
1	0.927	0.928	0.930	0.928

Tablo 2'ye göre; alfa katsayılarının en yüksek ortalama değerleri bir faktörlü (tekboyutlu) veri kümelerinden elde edilmiştir ve faktör sayısı arttıkça tüm katsayıların değerleri düşme göstermektedir. Alfa katsayılarının ortalama değerleri (faktör sayılarına göre) göz önüne alındığında  $\alpha_{Max}$  katsayısının diğer katsayılara göre daha yüksek olduğu gözlenmiştir.

Çizim 2'ye göre; tek faktörlü veri kümelerinde tüm katsayılar yaklaşık aynı değerler üretmiştir ( $\alpha_{Max} \cong \alpha_{STD} \cong \alpha_{VK} \cong \alpha_M \cong 0.93$ ). Ancak faktör sayısı arttıkça bu eşitlik bozulmaktadır ve özellikle 4 faktörlü veri kümelerinde alfa katsayılarının ortalama değerleri arasında ( $\alpha_{Max} > \alpha_{STD} > \alpha_{VK} > \alpha_M$ ) eşitsizlikleri ortaya çıkmaktadır. Burada vurgulanması gereken en önemli bulgunun ise Tablo 2'de görüldüğü gibi her 4 faktörde de  $\alpha_{Max}$  değerinin diğer alfa katsayılarından daha yüksek olduğu görülmüştür.

**Çizim 2: Farklı boyutlardaki örneklem veri kümelerinde alfa katsayılarının davranışı**

### 3.3. Ölçmelerin Yapısı ve Alfa Katsayıları:

Alfa katsayısı, ölçmelerin paralel, eşdeğer ya da eşbiçimli olduğu durumlarda güvenilirlik indeksini yansız bir şekilde kestirmektedir. Ancak alfa katsayısı konjenerik ölçmelerde gerçek güvenilirliğin altında bir değer ürettiği daha önce ifade edilmişti. Bu bölümde, araştırma kapsamına alınan dört farklı alfa katsayısının eşdeğer ve konjenerik ölçmelerdeki davranışı incelenerek konjenerik ölçmelerde en yüksek kestirime sahip alfa katsayısı araştırılmıştır. Bu amaçla; alfa katsayıları arasındaki fark elde edilmiştir.

$$\delta_{Max-VK} = \hat{\alpha}_{Max} - \hat{\alpha}_{VK} \quad (10)$$

$$\delta_{Max-STD} = \hat{\alpha}_{Max} - \hat{\alpha}_{STD} \quad (11)$$

$$\delta_{Max-M} = \hat{\alpha}_{Max} - \hat{\alpha}_M \quad (12)$$

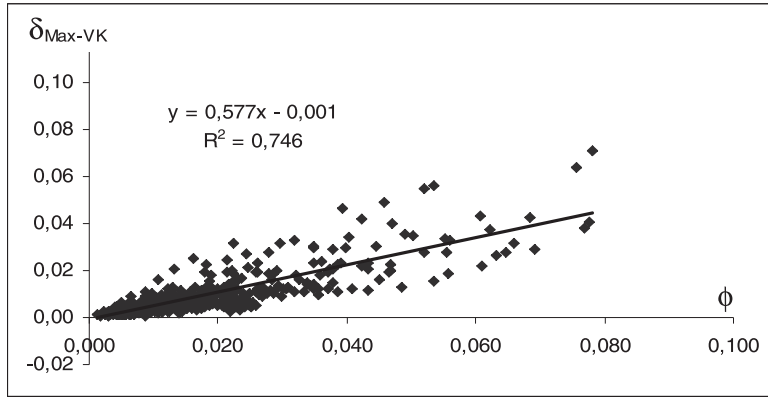
$$\delta_{STD-VK} = \hat{\alpha}_{STD} - \hat{\alpha}_{VK} \quad (13)$$

$$\delta_{\text{STD-M}} = \hat{\alpha}_{\text{STD}} - \hat{\alpha}_{\text{M}} \quad (14)$$

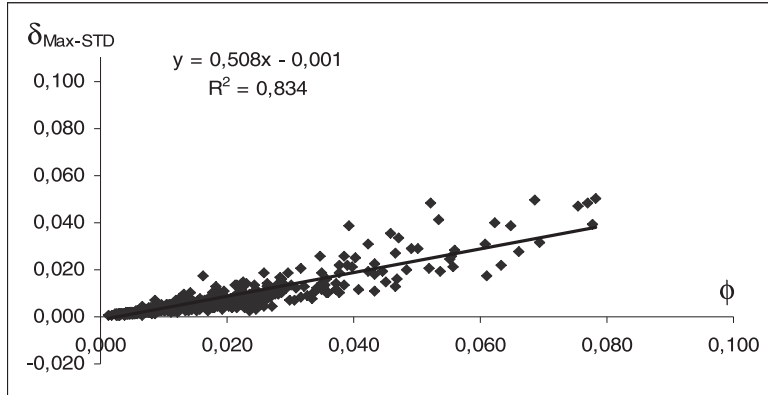
$$\delta_{\text{M-VK}} = \hat{\alpha}_{\text{M}} - \hat{\alpha}_{\text{VK}} \quad (15)$$

Bu katsayıların farkları Eşitlik 9'da verilen doğrulayıcı faktör analizinden elde edilen yüklerin varyanslarına ( $\phi$ ) göre ilişkisi incelenmiştir (Bkz: Dipnot 3).

Çizim 3'te  $\alpha_{\text{Max}}$  katsayısının diğer alfa katsayıları ile farklarının, faktör yüklerindeki sapmalara karşın davranışlarını içeren grafikler verilmiştir. Diğer katsayılar arasındaki karşılaştırma grafikleri ise Ek 1'de verilmiştir. Çizim 3'te görüldüğü gibi;  $\phi=0$  olduğu zaman  $\delta_{\text{Max-VK}}=\delta_{\text{Max-STD}}=\delta_{\text{Max-M}}=0$  olmaktadır. Bunun anlamı ise faktör yükleri eşit olduğu durumlarda (paralel veya eşdeğer ölçmeler)  $\alpha_{\text{Max}}$  katsayısı diğer katsayılar ile aynı değerleri almaktadır. Ancak faktör yüklerindeki sapmalar arttıkça,  $\phi>0$ , (konjenerik ölçmeler)  $\delta_{\text{Max-VK}}>0$   $\delta_{\text{Max-STD}}>0$   $\delta_{\text{Max-M}}>0$  olmaktadır; buna göre ölçme yapısına göre alfa katsayıların karşılaştırmalarında  $\alpha_{\text{Max}}>\alpha_{\text{VK}}$ ,  $\alpha_{\text{Max}}>\alpha_{\text{STD}}$  ve  $\alpha_{\text{Max}}>\alpha_{\text{M}}$  eşitsizlikleri gözlenmiştir. Diğer taraftan  $\phi$  ile  $\delta$  farkları arasındaki regresif bağıntı incelendiğinde veri kümesindeki konjenerik yapı arttıkça ( $\phi \rightarrow +\infty$ ) alfa katsayıları arasındaki en hızlı değişim  $\alpha_{\text{Max}}$  ile  $\alpha_{\text{M}}$  arasında gerçekleşmektedir.

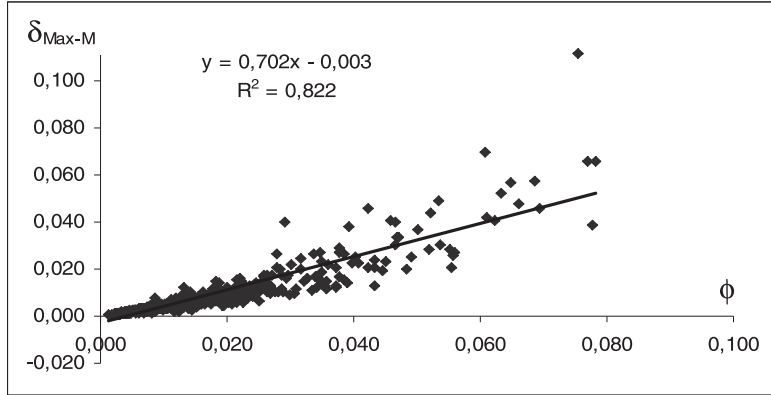


a)  $\phi$  ile  $\delta_{\text{Max-VK}}$  arasındaki bağıntı



b)  $\phi$  ile  $\delta_{\text{Max-STD}}$  arasındaki bağıntı



c)  $\phi$  ile  $\delta_{\text{Max-M}}$  arasındaki bağıntı

### Çizim 3: Konjenerik ölçmelerde $\alpha_{\text{Max}}$ katsayısının diğer alfa katsayılarına göre davranışı

#### 4. YORUM VE TARTIŞMA

Bu çalışmada klasik güvenilirlik katsayılarından olan alfa katsayısının dört farklı versiyonunu özellikleri karşılaştırılmıştır. Eğitimsel ve psikolojik araştırmalara ilişkin güvenilirlik literatüründe yalnızca  $\alpha_{\text{VK}}$ ,  $\alpha_{\text{STD}}$  ve  $\alpha_{\text{Max}}$  katsayıları büyüklükleri bakımından çeşitli karşılaştırmalar yapılmıştır. Bu araştırmalara göre;  $\alpha_{\text{VK}} \leq \alpha_{\text{STD}} \leq \alpha_{\text{Max}}$  türünden bir eşitsizlik ortaya konmuştur (Carmines & Zeller, 1979; Iacobucci & Duhachek, 2003; Osburn, 2000; Yurdugül, 2007). Bu araştırmada literatürdeki karşılaştırmalardan farklı olarak bu katsayıların istatistiksel özellikleri (sağlam kestiricilikleri), farklı boyutlardaki ve yapılardaki (eşdeğer ve konjenerik) davranışları incelenmiştir.

Farklı örneklemelerden elde edilen RMSE değerlerine göre en sağlam kestirici özelliğinin  $\alpha_{\text{Max}}$  katsayısında olduğu görülmüştür. Özellikle örneklem genişliği düşük olduğunda (N=30) en az hata içeren katsayı  $\alpha_{\text{Max}}$  katsayısı  $\text{RMSE}(\alpha_{\text{Max}}) < \text{RMSE}(\alpha_{\text{STD}}) < \text{RMSE}(\alpha_{\text{VK}}) < \text{RMSE}(\alpha_{\text{M}})$  iken, örneklem genişliği arttıkça (N=300, 500) her dört katsayının RMSE değerleri eşit olmaktadır.

Üzerinde çalışılan veri kümesinin boyut sayısına göre her dört alfa katsayısının nicel büyüklükleri ele alındığında; tekboyutlu veri kümelerinde dört katsayısının da değerlerinin yaklaşık eşit olduğu ( $\alpha_{\text{Max}} = \alpha_{\text{STD}} = \alpha_{\text{VK}} = \alpha_{\text{M}} = 0.93$ ) gözlenmiş iken, boyut sayısı arttıkça  $\alpha_{\text{Max}}$  katsayısının daha büyük değerler ürettiği gözlenmiştir ( $\alpha_{\text{Max}} \geq \alpha_{\text{STD}} \geq \alpha_{\text{VK}} \geq \alpha_{\text{M}}$ ).

Veri kümelerinde yer alan ölçmelerin yapıları göz önüne alındığında ise; paralel ve eşdeğer ölçmeler için  $\alpha_{\text{Max}}$  katsayısının değerleri ile diğer katsayıların değerleri yaklaşık aynı çıkarken, ölçmeler konjenerik yapıya yöneldiğinde  $\alpha_{\text{Max}}$  katsayısı diğer katsayıya göre daha büyük değerler üretmektedir.

Diğer taraftan, karşılaştırmalar standartlaştırılmış ve standartlaştırılmamış alfa katsayıları açısından ele alındığında; her ne kadar Iacobucci, ve Duhachek (2003) olarak standartlaştırılmış alfa katsayılarının rapor edilmesini öneri olarak ifade etmesine karşın küçük örneklem genişliğinde (Çizim 1) standartlaştırılmamış alfa katsayısının daha sağlam kestirici olduğu görülmektedir. Bununla birlikte veri kümesi tekboyutlu yapılardan çokboyutlu bir yapıya yöneldiğinde ise standartlaştırılmış alfa katsayısı güvenilirliği daha büyük değerler ile kestirmektedir (Çizim 2). Ancak ölçmeler paralel ölçmelerden konjenerik ölçmelere yöneldiğinde ise bu iki katsayının belirgin bir üstünlüğü ortaya çıkmamaktadır (Ek 1).

Son olarak, tüm karşılaştırmalarda en olumsuz özelliğe sahip alfa katsayısının McDonald (1985) tarafından tanımlanan  $\alpha_{\text{M}}$  katsayısı olduğu gözlenmiştir. Bunun olası nedenlerinden birisi ise  $\alpha_{\text{M}}$  katsayısının (Eşitlik 7) faktör yüklerinin ilgili ölçmenin varyansının kestirimine karşılık gelmesi ve böylelikle kestirilen ölçme varyansının doğrulayıcı faktör analizi tarafından gerçek ölçme varyansından yanlı olarak kestirilmesi olabilir.

## 5. SONUÇLAR VE ÖNERİLER

Bu çalışmada dört farklı alfa katsayısının optimallikleri üzerine üç aşamalı bir karşılaştırma yapılmıştır. Büyük örneklem genişliklerin, tekboyutlu ve paralel/eşdeğer veri kümelerinde tüm katsayılar özdeş olarak elde edilmiştir. Ancak düşük örneklem genişliğinde, çokboyutlu ve konjenerik ölçmelerde  $\alpha_{\text{Max}}$  katsayısı en iyi özelliklere sahip katsayı olarak belirlenmiştir. Bununla birlikte; en olumsuz katsayı ise  $\alpha_{\text{M}}$  katsayısı olarak bulunmuştur. Özellikle her üç aşamada ele alınan özellikler genel olarak ele alındığında araştırmacılara  $\alpha_{\text{Max}}$  katsayısının kullanımı önerilir.

### KAYNAKLAR

- Armor, D. J. (1974). *Theta reliability and factor scaling*. In Costner, H. L. (ed.), *Sociological Methodology*. Jossey-Bass, San Francisco. 17–50.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*, Sage, Beverly Hills, Calif.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98–104.
- Crano, W D., & Brewer, M. B. (1973). *Principles of research in social psychology*. New York: McGraw-Hill.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Cronbach, L. J. & Shavelson R. J. (2004). My current thoughts on coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64, 391-418.
- Cronbach, L. J., Schonemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement*, 25,29 1-3 12.
- 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.
- Gorusch, R. L. (1983). *Factor Analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Green, S. B., Lissitz, R. W., & Mulaik, S. (1977). Limitations of coefficient alpha as an index of text unidimensionality. *Educational and Psychological Measurement*, 37, 827–839.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255-282.
- Iacobucci, D., & Duhachek, A. (2003). Advancing alpha: Measuring reliability with confidence. *Journal of Consumer Psychology*, 13, 478-487.
- Jackson, P. H., & Agunwamba, C. C. (1977). Lower bounds for the reliability of total scores on a test composed of nonhomogeneous items: I. Algebraic lower bounds. *Psychometrika*, 42, 567-578.
- 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 uncorrelated error on coefficient alpha. *Applied Psychological Measurement*, 21, 337-348.
- Liu, Y., & Zumbo, B. D. (2007). The Impact of Outliers on Cronbach's Coefficient Alpha Estimate of Reliability: Visual Analogue Scales. *Educational and Psychological Measurement*, 67, 620-634.
- Lucke, J. F. (2005). The  $\alpha$  and  $\omega$  of congeneric test theory: An extension of reliability and internal consistency to heterogeneous tests. *Applied Psychological Measurements*. 29(1), 65-81.
- McDonald, R. (1985). *Factor analysis and related methods*. Hillsdale, N J: Erlbaum.
- 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.
- Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, 5, 343-355.
- Poon, W Y, Leung, K, and Lee, S. Y. (2002). The comparison of single item constructs by relative mean and relative variance. *Organizational Research Methods*, 5, 275-298.
- Rae. G. (2006). Correcting Coefficient Alpha for Correlated Errors: Is  $\alpha_K$  a Lower Bound to Reliability? *Applied Psychological Measuremet*, 30(1), 56-59.
- Raykov, T. (1997). Scale reliability, Cronbach's coefficient alpha, and violations of essential tauequivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329-353.
- Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement*, 22, 375-385.
- Raykov, T. (2001). Bias of coefficient  $\alpha$  for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25, 69-76.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8, 350–353.

- Scholz, U., Gutiérrez Doza, B., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment, 18*, 242-251.
- Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy Scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-Nelson.
- Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72a101.
- Traub, E. R. (1994). *Reliability for the social sciences: Theory and Applications*. Measurement methods for the social sciences. Sage Publications, 1994.
- van Zyl, J. M., Neudecker, H., & Nel, D. G. (2000). On the distribution of the maximum likelihood estimator of Cronbach's alpha. *Psychometrika, 65*, 271-280.
- Yuan, K., & Bentler, P. M. (2002). On robustness of the normal-theory cased asymptotic distributions of three reliability coefficient estimates. *Psychometrika, 67*, 251-259.
- 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)*, 21-47
- Yurdugül, H. (2006). Paralel, eşdeğer ve konjenerik ölçmelerde güvenirlik katsayılarının karşılaştırılması. *A.Ü. Eğitim Bilimleri Fakültesi Dergisi, 39(1)*, 15-37.
- Yurdugul, H. (2009). SIMREL: A Software for the coefficient alpha and its confidence intervals with Monte Carlo studies, *Applied Psychological Measurement, 33* (4), 325-328.
- Zimmerman, D. W., Zumbo, B.D. & Lalonde, C. (1993). Coefficient Alpha as an estimate of test reliability under violation of two assumptions. *Educational and Psychological Measurement, 53* (1), 33-49.
- Zinbarg, R. E., Revelle, W., Yovel, I. & Li, W. (2005). Cronbach's  $\alpha$ , Revelle's,  $\beta$  and McDonalds  $\omega$ : their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70(1)*, 1-11.

### EXTENDED ABSTRACT

Reliability in educational and psychological literature can be expressed in terms of stability, equivalence, and consistency. Unlike test-retest for stability and alternate form for equivalence, only a single test is needed for estimating internal consistency. Coefficient alpha suggested by Guttman (1945) and Cronbach (1951) is one of the most commonly used methods in the educational and psychological studies for evaluating internal consistency as an indicator of test quality. Cronbach (2004) reported that his article published in 1951 had been cited no less than 5590 times and, in recent years, had been cited approximately 325 times per year in the Social Sciences Citation Index (Liu, & Zumbo, 2007).

However, in the related literatures, four different coefficient alphas are defined. The coefficient alphas were ranked as: a) *unstandardized alpha* ( $\alpha_{VK}$ ) based on variance-covariance matrix of data set, b) *standardized alpha*  $\alpha_{STD}$  based on correlation matrix of data set, c) *maximized alpha* ( $\alpha_{Max}$ ) based on the first eigenvalue obtained principal component analysis (Armor, 1974), and d) *estimated alpha* ( $\alpha_M$ ) by McDonald (1985) based on confirmatory factor analysis.

Which coefficient alpha should we use in the calculating reliability of measurement tools? In this study, the optimality of four different coefficient alphas defined above was investigated. For this aim of the study, the research design was conducted within three stages: a) Firstly, the robust estimator properties of coefficient alphas were taken into consideration. The robust estimator properties of those were investigated by root mean square error (RMSE) obtained from Monte-Carlo simulations for different sample sizes (N=30, 100, 300, 500), b) secondly the comparison of coefficient alphas' values obtained from different dimensional data set, and c) finally comparisons of coefficient alphas' performance in different data sets which include tau-equivalent and congeneric measures. The test of whether the measures are tau-equivalent (i.e., the factor loadings are equal) or congeneric (i.e., the factor loadings are not equal) measures, was investigated in deviation of unstandardized factor loadings given in Equation 9.

The investigations in the second (b) and third (c) stages were taken into consideration as weakness of coefficient alpha. As known, the coefficient alpha yields a low boundary of real reliability in multidimensional and/or congeneric data set. A necessary and sufficient condition for coefficient alpha to equal the reliability of the test score is that the measures are parallel, tau-equivalent, and

essentially tau-equivalent. An essentially tau-equivalent model includes the essentially tau equivalent measures in which simply the factor loadings are equal for all of measures. The model involves that the off-diagonal elements in the covariance matrix are all equal, but that the variances of measures are not equal for all measures. The congeneric measures may conform to a single-factor model with distinct factor loadings or conform to a multi-factor model. Therefore, this study deals with the coefficient alphas' performance in dimensionality of data set and nature of measures separately. Due to the bias of coefficient alpha under those conditions, the comparisons among coefficient alphas were included in this study. The data set in this study was obtained by 'the general self-efficacy scale' which developed by Schwarzer and Jerusalem (1995). The psychometric properties of data set were reported by Scholz, Gutiérrez Doña, Sud, & Schwarzer (2002). The scale includes 10 items and the data set includes 19120 observations from 25 countries. The data set was used as a population data and several repeated samples were drawn from those population data and those samples were used in Monte-Carlo simulations of this study.

According to results of Monte-Carlo simulation, the smallest RMSE value of coefficient alphas was observed for ( $\alpha_{Max}$ ) in the smallest sample size ( $N=30$ ), and the ordering of RMSE values in  $N=30$  are  $RMSE(\alpha_{Max}) < RMSE(\alpha_{STD}) < RMSE(\alpha_{VK}) < RMSE(\alpha_M)$ . In other sample size ( $N=100, 300, 500$ ) all RMSE values were approximately the same (Table 1 and Figure 1).

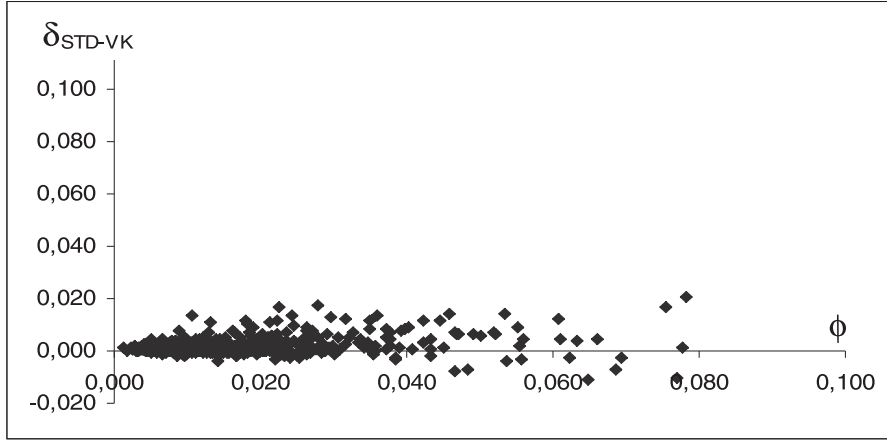
According to arithmetical means of coefficient alphas' values in unidimensional data sets, all of the coefficient alpha were approximately equal ( $\alpha_{Max} = \alpha_{STD} = \alpha_{VK} = \alpha_M = 0,93$ ). From single-factor data sets to multi-factor data sets, all of the coefficient alphas' were decreasing. However, the slowest decrease was observed for coefficient  $\alpha_{Max}$  and the slowest decrease was observed for coefficient  $\alpha_M$ . Specially in the four factor data sets, the relation of ( $\alpha_{Max} > \alpha_{STD} > \alpha_{VK} > \alpha_M$ ) for means of alphas' values was observed (Table 2 and Figure 1).

In stage (c); the coefficient alphas were compared in terms of measures' nature such as parallel, tau-equivalent, and congeneric. As known, the coefficient alpha underestimates the true reliability in congeneric measures. When those four alpha values were compared in congeneric measures, it is observed that the coefficient  $\alpha_{Max}$  yields the highest value. The lowest value was yielded by the coefficient  $\alpha_M$ . According to the results in this study, it is suggested that the coefficient  $\alpha_{Max}$  to be used in psychological and educational researches.

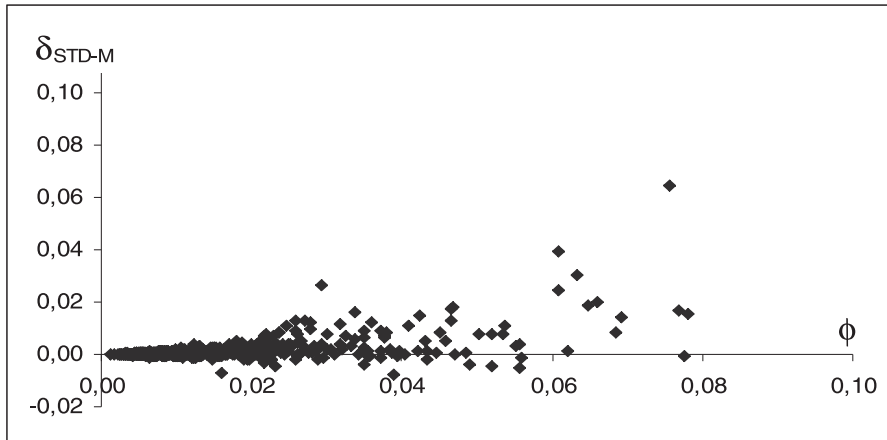


EK:

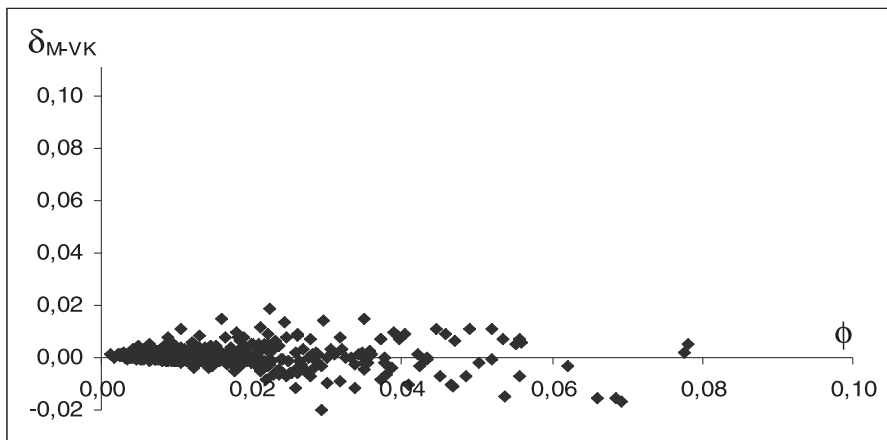
### Ölçmelerin Yapılarına Göre Alfa Katsayılarının Karşılaştırmaları



Yük Sapmalarına ( $\phi$ ) karşın  $\alpha_{STD}$  ve  $\alpha_{VK}$  farkının ( $\delta_{STD-VK}$ ) davranışı



Yük Sapmalarına ( $\phi$ ) karşın  $\alpha_{STD}$  ve  $\alpha_M$  farkının ( $\delta_{STD-M}$ ) davranışı



Yük Sapmalarına ( $\phi$ ) karşın  $\alpha_M$  ve  $\alpha_{VK}$  farkının ( $\delta_{M-VK}$ ) davranışı