

TARTIŞMAYI SONLANDIRMAK: CRONBACH'IN ALFA KATSAYISI, İKİ DEĞERLİ [0,1] ÖLÇÜMLENMiŞ MADDELER İLE KULLANILABİLİR

Vahit BADEMCİ*

Özet

Ölçümlerin güvenilirliğini kestirmek için değişik yöntemler vardır. Bunlardan birisi de Cronbach'ın alfa katsayısı yöntemidir. Cronbach'ın alfa katsayısı ile ilgili olarak da, bu katsayının sadece ağırlıklı ölçümlenmiş (örneğin, 1,2,3,4,5 gibi) maddelerle kullanılabileceği şeklinde bir görüş vardır. Ancak, bazı bilim adamlarınca ileri sürülen bu görüş doğru değildir. Çünkü, Cronbach'ın alfa katsayısı iki değerli (örneğin, 0,1 gibi) ölçümlenmiş maddelerle kullanılabilir.

Anahtar Sözcükler: Ölçüm Güvenirliği, Cronbach'ın Alfa Katsayısı

Abstract

To Put an End to the Discussion: Cronbach's Alpha Coefficient Can Be Used With Dichotomously Scored Items [0,1]

There are various methods to estimate the reliability of scores. One of them is the method of Cronbach's alpha coefficient. Related with the Cronbach's alpha coefficient, there is a view saying that this coefficient can be used only by weightily scored items (e.g., 1,2,3,4,5). However, this view suggested by some scientists is not true. Because, Cronbach's alpha coefficient can be used with dichotomously scored items (e.g., 0,1)

Key Words: Score Reliability, Cronbach's Alpha Coefficient

Güvenirlik, testin ya da ölçeğin bir özelliği değildir

Bir test [veya bir ölçek ya da bir ölçme aracı] güvenilir değildir; çünkü güvenilirlik, testin kendisinin değil, o testten elde edilen ölçümlerin bir özelliğidir (Bademci, 2001b; Bademci,2002; Bademci,2004; Bademci,2005a; Bademci,2005b; Bademci,2005c; Buhi, 2005; Caruso,2000; Caruso, Witkiewitz, Belcourt-Dittlof ve Gottlieb, 2001; Kieffer, 1999; Lane, White ve Henson, 2002; Thompson, 1994a; Thompson, 1994b; Thompson ve Vacha-Haase, 2000; Vacha-Haase,1998). Bir başka ifadeyle, "güvenilir veya güveniliriz olan testler değil, bir test veya ölçme aracından elde edilmiş ölçme sonuçları veya ölçümlerdir" (Bademci,2001b; Bademci, 2005b).

Güvenirlik, aracın kendisine değil, bir bellilendirme (assessment) aracı ile elde edilmiş ölçümlere işaret eder (Linn ve Gronlund, 2000; Linn ve Miller, 2005). Böylelikle, bir ölçme aracına [testin kendisine]*** işaret ettiği zaman söylenen "test**

* Yrd. Doç. Dr. Vahit Bademci, Gazi Üniversitesi, Endüstriyel Sanatlar Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Beşevler / Ankara. e-posta: bademci@gazi.edu.tr

güvenilirdir” veya “testin güvenilirliği” ifadelerini (Guthrie, 2000) kullanmak da doğru değildir, uygun değildir (Thompson,1994b; Thompson ve Vacha-Haase,2000). Daha doğru ve uygun ifade, ölçüm güvenilirliğidir ve doğru ifadelerin kullanılmasına yönelik olası örnekler, Bademci'nin (2004) bir çalışmasında bulunmaktadır. Bu bakış açısıyla, gözden kaçırılmaması gereken önemli bir nokta ise, güvenilir ölçümler ile güvenilir testler terimlerinin eş anlamlılıktan uzak olduğudur (Vacha-Haase, Kogan, Tani ve Woodall, 2001). Türkiye'deki profesör veya doçent unvanlı kimi öğretim üyelerinin, ölçme ile bağlantılı bu ve benzeri birtakım terimleri birbirlerinden ayırt edemedikleri de görülmektedir (Bademci, 2005c).

Yurt dışındaki çalışmalarda da yaygın kullanılan “test güvenilirdir” veya “testin güvenilirliği” ya da “aracın güvenilirliği” veya “ölçeğin güvenilirliği” gibi kısaltarak hatalı ifade etme biçimleri, Türk eğitim ve bilim topluluğunda da 1940'lerden bu yana yaklaşık 60 yılı aşkın bir süredir kullanılmaktadır (Bademci,2005a; Bademci,2005c) ve Bademci (2004) “Testin Güvenirliği” veya “Test Güvenilirdir” Diye İfade Etmek Doğru Değildir başlıklı bir çalışmasıyla, yaklaşık 60 yılı aşkın bir süredir süregelen bu doğru olmayan kullanım biçimine ve güvenilirliğin hatalı yorumlanış ve uygulama şekillerine de karşı çıkmıştır (Bademci, 2001b; Bademci,2002;Bademci,2004; Bademci,2005a; Bademci,2005b; Bademci,2005c). Güvenirliği daha doğru yorumlama ve de ifade etme biçimlerine örnekler veren Bademci (2004; 2005a), bilinenlerin aksine Cronbach alfa (α) katsayısının da, matematiksel olarak, negatif ve -1'den küçük olabileceğini (Bademci, 2001b; Bademci, 2002; Bademci,2005b), yine örnekleriyle [örneğin, $\alpha = -7$] ortaya koymuştur (Bademci, 2005a; Bademci,2005d).

Bademci (2005a), Cronbach alfa katsayısını da ağırlıklı içeren çalışmasında, Cronbach alfa katsayısıyla ilgili verdiği örnekte, iki değerli (dichotomously) ölçümlenmiş maddeler kullanmıştır. Bademci'nin (2005a) bu çalışmasından sonra kendisine sıklıkla yöneltilen sorulardan birisi de, Cronbach alfa katsayısının iki değerli [0,1] ölçümlenmiş maddeler ile kullanılıp, kullanılamayacağı yönündedir.

Türkçe Literatürdeki Mevcut Durum ve Bazı Örnekler

Cronbach alfa katsayısıyla ilgili olarak, Öncü (1994:49) “...Cronbach tarafından geliştirilen α (alpha) katsayısı yöntemi, maddeler doğru-yanlış olacak şekilde puanlanmadığında, kullanılması gereken bir iç tutarlılık tahmini metodudur. Puanlama kuralının (0,1) olmadığı durumlarda (1,3; 1,4; 1,5; gibi) Cronbach α katsayısını veren formül kullanılır” biçiminde bir ifade kullanmıştır. Bulduk (2003:76) ise, “*Kuder-Richardson güvenilirliği*...bir testin veya ölçeğin maddelerinin homojen olduğu ve sadece doğru/yanlış gibi iki seçeneğin olduğu testlerde kullanılır. *Cronbach alfası katsayısı* ise ikiden fazla cevap seçeneği bulunan testlerde kullanılabilen bir iç tutarlılık bulma formülüdür” şeklinde bir ifade de bulunmuştur. Erkuş (2003:69), daha kesin ifadelerle, “eğer test maddeleri 1-0 (dichotomously) şeklinde puanlanıyorsa, Kuder-Richardson; 1-5

gibi çoklu (polytomously) puanlanıyorsa Cronbach Alfa içtutarlılık katsayıları hesaplanmalıdır” biçiminde, görüşünü belirtmiştir. Tan ve Erdoğan (2004:183) ise, “alfa (α) güvenilirlik katsayısı... ağırlıklı puanlama veya dereceleme yöntemiyle puanlama uygulandığı [1’den 5’e kadar veya 0’dan 4’e kadar gibi] durumlarda kullanılacak bir güvenilirlik bulma tekniğidir” şeklinde bir ifade kullanmıştır. Büyüköztürk’de (2005:170-171), “test maddelerine verilecek cevapların doğru/yanlış, evet/hayır gibi iki seçeneği olması durumunda KR-20, pek çok kişilik testlerinde olduğu gibi üç veya daha fazla olması durumundan Cronbach tarafından geliştirilmiş olan alfa (α) katsayısı kullanılır” biçiminde görüşünü belirtmiştir. Yine bu hususla ilgili olarak, Tavşancıl’da (2005:28-29), kitabında, “ölçekteki sorulara ya da maddelere verilen cevapların doğru ya da yanlış olarak değerlendirilmediği, örneğin Likert ölçeğindeki gibi “Tamamen Katılıyorum”, “Katılıyorum”, “Kararsızım”, “Karşıyım”, “Tamamen Karşıyım” vb. gibi derecelendirilmiş tutum ve kişilik testlerinde, ilgi envanterlerinde Cronbach Alfa Katsayısı hesaplanır (Thorndike, Cunningham, Thorndike ve Hagen, 1991, s.318)” şeklinde bir ifade kullanmıştır. Tüm bu görüşler, evet-hayır, doğru-yanlış, katılıyorum-katılmıyorum gibi iki değerli [0,1] ölçümlenmiş maddelerle ilgili olarak KR-20’nin; çok değerli (polytomously) veya olası değerlerin bir dizi genişliğiyle ağırlıklı (örneğin, bir tutum ölçeği üzerindeki maddeleri 1’den 5’e -1,2,3,4,5- kadar) ölçümlenmiş maddelerle ilgili olarak da Cronbach alfa katsayısının kullanılması yönündedir; ki bu görüşler, yani Cronbach alfa katsayısının sadece ağırlıklı [1,2,3,4,5 gibi] ölçümlenmiş maddelere rezerv edilmesi yönündedir ve bu doğru değildir ve de tüm bu yargıların bilimsel olarak da, klasik test kuram modeli (Lord ve Novick, 1968) açısından da herhangi bir dayanağının bulunduğu görülememektedir.

Cronbach alfa katsayısının evet-hayır [0,1] gibi [iki yanıtı olarak] ölçümlenmiş maddelere sahip ölçme araçlarının ölçüm güvenliğinin kestirilmesinde kullanılmayacağına dair benzeri hatalı görüşler öyle etkili olmuştur ki, araştırmacıları kendi yaptıkları işlerden dahi kuşku duyar hale getirmiştir. Örneğin, Boratav (2003), kısaltılmış adı *Bordepta* olan *Boratav Depresyon Tarama Ölçeği*ni geliştirme sürecinde, Ölçek ölçüm güvenliğini kestirirken [uygun biçimde] Cronbach alfa katsayısı yöntemini kullanmıştır. Ancak ve de ilginç olan, Boratav’ın (2003:180) kendi araştırmasında kullandığı Cronbach alfa katsayısı yöntemi ile ilgili olarak, “burada tartışmaya açık olan konu bir evet-hayır ölçeği olan *Bordepta* ölçeğinde Cronbach alfa yönteminin uygulanmış olmasıdır” şeklinde de bir ifade kullanmış olmasıdır; doğru yapılan bir işlemde dahi kuşku duyulur hale gelmesi, yerinde kullanılan bir güvenilirlik kestirim yönteminin tartışmaya açık hale getirilmesi, Cronbach alfa katsayısı ile ilgili -gittikçe çoğalan- bazı mitlerin [hayali öykülerin] araştırmacılar üzerinde ne kadar etkili olduğunun bir göstergesi olarak da düşünülebilir.

Gerçek şudur ki, *Cronbach alfa katsayısı, ağırlıklı* (örneğin, kesinlikle katılım, katılım, kararsızım, reddederim, kesinlikle reddederim gibi 1’den 5’e kadar [-1,2,3,4,5-] [çok değerli]) *ölçümlenmiş maddelerle* veya *iki değerli ölçümlenmiş* (örneğin, evet-hayır, doğru yanlış, uygun-uygun değil, başarılı-başarısız, katılım-katılmam [0,1] gibi) *maddelerle kullanılabilir*.

Kısaca, Cronbach alfa katsayısı, test maddelerinin tüm çeşitleriyle kullanılabilir (Worthen, White, Fan ve Sudweeks,1999).

Cronbach'ın alfa katsayısı, iki değerli [0,1] ölçümlenmiş maddelerle kullanılabilir

Test ölçümlerinin güvenilirliğini kestirmenin yöntemlerinden birisi, Cronbach'ın (1951) alfa katsayısı yöntemidir (Mehrens ve Lehmann,1991; Linn ve Gronlund,2000). Alfa katsayısı, aşağıdaki formül yoluyla hesaplanmaktadır (Crocker ve Algina,1986; Reinhardt, 1996; Worthen, White, Fan ve Sudweeks,1999).

$$\alpha = k / (k-1) * [1 - (\Sigma \sigma_i^2 / \sigma_T^2)] \quad (\text{Formül 1})$$

k = test üzerindeki madde sayısı
 σ_i^2 = i madde ölçüm varyansı [ya da bir madde üzerindeki bir grup bireyden elde edilen ölçümlerin varyansı]
 $\Sigma \sigma_i^2$ = i madde ölçüm varyanslarının toplamı
 σ_T^2 = toplam test ölçümlerinin varyansı

Yine, test ölçümlerinin güvenilirliğini kestirmenin yöntemlerinden bir diğeri de, Kuder Richardson yöntemidir ve Kuder ve Richardson (1937:158) tarafından türetilen ve KR-20 şeklinde yaygın olarak bilinen ve kullanılan formül, aşağıdaki şekilde de yazılabilmektedir (Crocker ve Algina,1986; Henson,2001; Worthen, White, Fan ve Sudweeks,1999).

$$\text{KR-20} = k / (k-1) * [1 - (\Sigma p_i q_i / \sigma_T^2)] \quad (\text{Formül 2})$$

k = test üzerindeki madde sayısı
 p_i = i maddesini doğru cevaplayan kişilerin oranı
 q_i = i maddesini doğru cevaplayamayan kişilerin oranı
 $p_i q_i$ = i madde ölçüm varyansı [ya da bir madde üzerindeki bir grup bireyden elde edilen ölçümlerin varyansı]
 $q_i = 1 - p_i$
 $\Sigma p_i q_i$ = i madde ölçüm varyanslarının toplamı
 σ_T^2 = toplam test ölçümlerinin varyansı

Formül 1 ile Formül 2 alt alta tekrar yazıldığında, madde ölçüm varyansının türetimi hariç, KR-20 formülü ile Cronbach alfa katsayısı formülünün aynı olduğu görülebilecektir.

$$\alpha = k / (k-1) * [1 - (\Sigma \sigma_i^2 / \sigma_T^2)] \quad ($$

Formül 1)

$$\text{KR-20} = k / (k-1) * [1 - (\Sigma p_i q_i / \sigma_T^2)] \quad (\text{Formül 2})$$

Tüm maddeler iki değerli [0,1] ölçümlenmiş olduğu zaman, σ_i^2 matematiksel olarak $p_i q_i$ 'ye eşittir (Suen, 1990), yani $\sigma_i^2 = p_i q_i$. Buradan, tüm maddeler iki değerli [0,1]

ölçümlenmiş olduğunda, $\Sigma\sigma^2_i = \Sigma p_i q_i$ 'de yazılabilir (Worthen, White, Fan ve Sudweeks, 1999).

O halde, tüm maddeler iki değerli [0,1] ölçümlenmiş olduğu zaman Cronbach'ın alfa katsayısı,

$$\alpha = k / (k-1) * [1 - (\Sigma p_i q_i / \sigma^2_T)] \quad (\text{Formül 3})$$

şeklinde de, yeniden yazılabilir.

Yine tüm maddeler iki değerli [0,1] ölçümlenmiş olduğunda, Formül 2 ile Formül 3'ün eşit olduğu, alt alta yazıldığında, daha da açık görülebilecektir.

$$KR-20 = k / (k-1) * [1 - (\Sigma p_i q_i / \sigma^2_T)] \quad (\text{Formül 2})$$

$$\alpha = k / (k-1) * [1 - (\Sigma p_i q_i / \sigma^2_T)] \quad (\text{Formül 3})$$

Formül 2 ve Formül 3'den görüleceği üzere, *tüm maddeler iki değerli [0,1] ölçümlenmiş olduğu zaman, $\alpha = KR-20$ 'dir.*

Sonuç

KR-20'nin, *yalnızca* iki değerli [0,1] ölçümlenmiş maddelere uygun olduğu bilinmektedir (Crocker ve Algina, 1986; Henson,2001; Reinhardt,1996; Worthen, White, Fan ve Sudweeks,1999). Ancak, iki değerli [0,1] ölçümlenmiş maddelerle sadece KR-20 değil, Cronbach'ın alfa katsayısı da kullanılabilir.

Erkuş (2000:16), bir başka çalışmasında bu hususla ilgili olarak, “bilindiği gibi, maddeleri çoklu puanlanan (polytomously) homojen ölçek ya da alt ölçeklerin iç tutarlılık katsayılarını hesaplarken Cronbach alfa, ikili (dichotomously) puanlanan homojen ölçek ya da ölçeklerin iç tutarlılıklarını hesaplarken de Kuder-Richardson-20 (veya bazı sayıtların karşılanmasıyla 21) formülünü kullanmak gerekir (Crocker ve Algina, 1986)” şeklinde ifadeler kullanmış ve kullandığı bu ifadeler için de Crocker ve Algina'nın (1986) kitabını referans göstermiştir. Ancak, Erkuş (2000) tarafından kullandığı ifadelere referans gösterilen Crocker ve Algina'ya (1986) ait kitabın 138. sayfasında, Cronbach alfa katsayısı ile ilgili olarak, Erkuş'un (2000) söylediği ifadenin karşıtı bir ifade bulunmaktadır. Cronbach alfa katsayısıyla ilgili olarak, Crocker ve Algina'nın (1986:138) kitabındaki ifade, “[Cronbach] alfa [katsayısı], yazılı sınavlar veya bazı tutum envanterlerinde olduğu gibi, ölçümleme ağırlıklarının geniş bir yayılım genişliğine sahip [bir başka ifadeyle çok değerli] maddeler[in] veya iki değerli ölçümlenmiş maddelerin iç tutarlılığının kestirilmesinde kullanılabilir” şeklindedir. Görüldüğü üzere, Erkuş'un (2000) ifadelerinin aksine, Crocker ve Algina (1986:138), Cronbach alfa katsayısının iki değerli [0,1] ölçümlenmiş maddelerin iç tutarlılığının kestirilmesinde kullanılabildiğini belirtmektedir.

Sonuç olarak, iki değerli [0,1] ölçümlenmiş maddelerle Cronbach alfa katsayısı kullanılabilir (Crocker ve Algina, 1986; Henson, 2001; Nitko, 2001;Reinhardt,1996; Worthen, White, Fan ve Sudweeks, 1999). Zira, eğer maddeler 0 ve 1 şeklinde ölçümlenmişse, KR-20 ve Cronbach alfa katsayısı aynı [özdeş] sonuçlar verir (Sax, 1997). Bir başka ifadeyle, iki değerli [0,1] ölçümlenmiş bir testten elde edilen ölçümlerin güvenilirliğini kestirmede [Cronbach] alfa katsayısı kullanıldığı zaman, sonuç, KR-20 kullanılarak hesaplanmış olanla tamamen aynı olacaktır (Ebel ve Frisbie, 1991).

Yorum

Bundan 55 yıl önce de, yani alfa katsayısıyla ilgili özgün makalesinde Cronbach (1951:299), tüm maddeler 0 veya 1 şeklinde ölçümlenmiş olduğu zaman, kendi formülünün [α], KR-20 [yani α 'nın özel durumu] haline geldiğini ifade etmiştir; KR-20 ve α katsayısı ile ilgili Türkçe literatür genişlemesine incelendiğinde, Cronbach'ın (1951:299) bu önemli ifadesinin çoğu bilim insanınca gözden kaçırıldığı veya Cronbach'ın (1951) özgün makalesinin daha dikkatlice ya da hiç okunulmadığı düşünülmektedir. Tüm maddeler iki değerli ölçümlenmiş olduğu zaman, [Cronbach] alfa katsayısının, KR-20 haline geldiği, Thorndike'in (1982) ve Feldt ve Brennan'ın (1989) çalışmalarında da tekrar belirtilmiştir.

Ayrıca, yine Cronbach (1990:202-204), *Essentials of Psychological Testing* adlı ve bir klasik olarak nitelendirilebilecek kitabında da, 1 veya 0 şeklinde ölçümlenmiş [maddelerden oluşan] bir grup veri ile ilgili ölçüm güvenirligi hesaplama [kestirme] örnekleri vermiş, 1 veya 0 şeklinde ölçümlenmiş bu maddelere, hem KR-20, hem de α katsayısı [formülleri] kullanmış ve de 1 veya 0 şeklinde ölçümlenmiş bu maddelerle hesaplanmış KR-20 ile α katsayısının [ve de "varyans[ın] bileşenlerinden" hesaplanan α_k 'nin] aynı [özdeş] sonuçları verdiğini hem belirtmiş, hem de örnekleri üzerinde göstermiştir.

Kısacası ve açıkça, güvenilirlik kestirimleri ile ilgili çalışmalarda, Cronbach'ın alfa katsayısı, yalnızca iki değerli [0,1] ölçümlenmiş maddelerle veya ağırlıklı [1,2,3,4,5] ölçümlenmiş maddelerle kullanılabilir.

**** Ölçüm; (score).**

M. Fuat Turgut, ölçme işlemleri sonunda elde edilen sayılara *ölçüm* denilmesini önermektedir (Bademci,1999:7-8).

Ölçümleme; (scoring). (Bademci,2004:370).

Bellilendirme; (assessment). (Bademci, 2001a: 204).

*** Metin içindeki [...] arasındaki ifadeler yazar tarafından eklenmiştir.

KAYNAKLAR

- Bademci, V. (1999). *Hedefin Davranışlara Çevrilmesi, Davranışlardan Seçmeli Test Maddeleri Yazılması*. (Geliştirilmiş Üçüncü Baskı). Ankara: Gazi Kitabevi.
- Bademci, V. (2001a). *Türkiye'deki Okullar Ne İşe Yarar?* (Geliştirilmiş İkinci Baskı). Ankara: Alp Yayınevi.
- Bademci, V. (2001b). "Düşünmenin Öğretilmesi ve Öğretimde Kullanılan Yöntemler – Teknikler." Düzenleyen: TÜRMÖB. Bursa: Bursa SMMM Odası Konferans Salonu, 9 Kasım.
- Bademci, V. (2002). "Türkiye'deki Okullar Ne İşe Yarar? Türkiye'nin Anomi, Yabancılaşma, Ekonomik Büyüme, Demokratikleşme Sorunlarına Çözüm Önerisi." Düzenleyen: ESEF Öğrenci Bilimsel Faal. Org. Kom. Ankara: G.Ü. Mesleki Eğitim Fakültesi Konferans Salonu, 30 Mayıs.
- Bademci, V. (2004). "Testin Güvenirliği" veya "Test Güvenilirdir" Diye İfade Etmek Doğru Değildir. *Türk Eğitim Bilimleri Dergisi*, Cilt 2 , 367-372.
- Bademci, V. (2005a). Araştırmalarda Ölçme İle İlgili Bazı Büyük Hataları Düzeltmek ve Bir Reformu Başlatmak: Güvenirlik, Testlerin Bir Özelliği Değildir. *Eğitim Fakültelerinde Yeniden Yapılandırmanın Sonuçları ve Öğretmen Yetiştirme Sempozyumunda Sunulan Bildiri*. Ankara: Gazi Üniversitesi, Gazi Eğitim Fakültesi, 22-23-24 Eylül.
- Bademci, V. (2005b). Testler Güvenilir Değildir: Ölçüm Güvenirliğine Yeterli Dikkat ve Güvenirlik Çalışmaları İçin Örneklem Büyüklüğü. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*, Sayı 17, 33-45.
- Bademci, V. (2005c). Hakemlerin Değerlendirmelerindeki Hatalar Üzerine: Fisher'in Z Dönüşümü ve Güvenirlik Çalışmaları İçin Örneklem Büyüklüğü. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*, Sayı 17, 46-75.
- Bademci, V. (2005d). Güvenirliği Doğru Anlamak ve Bazı Klişeleri Yıkma : Bilinenlerin Aksine, Cronbach'ın Alfa Katsayısı, Negatif ve -1'den Küçük Olabilir. *Yayıma Hazırlanmış Makale*.
- Boratav, C. (2003). Boratav Depresyon Tarama Ölçeği (Bordepta): Epidemiyolojik Çalışmalar ve Birinci Basamak Sağlık Hizmetinde Depresyonu Tanımaya Duyarlı Bir Ölçek. *Türk Psikiyatri Dergisi*, Cilt 14 (3), 172-183.
- Buhi, E.R. (2005). Reliability Reporting Practices in Rape Myth Research. *Journal of School Health*, Vol 75 (2), 63-66.
- Bulduk, S. (2003). *Yeni Başlayanlar İçin Deneysel Psikolojide Araştırma Yöntemleri*. İstanbul: Çantay Kitabevi.
- Büyüköztürk, Ş. (2005). *Sosyal Bilimler İçin Veri Analizi El Kitabı*. (Gözden Geçirilmiş 5. Baskı). Ankara: PegemA
- Caruso, J.C. (2000). Reliability Generalization of the Neo Personality Scales. *Educational and Psychological Measurement*, Vol. 60, 236-254.
- Caruso, J.C., Witkiewitz, K., Belcourt-Dittloff, A. & Gottlieb, J.D. (2001). Reliability of

- Scores from the Eysenck Personality Questionnaire: A Reliability Generalization Study. *Educational and Psychological Measurement*, Vol. 61, 675-689.
- Crocker, L. & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Fort Worth: Holt, Rinehart and Winston.
- Cronbach, L. J.(1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, Vol. 16, 297-334.
- Cronbach, L.J. (1990). *Essentials of Psychological Testing*. (Fifth Edition). New York: HarperCollins.
- Ebel,R.L. & Frisbie, D.A. (1991). *Essentials of Educational Measurement*. (Fifth Edition). Englewood Cliffs, New Jersey: Prentice Hall.
- Erkuş, A. (2000). Sık Kullanılan Bazı Psikolojik Ölçeklerin Güvenirliklerinin İrdelenmesi. *Türk Psikoloji Yazıları*, Cilt 3 (6), 3-17.
- Erkuş, A. (2003). *Psikometri Üzerine Yazılar*.(1. Basım). Ankara: Türk Psikologlar Derneği Yayınları,24
- Feldt,L.S. & Brennan,R.L. (1989). Reliability. *Educational Measurement*. (Third Edition). Ed.R.L. Linn. New York: American Council on Education ve Macmillan
- Guthrie, A.C. (2000). *A Review of Coefficient Alpha and Some Basic Tenets of Classical Measurement Theory*. (ERIC Document Reproduction Service No. ED 438 307).
- Henson, R. K. (2001). Understanding Internal Consistency Reliability Estimates: A Conceptual Primer on Coefficient Alpha. *Measurement and Evaluation in Counseling and Development*, Vol. 34, 177-189.
- Kieffer, K.M. (1999). Why Generalizability Theory is Essential and Classical Test Theory is Often Inadequate. *Advances in Social Science Methodology, Volume 5*. Ed. B. Thompson. Stamford,Connecticut: JAI Press .
- Kuder, G.F. & Richardson, M.W. (1937). The Theory of the Estimation of Test Reliability. *Psychometrika*, Vol. 2, 151-160
- Lane, G.G., White, A.E. & Henson, R.K. (2002). Expanding Reliability Generalization Methods with KR-21 Estimates: An RG Study of the Coopersmith Self-Esteem Inventory. *Educational and Psychological Measurement*, Vol. 62, 685-711.
- Linn, R. L. & Gronlund, N.E. (2000). *Measurement and Assessment in Teaching*. (Eight Edition). Upper Saddle River, New Jersey : Merrill.
- Linn, R.L. & Miller, M. D. (2005). *Measurement and Assessment in Teaching*. (Ninth Edition). Upper Saddle River, New Jersey: Pearson
- Lord, F.M. & Novick, M.R. (1968). *Statistical Theories of Mental Test Scores*. Reading, Massachusetts: Addison-Wesley.
- Mehrens, W. A. & Lehmann, I. J. (1991). *Measurement and Evaluation in Education and Psychology*. (Fourth Edition). Fort Worth : Harcourt Brace.
- Nitko, A.J. (2001). *Educational Assessment of Students*. (Third Edition). Upper Saddle River, New Jersey: Merrill/Prentice Hall.
- Öncü, H. (1994). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Matser Basım.
- Reinhardt, B. (1996). *Factors Affecting Coefficient Alpha: A Mini Monte Carlo Study*.

- Advances in Social Science Methodology, Volume 4.* Ed. B. Thompson. Greenwich, Connecticut: JAI.
- Sax, G. (1997). *Principles of Educational and Psychological Measurement and Evaluation.* (Fourth Edition). Belmont, CA: Wadsworth.
- Suen, H.K.(1990). *Principles of Test Theories.* Hillsdale, New Jersey: Lawrence Erlbaum.
- Tan, Ş. & Erdoğan.A. (2004). *Öğretimi Planlama ve Değerlendirme.* (Genişletilmiş 5. Baskı). Ankara: PegemA
- Tavşancıl, E. (2005). *Tutumların Ölçülmesi ve SPSS ile Veri Analizi.* (2. Baskı). Ankara: Nobel Yayın Dağıtım.
- Thompson, B. (1994a). *It is Incorrect to Say "The Test is Reliable": Bad Language Habits Can Contribute to Incorrect or Meaningless Research Conclusions.* (ERIC Document Reproduction Service No. ED 367 707).
- Thompson, B. (1994b). Guidelines for Authors. *Educational and Psychological Measurement*, Vol. 54, 844-847.
- Thompson, B. & Vacha-Haase, T. (2000). Psychometric is Datametrics: The Test is Not Reliable. *Educational and Psychological Measurement*, Vol. 60, 174-195.
- Thorndike, R.L. (1982). *Applied Psychometrics.* Boston: Houghton Mifflin.
- Vacha-Haase, T. (1998). Reliability Generalization: Exploring Variance in Measurement Error Affecting Score Reliability Across Studies. *Educational and Psychological Measurement*, Vol. 58, 6-20.
- Vacha-Haase, T., Kogan, L.R., Tani, C.R. & Woodall, R.A. (2001). Reliability Generalization: Exploring Variation of Reliability Coefficients of MMPI Clinical Scales Scores. *Educational and Psychological Measurement*, Vol. 61, 45-59.
- Worthen, B. R., White,K.R., Fan,X & Sudweeks,R.R. (1999). *Measurement and Assessment in Schools.* (Second Edition). New York: Addison Wesley Longman.