



The Significance of Congruence Coefficients: A Comment and Statistical Test

Arthur G. Bedeian
Louisiana State University
Achilles A. Armenakis
Auburn University
W. Alan Randolph
University of Baltimore

Contemporary researchers have mistakenly continued to assert that no procedure exists for assessing the statistical significance of congruence coefficients. Such a procedure, however, has existed for more than a decade. This study presents and illustrates the procedure and discusses the problems that can arise as information is disseminated through academic channels.

The system of communication used in management or any other discipline should ideally provide a mechanism for information building. Indeed, Ziman (1969) points out that "the invention of a mechanism for the systematic publication of fragments of scientific work may well have been the key event in the history of modern science" (p. 318).

Although Ziman's (1969) observation rings true, it should also be recognized that the dissemination of information through academic channels is equally subject to the so-called serial transmission effect (Bedeian, 1986). As both research findings (and techniques) are transmitted across time, details are often inadvertently omitted, new information overlooked, and inaccuracies in reporting become solidified. Evaluations of presentations in journals and textbooks have revealed widespread inaccuracies in the recounting of even so-called classic experiments (e.g., Berkowitz, 1971; Patzig & Wisdom, 1986; Patzig & Zimmerman, 1985-86; Pethia, 1983; Zimmerman, 1978). The impact of these errors is felt as they are transmitted from generation to generation of scholars and ultimately accepted as fact, even in the face of factual information (e.g., Cartwright, 1973). The purpose of this article is to comment upon a contemporary occurrence of the serial transmission effect involving a particular research technique—the

The helpful comments of Arthur P. Brief, Jerry I. Porras, Elizabeth A. Ravlin, and David M. Schweiger on various earlier draft manuscripts are gratefully acknowledged.

Address all correspondence to Arthur G. Bedeian, Department of Management, Louisiana State University, Baton Rouge, LA 70803-6312.

Copyright 1988 by the Southern Management Association 0149-2063/88/\$2.00.

coefficients of congruence method for evaluating the similarity of factor analytic solutions (Tucker, 1951). An equally important purpose, however, is to share the Korth and Tucker (1975) procedure for assessing the statistical significance of congruence coefficients.

The Coefficients of Congruence Method

Factor analysis has been used as a data analytic technique for scientific purposes ranging from data reduction to hypothesis testing. Often researchers find it necessary to compare two or more factor structures to determine their similarity, either because of differences in subject samples or possible differences in temporal stability. For example, in the instance of cross-sectional investigations, a researcher may find it necessary: (a) to employ factor comparison methods to cross validate factor solutions across samples (e.g., Angle & Perry, 1981; Dunham & Pierce, 1986; Fullagar, 1986; Hatfield & Huseman, 1982; Ivancevich & McMahon, 1977; Jung, Dalessio, & Johnson, 1986; Ladd, Gordon, Beauvais, & Morgan, 1982; Whitely & England, 1977); (b) to quantitatively evaluate questionnaire responses to determine whether to combine data obtained from two or more survey sites (e.g., Golembiewski & Yeager, 1978) or from groups with differing demographic characteristics (e.g., Dunham, Aldag, & Brief, 1977); or (c) to assess the stability of baseline and experimental measures (e.g., Cascio, 1976; Lorenzi, Sims, & Slocum, 1981). Similarly, in the case of longitudinal investigations, it may be necessary to compare factor solutions to test hypotheses regarding the emergence of different types of change across time (e.g., Armenakis & Smith, 1978; Golembiewski & Billingsley, 1980; Golembiewski, Billingsley, & Yeager, 1976; Randolph, 1982; Schmitt, 1982).

Although limitations associated with the sensitivity of congruence coefficients have been noted, their use is generally accepted as a "practical guide" in the absence of information necessary for more refined techniques (Gorsuch, 1974, p. 254). A coefficient of congruence matrix summarizes interrelationships among factors in one factor structure with those in another factor structure. Coefficients are computed from the following formula:

$$\phi_{pq} = \frac{\sum_{j=1}^n (1^{ajp} \cdot 2^{ajq})}{\sqrt{\left(\sum_{j=1}^n 1^{a^2jp}\right) \left(\sum_{j=1}^n 2^{a^2jq}\right)}} \quad [1]$$

where: ϕ_{pq} = coefficients of congruence
 p = number of factors in sample 1
 q = number of factors in sample 2
 1^{ajp} = factor loading for the j th item of factor p in sample 1
 2^{ajq} = factor loading for the j th item of factor q in sample 2
 n = number of items factored in each sample

With a range from -1.00 to $+1.00$, coefficients of congruence resemble

product-moment correlation coefficients. However, because the factor loadings in Formula 1 are divided by the number of items factored rather than number of subjects, they are *not* measures of dispersion. Consequently, a test of significance similar to that used in evaluating correlation coefficients cannot be employed.

Writing in each of the three editions (1960, 1967, 1976) of his influential reference book *Modern Factor Analysis*, Harman (1976) observes that though developing a test of significance for congruence coefficients sounds like a classical problem in the theory of statistical sampling "little progress has been made toward its resolution" (p. 346). Although one cannot be certain in all instances, this statement appears to have been transmitted from generation to generation of researchers, and essentially accepted as fact (e.g., Armenakis & Smith, 1978; Griffin, Moorhead, Johnson, & Chonko, 1980; Mulaik, 1972). This acceptance is easily understandable given the influential nature of the various editions of Harman's book. For years, researchers have relied on Harman's statement without being aware of new information. Thus, we find that "Harman (1976) has indicated that congruence coefficients in the mid-80s and above should be considered similar" (Fried & Ferris, 1986, p. 421) or that it can be "inferred from Harman's examples and discussion ... that coefficients above .80 are entirely acceptable and coefficients below .50 are unacceptable" (Griffin, 1981, p. 107). Other researchers have simply stated that the congruence coefficients they report are "high" (Jung, Dalessio, & Johnson, 1986) or that comparative factor-loading patterns are "virtually identical" (Angle & Perry, 1981, p. 4).

Significance Testing of Congruence Coefficients

What appears to have gone unnoticed is that Korth and Tucker (1975) have developed a procedure for assessing the significance of congruence coefficients. The fact that the procedure was reported in a scholarly outlet (i.e., *Psychometrika*) not widely read by management researchers, may account for it being generally overlooked. Informed use of the Korth and Tucker (1975) procedure will permit management researchers to present their results with greater statistical confidence. The Korth and Tucker (1975) procedure is presented in the form of a Fisher Z transformation involving the arc hyperbolic tangent of a congruence coefficient. The calculations necessary for this procedure can be mathematically reversed from a Fisher Z to determine the minimum value necessary for a congruence coefficient to reach statistical significance. Generated by the authors following conversations and correspondence with Korth (personal communication, 1979) and Tucker (personal communication, November 22, 1986), this reversal in direction requires five steps:

Step 1. Following rotation to maximal congruence, evaluate the similarity of the focal factor structures using the coefficients of congruence formula presented earlier (Korth & Tucker, 1975). An algorithm for this purpose is available (Armenakis, Feild, & Wilmoth, 1977).

Step 2. Determine the appropriate number of degrees of freedom and obtain the critical *t*-value necessary for an initial congruence coefficient (ϕ_{11}) to be statistically significant.

Degrees of freedom are equal to the difference between the number of items

factored in an analysis and the number of factors in a factor solution for Time-1 (L.R. Tucker, personal communication, November 22, 1986). That is, the degrees of freedom is the number of free dimensions in a space with dimensionality equal to the number of items factored. After the degrees of freedom are computed, the critical value of t can be located in a standard t -table.

Step 3. Compute the minimum value (MINVAL) an initial coefficient of congruence (ϕ_{11}) must attain to be significant ($p < .05$). MINVAL is defined by the formula:

$$\text{MINVAL} = \text{arc hyperbolic tangent of } x = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad [2]$$

where $x = [(\text{critical value of } t) \cdot (\sigma_z)] + \bar{X}_z$

$e = 2.7183$ (base of the natural system of logarithms).

Values of σ_z (the standard deviation of the congruence coefficients converted to Fisher Z 's) and \bar{X}_z (the mean of the congruence coefficients converted to Fisher Z 's) are given in Korth and Tucker (1975, Table 4). Interpolation may be necessary depending on the number of items and factors in a particular data set.

Step 4. Determine if the initial congruence coefficient (ϕ_{11}) computed in Step 1 equals or exceeds the required MINVAL computed in Step 3. If so, the content or "meaning" of the associated factors being compared is statistically similar.

Step 5. Repeat Steps 1 through 4 for any remaining factors.

As Korth and Tucker (1975) acknowledge, because the minimum value (MINVAL) a congruence coefficient must attain to be significant is empirically determined, cautious interpretation is necessary. The required MINVAL calculated in Step 3 will vary with both the number of items and factors in a solution, as well as the order of factor extraction. In some cases, a coefficient may need only exceed .25 to indicate factor structure similarity, whereas in other cases a coefficient less than .90 may indicate lack of similarity. This underscores the risk of following "rules of thumb" that have been passed on by succeeding generations of researchers.

An Illustration

An illustration of the Korth and Tucker (1975) technique may be helpful in appreciating the usefulness of the proposed significance test. The illustration involves an organization development (OD) intervention in a university student affairs division. A 26-item, paper-and-pencil instrument, titled the Organizational Questionnaire (OQ) was developed as a diagnostic tool for the intervention. It was initially completed by 196 staff members (an 85% response rate) and administered again 18 months later, following implementation of a rather extensive OD intervention. Because the intervention was designed to have a pronounced effect on inter-departmental communications, interdepartmental coordination, staff participation opportunities, physical barriers to communication, and clarity of job expectations, gamma change was expected, being evidenced by a transformation in the OQ factor structure (Golembiewski et al., 1976). A more detailed account of the methodology employed is described in Randolph and Elloy (1983).

Factor analysis of responses to the 26 OQ items at Time-1 and Time-2 pro-

duced seven factors, as determined by both the classic eigenvalue cutoff of 1.0 (Kaiser, 1958) and a scree test (Cattell, 1966). Content analysis of items at Time-1 suggested the following labels: Factor One—*Intradepartmental Communications*; Factor Two—*Interdepartmental Coordination*; Factor Three—*Interdepartmental Communications*; Factor Four—*Staff Participation in Decisions*; Factor Five—*Perceived Physical Barriers to Interdepartmental Communication*; Factor Six—*Clarity of Job Expectations*; and Factor Seven—*Staff Development Opportunities*.

As noted earlier, the rather extensive intervention was directed at altering interdepartmental communications, interdepartmental coordination, and staff participation opportunities. And in the opinions of the intervention task force and upper management, these three areas were profoundly affected by the planned intervention. Hence, a change in the structure of the factors representing these areas was expected. On the other hand, little if any effort was expended toward changing intra-departmental communications and staff development opportunities. Thus, little change in the structure of the factors tapping these areas was anticipated.

Results in Table 1 provide a statistical means to test these expectations. The boxed congruence coefficients indicate the Time-1 - Time-2 matching deemed most appropriate (largest coefficient in each row, without twice matching a Time-1 factor). An inspection of the factor loadings confirmed that this matching made sense conceptually. The congruence coefficients for the seven factors were considered in descending order and compared to the MINVALs shown in Table 1 to determine if a change in factor structure had occurred.

The initial coefficient ($\phi_{11} = .92$) relates to *Intradepartmental Communications*. With a MINVAL of .87, as predicted, it suggests no change in factor structure. In contrast, the second ($\phi_{22} = .53$) and third ($\phi_{36} = .49$) congruence coefficients relating to *Interdepartmental Coordination* and *Interdepartmental Communications*, respectively, both suggest factor structure change, because they are both less than their required MINVALs (.74 and .71). The same is true for *Clarity of Job Expectations* ($\phi_{63} = .39$, MINVAL = .42), and *Perceived Physical Barriers to Interdepartmental Communication* ($\phi_{57} = .35$, MINVAL = .36). Both coefficients suggest changes in factor structure, as might be expected from the intervention.

For the remaining factors, the intervention did not operate as intended. For *Staff Development Opportunities* no factor structure change was predicted, but appears to have occurred ($\phi_{74} = .13$, MINVAL = .14). And, for *Staff Participation in Decisions*, factor structure change was predicted, but none is indicated ($\phi_{45} = .26$, MINVAL = .22).

Using a common rule of thumb such as a .80 cutoff would have been misleading in interpreting this final factor. Although a .26 coefficient appears quite low, it exceeds the required .22 MINVAL, indicating that the content or meaning of the associated factors at Time = 1 and Time = 2 are statistically similar. Moreover, in the case of the present illustration, several additional interpretive errors would have been possible because only the initial MINVAL is above .80. Indeed, some of the MINVALs are quite low, yet still reach acceptable levels of statistical

Table 1
Coefficients of Congruence Method Results

Time-1 Factors	Time-2 Factors							MINVAL ^a
	1	2	3	4	5	6	7	
1. Intradepartmental Communications	.92 ^b	.04	-.13	.05	.21	.10	.13	.87
2. Interdepartmental Coordination	.00	.53	-.19	.00	-.47	.06	-.10	.74
3. Interdepartmental Communications	-.11	.37	.03	.17	.11	.49	.21	.71
4. Staff Participation in Decisions	-.17	.14	-.03	.03	.26 ^b	.02	.29	.22
5. Physical Barriers to Interdepartmental Communications	.02	.00	.19	.11	-.24	.25	.35	.36
6. Clarity of Job Expectations	.07	.11	.39	.26	.03	.02	.16	.42
7. Staff Development Opportunities	-.05	.26	-.08	.13	.09	.10	.09	.14

^aComputed using the procedure described in Korth and Tucker (1975) and in the present manuscript.

^bIndicates factor structure similarity ($p \leq .05$) for Time-1 versus Time-2.

significance. On the other hand, had the first coefficient of congruence (ϕ_{11}) been .85, the .80 rule of thumb would have suggested similarity of factor structures when such was not the case, given a .87 MINVAL.

Conclusion

An immediate implication of this note is that though contemporary researchers have understandably thought otherwise, a technique for assessing the statistical significance of congruence coefficients has existed for more than 10 years. It is all too easy for researchers to be unaware of the many techniques existing in the vast psychometric literature. Given that researchers must often rely on reference books or the methods sections of journal articles, it is easy to understand how serial transmission errors involving measurement techniques occur. Ideally, this note will aid future management researchers by making known a valuable technique, as well as by making them more careful in avoiding serial transmission errors.

References

- Angle, H.L., & Perry, J.L. (1981). An empirical assessment of organizational commitment and organizational effectiveness. *Administrative Science Quarterly*, 26, 1-14.
- Armenakis, A.A., Feild, H.S., & Wilmoth, J.N. (1977). An algorithm for assessing factor structure congruence. *Educational and Psychological Measurement*, 37, 213-214.
- Armenakis, A.A., & Smith, L. (1978). A practical alternative to comparison group designs in OD evaluations: The abbreviated time series design. *Academy of Management Journal*, 21, 499-507.

- Bedeian, A.G. (1986). The serial transmission effect: Implications for Academe. *Bulletin of the Association for Business Communication*, 44 (1), 34-36.
- Berkowitz, L. (1971). Reporting an experiment: A case study in leveling, sharpening, and assimilation. *Journal of Experimental Social Psychology*, 7, 237-243.
- Cartwright, D. (1973). Determinants of scientific progress: The case of research on the risky shift. *American Psychologist*, 28, 222-231.
- Cascio, W.F. (1976). An empirical test of factor structure stability in attitude measurement. *Educational and Psychological Measurement*, 36, 847-854.
- Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Dunham, R.B., Aldag, R.J., & Brief, A.P. (1977). Dimensionality of task design as measured by the Job Diagnostic Survey. *Academy of Management Journal*, 20, 209-223.
- Dunham, R.B., & Pierce, J.L. (1986). Attitudes toward work schedules: Construct definition, instrument development, and validations. *Academy of Management Journal*, 29, 170-182.
- Fried, Y., & Ferris, G.R. (1986). The dimensionality of job characteristics: Some neglected issues. *Journal of Applied Psychology*, 71, 419-426.
- Fullagar, C. (1986). A factor analytic study on the validity of a union commitment scale. *Journal of Applied Psychology*, 71, 129-136.
- Golembiewski, R.T., & Billingsley, K. (1980). Measuring change in O.D. panel designs: A response to critics. *Academy of Management Review*, 5, 97-104.
- Golembiewski, R.T., Billingsley, K., & Yeager, S. (1976). Measuring change and persistence in human affairs: Types of change generated by OD designs. *Journal of Applied Behavioral Science*, 12, 133-137.
- Golembiewski, R.T., & Yeager, S. (1978). Testing the applicability of the JDI to various demographic groupings. *Academy of Management Journal*, 21, 514-519.
- Gorsuch, R.L. (1974). *Factor analysis*. Philadelphia: Saunders.
- Griffin, R.W. (1981). A longitudinal investigation of task characteristics relationships. *Academy of Management Journal*, 24, 99-113.
- Griffin, R.W., Moorhead, G., Johnson, B.H., & Chonko, L.B. (1980). The empirical dimensionality of the Job Characteristics Inventory. *Academy of Management Journal*, 23, 772-777.
- Harman, H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.
- Hatfield, J.D., & Huseman, R.C. (1982). Perceptual congruence about communication as related to satisfaction: Moderating effects of individual characteristics. *Academy of Management Journal*, 25, 349-358.
- Ivancevich, J.M., & McMahon, J.T. (1977). Black-white differences in a goal-setting program. *Organizational Behavior and Human Performance*, 20, 287-300.
- Jung, K.G., Dalessio, A., & Johnson, S.M. (1986). Stability of the factor structure of the Job Descriptive Index. *Academy of Management Journal*, 29, 609-616.
- Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- Korth, B., & Tucker, L.R. (1975). The distribution of chance congruence coefficients from simulated data. *Psychometrika*, 40, 361-372.
- Ladd, R.T., Gordon, M.E., Beauvais, L.L., & Morgan, R.L. (1982). Union commitment: Replication and extension. *Journal of Applied Psychology*, 67, 640-644.
- Lorenzi, P., Sims, H.P., Jr., & Slocum, J.W., Jr. (1981). Perceived environmental uncertainty: An individual or environmental attribute? *Journal of Management*, 7, 27-41.
- Mulaik, S.A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.
- Patzig, W.D., & Wisdom, B.L. (1986). Some words of caution about having students read the classics. In D.F. Ray (Ed.), *Proceedings of the Southern Management Association* (pp. 117-119). Mississippi State, MS: SMA.
- Patzig, W.D., & Zimmerman, D.K. (1985-86). Accuracy in management texts: Examples in reporting the works of Maslow, Taylor, and McGregor. *Organizational Behavior Teaching Review*, 10 (2), 39-47.
- Pethia, P.F. (1983). *Pseudo-history in management textbook: Illustrations and comments*. Paper presented at the meeting of the Academy of Management, Dallas.

- Randolph, W.A. (1982). Planned organizational change and its measurement. *Personnel Psychology*, 35, 117-139.
- Randolph, W.A., & Elloy, D.F. (1983). *Measuring organizational change: A comparison of two approaches*. Unpublished manuscript, University of South Carolina, Department of Management, Columbia.
- Schmitt, N. (1982). The use of analysis of covariance structures to assess beta and gamma change. *Multivariate Behavioral Research*, 17, 343-358.
- Tucker, L. (1951). *A method for synthesis of factor analysis studies* (Personnel Research Section Report, No. 984). Washington, DC: Department of the Army.
- Whitely, W., & England, G.W. (1977). Managerial values as a reflection of culture and the process of industrialization. *Academy of Management Journal*, 20, 439-453.
- Ziman, J.M. (1969, October 25). Information, communication, knowledge. *Nature*, 318-324.
- Zimmerman, D.K. (1978). Participative management: A reexamination of the classics. *Academy of Management Review*, 3, 896-901.
-