IS COEFFICIENT ALPHA CORRECTLY REPORTED? PRELIMINARY EVIDENCE FROM THE SCHOLARLY LITERATURE

Anie Aghababyan, Utah State University, 3555 Old Main Hill, Logan, UT, 84322, 435-754-6285, anie.aghababyan@gmail.com
Kenneth R. Bartkus, Utah State University, 3555 Old Main Hill, Logan, UT, 84322, 435-797-3891, ken.bartkus@usu.edu

ABSTRACT

Coefficient alpha is perhaps the most common measure of reliability in the social sciences. To ensure good quality research, therefore, it is important that alpha be accurately reported. This research explores reporting practices from a review of articles published in hospitality and tourism management journals. The results reveal that researchers often do not provide references when citing alpha. Additionally, it was found that alpha is sometimes not reported at all. Alpha is also sometimes reported as simply a range or reference by a range rather than individually. Finally, some researchers have misinterpreted reference citations. Guidelines for proper reporting of alpha are presented.

INTRODUCTION

Although numerous coefficients of reliability are available for use, coefficient alpha has been widely cited as the most common in social science research. Indeed, when it was first introduced, Cronbach (1951) referred to coefficient alpha as "a tool that we expect to become increasing prominent in the literature" [3, p. 299]. Since that time, references to alpha as a widely used reliability coefficient have been noted. For example, Cortina (1993) described alpha as "one of the most important and pervasive statistics in research involving test construction and use." [6, p. 98] Similarly, Peterson (1994) maintained that: "By far the most commonly used reliability coefficient is coefficient alpha." [22, p. 382]

Empirical reviews have largely supported these propositions. Hogan, Benjamin and Brezinski (2000) examined the frequency of use for various reliability coefficients for 2,078 tests obtained from research in the *Directory of Unpublished Experimental Mental Measures, Volume 7* (Goldman, Mitchel, & Egelson, 1997). [13] [10] They found that alpha was the overwhelming favorite with 66.5 percent of the total. Charter (2003) examined reliability coefficients in published research and found that in the time period before 1960, alpha represented only 29 percent of reliability coefficients whereas alpha accounted for 75 percent of published coefficients for the period beyond 1990. [2]

It is not surprising, therefore, that alpha has been widely studied. Some of the research has investigated the characteristics of alpha through meta-analysis [21] [22] [5] [25] [32]. For example, Peterson (1994) found that, with few exceptions, there were no substantive relationships between the magnitude of alpha and the research design characteristics investigated (e.g., type of sample, scale type, type of research). [22]

Other research has provided guidance on the proper *use* and *interpretation* of alpha. For example, Streiner (2003) notes that because "α is affected by the length of the scale, high values do not guarantee internal consistency or unidimensionality." [31, p. 103] Similarly, Schmitt (1995) has argued that it is

inappropriate to use α as a measure of unidimensionality. [28] Raykov (2001) demonstrated that when errors of measurement have a positive correlation, α can overestimate the true reliability of a composite. 23] Finally, Graham (2006) reminds us that α is based "on the essentially tau-equivalent measurement model, a measurement model that requires a number of assumptions to be met for the estimate to accurately reflect the data's true reliability (Raykov, 1997)". [11], p. 930-931)

Clearly, the nature and scope of coefficient α has been widely examined. However, one area that has not been examined is whether or not α is being properly reported in the literature. The notion that research should be properly reported and cited is a *sine qua non* of scholarship. As Wright and Armstrong (2008) note, the "prevalence of faulty citations impedes the growth of scientific knowledge." [33, p. 125]

Is α being properly reported? The purpose of this study is to provide preliminary evidence through a review of two premier journals in the area of travel and tourism. The study begins with the development of guidelines for reporting α. Next, journal articles are harvested for the years 2008-2010 for each of the two journals. The data is then content analyzed and coded. The study concludes with a presentation of results and a discussion of their implications for future research.

LITERATURE REVIEW

The most widely cited guideline for α is the recommended level. In this regard, Peterson (1994) summarized selected authors. [22] These sources indicate that, with the exception of Nunnally (1967) and Murphy and Davidshofer (1988), alpha levels below .70 are not acceptable. [18] [17] The Nunnally (1967) source can be discounted as more recent alpha level recommendations have been provided by Nunnally (1978) and Nunnally and Bernstein (1994). [19] [20] Hence, it appears that alpha reliability above .70 is the generally accepted standard.

It is important to note, however, that these recommendations involve the context of the research (i.e., preliminary, basic, applied). Lance, Butts, and Michels (2006) provide examples where Nunnally (1978) was cited without regard for the context. [14] They provided the following examples:

- Rothbard and Edwards (2003) reported that "all reliabilities exceeded the .70 criterion suggested by Nunnally (1978) and were considered acceptable" [26, p. 713]
- McAllister and Bigley (2002) wrote that "reliability assessments for all scales exceeded the minimum standard of .70 suggested by Nunnally (1978)" [15, p. 898]
- Spector et al. (2002) reported that "these scales maintained adequate internal consistency reliabilities as assessed with the widely accepted .70 coefficient alpha standard (Nunnally, 1978)" [30, p. 458]
- Schilling (2002) wrote that "reliabilities (Cronbach's alphas) were well above the recommended value of .70 . . . indicating that the scales had sufficient internal reliability" [29, p. 393]

What did Nunnally (1978) actually say? He elaborates on the context as follows:

> "What a satisfactory level of reliability is depends on how a measure is being used. In the early stages of research on predictor tests or hypothesized measures of a construct, one saves time and energy by working with instruments that have only modest reliability, for which purpose reliabilities of .70 or higher will suffice….For basic research, it can be argued that increasing reliabilities beyond .80 is often wasteful of time and funds….In contrast to the standards in basic research, in many applied settings a reliability of .80 is

> not merely high enough….In those applied settings where important decisions are made with respect to specific test scores, a reliability .90 is the minimum that should be tolerated, and a reliability of .80 is not merely high enough….In those applied settings where merely high enough….In those applied settings where important decisions are made with respect to specific test scores, a reliability .90 is the minimum that should be tolerated, and a reliability of .95 should be considered the desirable standard." (245-246)

Lance, Butts, and Michels (2006) argued that since most researchers would likely claim to be conducting basic (or perhaps applied) research, rather than preliminary research, the recommended reliability standard is actually .80, not .70. They also remind us that Carmines and Zeller (1979) made a similar recommendation: "As a general rule, we believe that reliabilities should not be below .80 for widely used scales" [1, p. 51]. As such, unless researchers state the preliminary nature of the research, the minimum recommended level is .80.

Clearly, there is already some evidence that alpha is sometimes inappropriately reported. To add structure to the analysis, Schmitt (1996) provides the following set of recommendations [28]:

- Alpha should be individually reported for each construct
- An authoritative source should be included.
- Alpha should not be reported as a range (e.g., all alphas were between .70 and .94)
- Alpha should not be reported as simply meeting a threshold (e.g., all alphas were above .70)

METHOD

Because the intent of this research is to provide preliminary evidence regarding the reporting for α, a convenience sample from two leading scholarly journals in travel and tourism were evaluated for three years (i.e., 2008-2010). In total, 49 articles were harvested for review. These articles represented all research where alpha was relevant. Included in the analysis of the articles were the citations and the verbatim describing the level of reliability.

RESULTS

The majority of articles (57.1 percent) reviewed in this study provided no citation. Of the remaining articles, 3 (6.1 percent) used an outdated citation for Nunnally (i.e., 1967). For two of these articles, the alphas were greater than .70 indicating that the citation did not adversely support a suboptimal alpha. In one article, a single alpha was below .50. The authors cited a relatively obscure citation noting that lower reliability coefficients (in the range of .3 to .5) were acceptable. Further review of the source article found that it did not directly support this argument.

Given that Nunnally (1978) and Nunnally and Bernstein (1994) are relatively well-known authorities on the subject, it was expected that they would be better represented in the sample. The results indicated that they comprised the majority of the citations with different editions of the Hair et. al textbook being the second most cited. [e.g., 12]

Finally, the results suggest that numerous "other" citations were sometimes used that were not authoritative. In one study, the authors merely cite that a previous study found acceptable levels of reliability for the scales.

In several other studies, the researchers used a secondary source. For example, one study cited Churchill (1979) as stating that the recommended minimum level for alpha is .70. [4] A review of Churchill, however, reveals that he was referencing Nunnally (1967). Similarly, one article referenced Ruekert and Churchill (1984), but a review of that article revealed that they were citing Nunnally (1978) as the source for the alpha level. [27]

In another study, low alpha levels (i.e., <.70) were justified based on the fact that alpha is sensitive to the length of the scale. The lows alphas were thereby justified on the grounds that they contained a small number of items.

One study cited Lance, Butts, and Michels (2006) in support of an acceptable level of .70. [14] A review of the citation, however, reveals that Lance et. al were arguing that their own interpretation of Nunnally (1978) was that the minimum should be .80 not .70.

Finally, while just over 61 percent of those reporting alpha did so for each individual construct, the remainder (approximately 40 percent) reported on a range for the alphas in the study or simply noted that all of the alphas exceeded the recommended minimum level.

CONCLUSIONS

This study examined the nature of reporting for coefficient α in the scholarly literature. While the study only covered two journals over a three-year time period, the results nonetheless suggest the potential for improper reporting on a wider scale. At a minimum, given that few studies have previously focused on this issue, the results would appear to provide additional guidance.

Perhaps most importantly, this study highlights the issue of the appropriate cut-off value for alpha. Since many researchers simply reported that the alpha levels were acceptable based on a .70 criteria, it implies that the research was preliminary. This would seem highly unlikely. Since this study did not collect and examine verbatim with regard to the preliminary nature of each studies, future research make want to examine this issue in greater detail.

It is also important to recognize that α has received a considerable amount of attention that transcends the issue of reporting. Streiner (2003), for example, notes that a high α is not always a measure of high internal consistency because alpha is strongly affected by the length of the scale. [31] He concludes that "even though a scale may consist of two or more independent constructs, α could be substantial as long as the scale contains enough items. The bottom line is that a high value of α is a prerequisite for internal consistency, but does not guarantee it; long, multidimensional scales will also have high values of α." [31, p. 102]

Striener (2003) also reminds us that that high levels of α may "reflect unnecessary duplication of content across items and point more to redundancy than to homogeneity; or, as McClelland (1980) put is, 'asking the same question many different ways.' (p. 30)" [31, p. 102] [16]

Others have argued that additional information should be provided to assess the quality of α. In particular, Duhachek and Iacobucci (2004) recommend that all α levels be reported in conjunction with standard error or confidence interval estimates, a recommendation supported by others (e.g., Lance, Butts, and Michels 2006). [8] As Duhachek, Coughlan, and Iacobucci 2005 note "One reason confidence intervals are informative is that researchers with lower alphas need not always be tentative in their interpretations,

depending on factors such as sample size. Similarly, a higher alpha estimate, but accompanied by a high variance, would limit researchers in their data interpretations." [7, p. 295]

One might also question the use of α when the research involves a structural equation model. In this situation, the more appropriate measure of internal consistency is the Fornell/Larcker Index (FLI) of composite reliability. [9] Future research might want to follow in the spirit of the current study by ascertaining whether or not researchers are utilizing the appropriate reliability coefficient.

In conclusion, the results of this study provide preliminary evidence regarding the reporting practices of alpha in scholarly journals in travel and tourism. As such, additional research is needed to more fully validate these results. Besides extending the sample to include a broader range of years, it might also be useful to examine (1) the nature of the research (preliminary, basic, applied), (2) the use of confidence intervals and standard errors for alpha, (3) the reporting of alternative reliability coefficients such as composite reliability (i.e., FLI), and (4) how alpha is reported in scholarly publications outside of travel and tourism. Doing so will help facilitate scientific progress in the discipline.

REFERENCES

[1] Carmines, E. G., & Zeller, R. A. Reliability and validity assessment. 1979, Newbury Park, CA: Sage.

[2] Charter, R.A. A Breakdown of Reliability Coefficients by Test Type and Reliability Method, and the Clinical Implications of Low Reliability. *Journal of General Psychology*, 2003, 130(3), 290-304.

[3] Cronbach, L. J. Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 1951, 16, 297-334.

[4] Churchill, G.A. A Paradigm for developing Better Measures of Marketing Constructs. *Journal of Marketing Research*, 1979, 16, 64-73.

[5] Conway J. M. & Jako, R.A. A Meta-Analysis of Interrater and Internal Consistency Reliability of Selection Interviews. *Journal of Applied Psychology*, 1995, 80(5), 565-579.

[6] Cortina, J.M. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 1993, 78(1), 98-104

[7] Duhachek, A., Coughlan, A.T. & Iacobucci, D. Results on the Standard Error of the Coefficient Alpha Index of Reliability," *Marketing Science*, 2005, 24(2), 294-301.

[8] Duhachek, A. & Iacobucci, D. Alpha's Standard Error (ASE): An Accurate and Precise Confidence Internal Estimate. *Journal of Applied Psychology*, 2005, 89(5), 792-808.

[9] Fornell C. and Larcker, D.F. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 1981, 18, 39-50.

[10] Goldman, B. A., Mitchel, D. F.,& Egelson, P. E. Directory of unpublished experimental measures (Vol. 7). Washington, DC: *American Psychological Association*.

[11] Graham, J.M. Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability: What They Are and How to Use Them. *Educational and Psychological Measurement*, 2006, 66, 930-944.

[12] Hair, J.F., Bush, R.P., & Ortinau, D.J. Marketing Research, 2008, 4th edition, McGraw-Hill Professional Publishing

[30] Hogan, T.P., Benjamin, A. & Brezinski, K.L. Reliability Methods: A Note on the Frequency of use of Various Types. *Educational and Psychological Measurement*, 2000, 66, 523-531.

[14] Lance, C.E., Butts, M.M., & Michels, L.C. The Sources of Four Commonly Reported Cutoff Criteria: What did they really say*? Organizational Research Methods*, 2006, 9, 202-220.

[15] McAllister, D.J., & Bigley, G.A. Work Context and the Definition of Self: How Organizational Care Influences Organization-Based Self-Esteem. *Academy of Management Journal*, 2002, 45, 894-904.

[16] McClelland, D. C. Motive Dispositions: The Merits of Operant and Respondent Measures. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 1; pp. 10–41), 1980, Beverly Hills, CA: Sage.

[17] Murphy, K.R. & Davidshofer, C.O. *Psychological testing: Principles and Applications*. 1988, Englewood Cliffs, NJ: prentice-hall.

[18] Nunnally, J. C. *Psychometric theory* (1st ed.), 1967, New York: McGraw-Hill.

[19] Nunnally, J. C. *Psychometric theory* (2nd ed.), 1978, New York: McGraw-Hill.

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

[21] Peter, J.P. Reliability: A Review of Psychometric Basics and Recent Marketing Practices. *Journal of Marketing Research*, 1979, 16, 6-17.

[22] Peterson, R.A. A Meta-analysis of Cronbach's Coefficient Alpha," *Journal of Consumer Research*, 1994, 21, 381-391.

[23] Raykov, T. Scale Reliability, Cronbach's Coefficient Alpha, and Violations of Essential Tau-Equivalence with Fixed Congeneric Components," *Multivariate Behavioral Research*, 1997, 32(4), 329-353.

[24] Raykov, T. Bias of Coefficient α for Fixed Congeneric Measures with Correlated Errors. *Applied Psychological Measurement*, 2001, 25, 69-76.

[25] Rodriguez M.C. & Maeda,Y. Meta-Analysis of Coefficient Alpha. *Psychological Methods*, 2006, 11(3), 306-322.

[26] Rothbard, N. P., & Edwards, J. R. Investment in Work and Family Roles: A Test of Identity and Utilitarian Motives. *Personnel Psychology*, 2003, 56, 699-729.

[27] Ruekert, R.W. & Chruchill, G.A. Reliability and Validity of Alternative Measures of Channel Member Satisfaction. *Journal of Marketing Research*, 1984, 21(2), 226-233.

[28] Schmitt, N. Uses and Abuses of Coefficient Alpha. *Psychology Assessment*, 1996, 8(4), 350-353.

[29] Schilling, M. A. Technology Success and Failure in Winner-take-all Markets: The Impact of Learning Orientation, Timing, and Network Externalities. *Academy of Management Journal*, 2002, 45, 387-398.

[30] Spector, P. E., Cooper, C. L., Sanchez, J. I., O'Driscoll, M., Sparks, K., Bernin, P. Locus of Control and Well-being at Work: How Generalizable are Western Findings? *Academy of Management Journal*, 2002, 45, 453-470.

[31] Streiner, D.L. Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, 2003, 80(1), 99-103.

[32] Vassar, M. & Bradley, G. A Reliability Generalization Study of Coefficient Alpha for the Life Orientation Test. Journal of Personality Assessment, 2010, 92(4), 362-370.

[33] Wright, M. & Armstrong, J.S. The Ombudsman: Verification of Citations: Fawlty Towers of Knowledge? *Interfaces*, 2008, 38(2), 125-139.