Cronbach’s Alpha is not a Measure of Unidimensionality or Homogeneity

Cronbach’ın Alfası, Tek Boyutluğun veya Benzeşikliğin (Homojenliğin) Bir Ölçüsü Değildir

Vahit BADEMCİ

Abstract

Reliability is a characteristic of scores, not tests or measurement instruments. One of the methods to estimate the reliability of the scores on a test or measurement instrument is Cronbach’s coefficient alpha method. In other words, Cronbach’s alpha is an estimate of the score reliability based on the internal consistency among the [item] scores. In addition to this, Cronbach’s coefficient alpha should not be interpreted as a measure of the test’s or the measurement instrument’s unidimensionality or taken as an index of the homogeneity of a measurement instrument. Briefly, Cronbach’s alpha is not a measure of unidimensionality or homogeneity.

Keywords: Cronbach’s alpha, score reliability, misconception, unidimensionality, factor analysis

Öz


Anahtar kelimeler: Cronbach’ın alfası, ölçüm güvenirlüğü, yanlışlık (kavram yanlıştırması), tek boyutlu (bir boyutlu), faktör analizi (etken çözümlemesi)

Misconceptions and their overcome within the context of this article has been brought to agenda in Turkey by Bademci (2001; 2006a; 2006b; 2006c).

1 This article is a part of the declaration presented by the author during the conference “Paradigm shift: Tests are not reliable” given at Gazi University on 28 April 2006 [Conference, in Turkish]. [For your information; Gazi University. (2006). Gazi Haber, 66, 64. - in Turkish].

2 Gazi University, bademci@gazi.edu.tr
On Cronbach’s Alpha

One of the methods to estimate the reliability of test scores is Cronbach’s (1951) coefficient alpha (\(\alpha\)) method (Bademci, 2006a; 2006b; Linn & Miller, 2005; Mehrens & Lehmann, 1991; Worthen, White, Fan & Sudweeks, 1999). In other words, Cronbach’s alpha is an estimate of the score reliability based on the internal consistency among the [item] scores. Coefficient alpha is calculated by the following formula (Crocker & Algina, 1986; Henson, 2000; Reinhardt, 1996).

\[
\alpha = \frac{k}{k-1} \left[ 1 - \frac{\sum \sigma_k^2}{\sigma_T^2} \right]
\]

[ Formula 1 ]

\(k = \text{the number of items on the test}\)

\(\sum \sigma_k^2 = \text{the sum of the } k \text{ item score variances}\)

\(\sigma_T^2 = \text{the variance of the total test scores}\)

Coefficient alpha is affected by item difficulty, sum of item score variances, and total test score variance (Helms, 1999). Reinhardt (1996) demonstrated that total test score variance has the greatest effect on Cronbach’s coefficient alpha. Thompson (2003) also noted the variance of total test scores is not only a function of the item score variances, but also of the covariances of the item scores with each other. Smaller total test score variance leads to smaller coefficient alpha and bigger total test score variance leads the coefficient alpha to increase (Arnold, 1996; Helms, 1999).

Because, classical true score theory reliability estimates are affected by the total test score variance (Capraro, Capraro & Henson, 2001), also the total test score variance is greatly affected by how homogeneous or heterogeneous are the group of examinees (Helms, 1999).

If a test is given to a homogeneous group, the variability in total test score will decrease and then the alpha coefficient will be smaller; if the same test is given to a heterogeneous group, the variability in the total test score will increase and then the alpha coefficient will also increase (Arnold, 1996; Helms, 1999). In this situation, when the same
measurement instrument is given to the more homogeneous or the more heterogeneous groups (or samples), differing score reliability will occur (Thompson, 1994).

**Reliability is a Property of Scores, not Measurement Instruments**

To say in other words, the alpha coefficients, calculated from the scores obtained from applying the same measurement instrument to different samples, will take changing values according to the sample characteristics (Guthrie, 2000). Cronbach’s alpha can (and will) vary from sample to sample (Bademci, 2001; 2006b; Knapp, 1991). Bademci (2006b) emphasized, “reliability is greatly effected by the sample characteristics” (p. 7). Briefly, reliability is population or sample dependent (Bademci, 2001; 2006c; Mellenberg, 1996). From this point, it can be stated that reliability is not a characteristic of tests, but a property of the scores on a test or measurement instrument for a *particular* population or sample of examinees (Bademci, 2001; 2006c; Caruso, 2000; Crocker & Algina, 1986; Thompson, 1994; 1999; Wilkinson & APA Task Force on Statistical Inference, 1999).

Therefore, to accept reliability as a characteristic of tests or measurement instruments but not the scores, to act in the direction of this “unconscious paradigmatic belief” (Thompson, 1994, p. 839) in the researches and to speak of “the reliability of the test” or to say “the test is reliable” when referring to an instrument of measurement is *not convenient* and is *not correct* (Bademci, 2004; 2005; Guthrie, 2000). Reliability refers to the reproducibility or the consistency of the scores on a test or measurement instrument for a *particular* population or sample of examinees (Bademci, 2001; 2006c). In addition, the scores to be used in estimating reliability coefficients can be collected by a variety of methods (e.g., test-retest method, Cronbach’s coefficient alpha method).

Briefly, reliability is a characteristic of scores, not an inherent or a stable property of tests or measurement instruments. J. E. Helms (2005) also emphasized that

When researchers wrongfully assume that reliability is a stable property of tests, they tend to erroneously report the results of their reliability analyses as well. Some of these wrongful practices include (1) reliability induction, inferring reliability from previous studies rather than conducting one’s own reliability analyses; (2)
assuming that reliability coefficients are “inadequate” according to some specified or unspecified standard (e.g., .70) without conducting tests of significance or computing confidence intervals; (3) assuming homogeneity of sample characteristics such that all samples are supposed to respond in the same manner regardless of their characteristics; and (4) misconstruing the relationships between reliability and validity coefficients (p. 368).

Thompson (1994) noted that “Too few researchers act on a conscious recognition that reliability is a characteristic of scores or the data in hand” (p. 839). However, many people [still] do not understand score reliability (Mittag & Thompson, 2000; Thompson & Vacha-Haase, 2000; Vacha-Haase, Kogan, & Thompson, 2000; Vacha-Haase, Ness, Nilsson, & Reetz, 1999; Yin & Fan, 2000).

Cronbach’s Coefficient Alpha is not a Measure of Unidimensionality or Homogeneity

Pedhazur and Schmelkin (1991) noted, “when homogeneity is understood to mean unidimensionality, it means that, for an instrument to be viewed of as consisting of homogeneous items, it has to be demonstrated that a general factor is sufficient to account for the relations among them.” (p.102) Another settled idea about Cronbach’s alpha is to suggest the relatively high alpha value as the evidence of scale unidimensionality or homogeneity or to take Cronbach’s alpha as a measure of the homogeneity or the unidimensionality of a measurement instrument or test items (e.g., Bolton & Humphreys, 2002; Caci, Nadalet, Staccini, Myquel, & Boyer, 1999; Cramer & Barry, 1999; Kütçükdeveci, Sahin, Ataman, Griffiths, & Tennant, 2004; Marx, Bombardier, Hogg-Johnson, & Wright, 1999; Pike & Hudson, 1998; Radovanovic & Alexandre, 2004; Tiesinga, Dassen, & Halfens, 1998). But, this is not also true. Cronbach’s coefficient alpha is not a measure of unidimensionality or homogeneity (Bademci, 2001; 2006c; Cortina, 1993; Gerbing & Anderson, 1988; Hattie, 1985; Miller, 1995; Netemeyer, 2001; Netemeyer, Bearden, & Sharma, 2003; Rogers, Schmitt, & Mullins, 2002; Schmitt, 1996). Hattie (1985) pointed out, “Alpha is

Vahit BADEMCI

22
dependent on the length of the test, ... the unidimensionality of a test should be independent of its length” (p. 145). Streiner (2003) also remarked,

“It is true that the higher the correlations among the items of a scale, the higher will be the value of $\alpha$. But, the converse of this - that a high value of $\alpha$ implies a high degree of internal consistency - is not always true... Because $\alpha$ is affected by the length of the scale, high values do not guarantee internal consistency or unidimensionality” (p. 101, 103).

Cortina (1993) explained, “Internal consistency refers to the degree of interrelatedness among the items... whereas homogeneity refers to unidimensionality...” (p. 100). Consequently, Cronbach’s alpha should not be interpreted as a measure of the test’s or the measurement instrument’s unidimensionality or taken as an index of the homogeneity of a measurement instrument (Crocker & Algina, 1986; Pedhazur & Schmelkin, 1991).

**Examine the Unidimensionality**

Nunnally and Bernstein (1994) suggested that “Factor analysis is at the heart of the measurement of psychological constructs” (p. 111). Unidimensionality may be examined using exploratory factor analysis or especially confirmatory factor analysis (see Pedhazur and Schmelkin, 1991; Stevens, 2002; Thompson, 2004). But, Cronbach’s alpha should not be used as a measure of unidimensionality.

Netemeyer (2001) also noted “...alpha should be used to assess internal consistency only after unidimensionality is established” (p. 57). In other words, Cronbach’s alpha should be used to estimate of the score reliability based on the internal consistency among the [item] scores after unidimensionality is examined.

**References**


*Cilt 2 Sayı 3 (2014)*
Bademci, V. (2004). It is incorrect to speak of “the reliability of the test” or “the test is reliable”. *The Journal of Turkish Educational Sciences*, 2 (3), 367-372 [in Turkish].


Bademci, V. (2006a). To put an end to the discussion: Cronbach’s alpha coefficient can be used with dichotomously scored items [0,1]. *Journal of Kâzım Karabekir Education Faculty*, 13, 438-446 [in Turkish].

Bademci, V. (2006b). To understand reliability properly and to overthrow some clichés: Contrary to the known facts, Cronbach’s coefficient alpha can be negative and smaller than “–1″. *İnönü University Journal of the Faculty of Education*, 7(12), 3-26 [in Turkish].


