Educational and Psychological Measurement

http://epm.sagepub.com/

A Monte Carlo Study of the Sampling Distribution of the Congruence Coefficient

Wendy J. Broadbooks and Patricia B. Elmore

Educational and Psychological Measurement 1987 47: 1

DOI: 10.1177/0013164487471001

EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT
1987, 47

A MONTE CARLO STUDY OF THE SAMPLING DISTRIBUTION OF THE CONGRUENCE COEFFICIENT

WENDY J. BROADBOOKS
Arthur Andersen & Co., St. Charles, IL

PATRICIA B. ELMORE
Southern Illinois University at Carbondale

The effects of sample size, number of variables, and population value of the congruence coefficient on the sampling distribution of the congruence coefficient were examined. Sample data were generated on the basis of the common factor model, and principal axes factor analyses were performed. The results indicated that when the population congruence coefficient is zero, the sampling distribution of the congruence coefficient is relatively stable and is similar to the sampling distribution of the correlation coefficient. The congruence coefficient was found to be positively biased for values of psi from .10 to .40 and negatively biased for the following values of psi: .50, .60, .90, and .99. The amount of bias (a) depends on the sample size and the number of variables used and (b) tends to decrease as sample size increases.

REPLICATIONS of studies utilizing factor analysis are just as necessary as replications of studies using any other statistical technique. In order to make comparisons of sets of factors and their interpretability across different studies, some objective basis is needed for the comparison in order to ascertain the degree of relationship between corresponding factors in the studies. Several measures of factor similarity are available and commonly used. The present study focused on one such similarity coefficient: the congruence coefficient. The congruence coefficient (referred to as the unadjusted correlation in some earlier studies) was developed as a measure of the similarity of the factor patterns for different samples of subjects (Burt, 1948; Tucker, 1951; Wrigley and Neuhaus, 1955). The congruence coefficient involves a comparison of two sets of factor loadings in terms of both the pattern and magnitude of the loadings. It has been used extensively as a descriptive statistic in research comparing factors across studies. In most of these studies the same set of variables but different samples of subjects were used. The formula used to calculate the congruence coefficient is as follows:

$$pq = \frac{\sum_{j=1}^{n} (a_{jp})\,(b_{jq})}{\sqrt{\left[\sum_{j=1}^{n} a_{jp}^{2}\right]\left[\sum_{j=1}^{n} b_{jq}^{2}\right]}}$$

where n is the number of variables common to the two studies, p is the factor found in study one, q is the factor found in study two, and a_{jp} and b_{jq} are factor loadings. The value of the congruence coefficient can range from $+1$ (perfect agreement) to -1 (perfect inverse agreement). A zero value for the congruence coefficient reflects a lack of agreement between the two vectors of factor loadings. Some studies have examined the distribution of the congruence coefficients under various conditions (Korth, 1978; Korth and Tucker, 1975, 1979; Nesselroade and Baltes, 1970; Nesselroade, Baltes, and Labouvie, 1971); however, none of these studies has examined the sampling distribution for varying values of the population congruence coefficient.

The purpose of the present research was to develop and to investigate an empirical sampling distribution of the congruence coefficient. The effects of sample size, the number of variables, and the magnitude of the population congruence coefficient were considered in the analyses.

Nesselroade and Baltes (1970) carried out a simulation study of factor matching based on random data. They examined the distribution of the congruence coefficient and the effects of various parameters on the magnitude of the congruence coefficient. The experimenters' conclusions included a statement that the congruence coefficients in factor comparison studies should be evaluated with reference to the number of variables used and the number of factors extracted. In a reanalysis of the data obtained in the Nesselroade and Baltes (1970) study, Nesselroade, Baltes, and Labouvie (1971) used Meredith's oblique rotation technique. Again, congruence coefficients were computed for each pair of factors and $3 \times 3 \times 2$

analyses of variance were used to analyze the data: (a) sample size (50, 100, 200); (b) number of variables (15, 30, 45); and (c) number of factors extracted (5, 10). As in the earlier study, significant main effects were found for the number of variables and the number of factors. Nesselroade, Baltes, and Labouvie (1971) suggested that their results (where an oblique rotation was used) were in substantial agreement with the results of the earlier Nesselroade and Baltes study (where an orthogonal rotation was used).

Korth and Tucker (1975, 1979) reported the results of a Monte Carlo study examining the distribution of chance congruence coefficients from simulated data. The results provided normative data about the distribution of the congruence coefficient. The tabled values could be used to evaluate the significance of congruence coefficients.

Tables of the distribution of chance congruence coefficients were constructed by Korth (1978) using a Monte Carlo method. The tables were developed to test the hypothesis that two independent sets of random vectors have been matched. The number of variables, for which values of chance congruence coefficients were presented, ranged from 10 to 70. In addition to the number of variables, the table contained the: (a) number of vectors; (b) number of replications; and (c) mean, standard deviation, and critical value of the congruence coefficient (at the .05 significance level). Korth concluded that rejecting the null hypothesis that two independent sets of random vectors have been matched still does not establish that two factors are identical.

The research performed in the present study is an extension of previous research. In this study a systematic examination of the distribution of the congruence coefficient was conducted for varying values of the population congruence coefficient. This study provides information not only for the situation where one can assume that there is no relationship between two factors (which would theoretically result from a population congruence coefficient of zero), but also for situations where one can assume that there is some relationship between two factors (which would result from a population congruence coefficient different from zero).

Method

For the purposes of the present study, simulated data were generated. Population congruence coefficients were specified by determining pairs of population factor loading vectors that resulted in the following values for the congruence coefficient: .00, .10, .20,

.30, .40, .50, .60, .70, .80, .90, and .99. Sample data were generated from these population factor loading vectors on the basis of the common factor model. In matrix notation, a raw data matrix for each sample was generated by using the following equation (Gorsuch, 1974, p. 48):

$$X_{nv} = F_{nf}P_{fv}' + U_{nv}D_{vv}' \tag{2}$$

where X_{nv} = standard score data matrix for n individuals on v variables; F_{nf} = common factor scores of n individuals on f factors; P_{fv}' = transpose of factor loading matrix containing weights to reproduce the v variables from the f factors; U_{nv} = unique factor scores for n individuals on v variables; and D_{vv}' = transpose of diagonal matrix of weights for the v unique factors.

Sample sizes examined were 50, 100, and 200. Number of variables used were 10, 30, and 50. The data were generated by using FORTRAN computer programs. Principal axes factor analyses were performed by using squared multiple correlations as the communality estimates, and one factor was extracted in each analysis. A sample congruence coefficient (r_c) and a sample Fisher z-transformed congruence coefficient (Z_{r_c}) were computed for each pair of sample factor loading vectors. Sampling distributions of the congruence coefficients and of Fisher z-transformed congruence coefficients were formed on the basis on 200 replications for each combination of psi (population congruence coefficient), sample size, and number of variables. The congruence coefficients were transformed by using Fisher's logarithmic transformation for the Pearson product-moment correlation coefficient (Fisher, 1921) based on the computational similarity of the congruence coefficient. For analyses where psi were not equal to zero, the obtained sample congruence coefficient was Fisher z-transformed by using the following formula (Glass and Stanley, 1970, p. 265):

$$Z_{r_c} = \log_e \sqrt{(1 + r_c) / (1 - r_c)} \tag{3}$$

where Z_{r_c} is the Fisher z-transformed congruence coefficient, $\log_e$ is the natural logarithm, and r_c is the sample congruence coefficient.

Characteristics of the empirical sampling distribution were examined (e.g., mean, variance, selected percentiles) along with the results of goodness of fit to normality tests. Goodness-of-fit tests to the normal distribution were performed on the sampling distributions of psi for the case where psi = 0 and on the sampling distributions of the Fisher z-transformed psi for psi $\neq$ 0. The Statistical Analysis System (SAS) was used to perform the goodness of fit of each sampling distribution to the normal distribution (Barr, Goodnight, Sall, Blair, and Chilko, 1979). The NORMAL option in the SAS UNIVARIATE procedure was used. The SAS procedure is a modified version of the Kolmogorov-Smirnov *D* statistic.

Chi-square tests and *t* tests were performed in order to determine whether the sampling distribution of the congruence coefficient and of the Fisher *z*-transformed congruence coefficient behaved similarly to the sampling distribution of the correlation coefficient and of the Fisher *z*-transformed correlation coefficient, respectively. Chi square tests were done to test for equal variances of the desired distributions, and *t* tests were done to test for equal means.

The chi-square test used to test whether the variance of the sampling distribution of Z_{r_c} was equal to the variance of the sampling distribution of Z_r was:

$$\chi^2 = \frac{(n_r - 1){s_{z_{r_c}}}^2}{1/(n - 3)}, \; df = n_r - 1 \quad (4)$$

where n is the sample size, n_r is the number of replications, and ${s_{z_{r_c}}}^2$ is the variance of the sampling distribution of the Fisher *z*-transformed congruence coefficient. The number of degrees of freedom for each test is 199. This formula can be used to test whether the variance of a given distribution is equal to a specific value (Blommers and Lindquist, 1960). This formula will test the null hypothesis that $\sigma^2_{z_{r_c}} = 1/(n - 3)$ where $1/(n - 3)$ is the variance of the sampling distribution of the Fisher *z*-transformed Pearson product-moment correlation coefficient. The chi square test used to ascertain whether the variance of the sampling distribution of r_c was equal to the variance of the sampling distribution of r was the same as the previously cited chi-square test except that the denominator used was $1/(n - 1)$. This denominator was used because the expected variance of r, when rho = 0, is $1/(n - 1)$ where n is the sample size. All statistical tests were performed by using a .01 level of significance.

The *t* test used to evaluate whether the mean of the sampling distribution of Z_{r_c} was equal to the expected mean of the sampling distribution of Z_r was

$$t = \frac{\bar{Z}_{r_c} - Z_{\psi}}{s_{z_{r_c}}/\sqrt{n_r}}, \; df = n_r - 1 \quad (5)$$

where $\bar{Z}_{r_c}$ is the mean Fisher *z*-transformed congruence coefficient for a given sampling distribution, $Z\psi$ is the Fisher *z*-transformed population congruence coefficient, $s_{z_{r_c}}$ is the standard deviation of the sampling distribution of the Fisher z-transformed congruence coefficient, and n_r is the number of replications. The t test used to ascertain whether the mean of the sampling distribution of r_c was equal to the expected mean of the sampling distribution of r was the same as the just stated t test except that the numerator used was $\bar{r}_c - \psi$, where $\bar{r}_c$ is the mean congruence coefficient for a given sampling distribution and ψ is the population congruence coefficient.

The influence of the number of subjects was taken into account in the present study. Most factor analysis books present the issue of sample size in factor analytic research in terms of the appropriate ratio of the number of subjects to the number of variables. Rules of thumb in some texts (e.g., Nunnally, 1978) suggest that a ratio between five-to-one and ten-to-one is desirable. The ratios of the number of subjects to the number of variables used in the present study ranged from 1.67-to-one to twenty-to-one. This range encompasses the ratio of the number of subjects to the number of variables used in most factor analytic studies.

Results and Discussion

The data generated for this study resulted in a total of 88 analyses. Each analysis performed was based on pairs of samples generated. Each sample set of data generated was based on a vector of population factor loadings. These two vectors were specified such that they would result in a particular value for a population congruence coefficient (psi). In order to examine the sampling distribution of the congruence coefficient, each analysis consisted of 200 replications with sample size, number of variables, and psi remaining constant. In order to determine the number of replications needed to obtain relatively stable results, preliminary analyses were conducted. The number of replications performed was varied from

TABLE 1
Characteristics of the Sampling Distributions for Selected Values of psi

		psi = .00		psi = .10		psi = .20		psi = .30		psi = .40	
NV[a]	NS[b]	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
10	50	.013	.020	.135	.018	.256	.015	.390	.011	.568	.007
10	100	.005	.012	.127	.012	.251	.010	.388	.007	.571	.004
10	200	−.010	.005	.116	.005	.244	.004	.386	.003	.574	.002
30	50	−.001	.006	.116	.005	.236	.005	.377	.004	.567	.002
30	100	.002	.003	.121	.003	.242	.003	.384	.002	.577	.001
30	200	.000	.002	.121	.002	.244	.001	.388	.001	.585	.001
50	100	.000	.002	.121	.002	.243	.001	.385	.001	.577	.001
50	200	.001	.001	.121	.001	.244	.001	.388	.001	.585	.001

50 to 500 (values examined were 50, 75, 150, 200, 250, 300, 350, 400, 450, and 500). Three sample values for psi were chosen for these preliminary analyses (.00, .50, and .90). It was found that when 200 replications were performed the standard error of the mean was less than .01. Increasing the number of replications beyond 200 did not result in any substantial reduction in the standard error; therefore, the number of replications for each analysis was set at 200. As the standard error of the mean was less than .01, the mean of the sample congruence coefficients would be accurate to at least the nearest one-hundredth.

The characteristics of the sampling distribution of the congruence coefficient and of the Fisher z-transformed congruence coefficient for various combinations of the values of the population congruence coefficient, sample size, and number of variables are presented in Tables 1 and 2, respectively.

The following results were found concerning the characteristics of the sampling distributions of the congruence coefficient and of the Fisher z-transformed congruence coefficient: (a) the congruence coefficient is positively biased (i.e., the mean for each distribution is greater than the population value) for values of psi from .10 to .40; (b) the congruence coefficient is negatively biased (i.e., the mean for each distribution is less than the population value) for the following values of psi: .50, .60, .90, and .99; (c) the skewness and kurtosis values of most of the sampling distributions seemed to be near those expected for a normal distribution; (d) almost all of the sampling distributions examined did approximate a normal distribution (of the 88 goodness-of-fit to normality tests only six tests led to a rejection of the null hypothesis); (e) regardless of the value of psi, as sample size became larger the variance of each distribution either stayed the

TABLE 1 (Continued)

psi = .50		psi = .60		psi = .70		psi = .80		psi = .90		psi = .99	
Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
.474	.011	.567	.008	.671	.006	.789	.004	.860	.002	.950	.001
.481	.007	.577	.005	.685	.004	.806	.002	.877	.001	.970	.000
.477	.003	.577	.002	.688	.001	.814	.001	.883	.001	.980	.000
.466	.003	.566	.002	.675	.002	.790	.001	.863	.001	.952	.000
.482	.001	.584	.001	.695	.001	.812	.001	.881	.000	.972	.000
.486	.001	.590	.001	.703	.001	.822	.000	.890	.000	.982	.000
.481	.001	.587	.001	.693	.001	.810	.000	.882	.000	.973	.000
.487	.001	.591	.000	.704	.000	.822	.000	.890	.000	.982	.000

Note. For each distribution 200 replications were performed.
[a] NV = Number of variables.
[b] NS = Number of subjects.

TABLE 2
Characteristics of the Sampling Distributions for Selected Values of Z_{psi}

		Zpsi = .100 (psi = .10)		Zpsi = .203 (psi = .20)		Zpsi = .310 (psi = .30)		Zpsi = .424 (psi = .40)		Zpsi = .549 (psi = .50)	
NV[a]	NS[b]	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
10	50	.138	.019	.266	.018	.418	.016	.653	.015	.524	.018
10	100	.130	.012	.259	.011	.414	.010	.654	.010	.530	.012
10	200	.117	.005	.250	.005	.409	.004	.656	.004	.521	.005
30	50	.117	.006	.242	.005	.398	.005	.647	.005	.507	.005
30	100	.123	.003	.248	.003	.406	.003	.659	.003	.527	.003
30	200	.121	.002	.249	.002	.410	.002	.671	.002	.532	.002
50	100	.122	.002	.249	.002	.406	.001	.659	.001	.525	.002
50	200	.121	.001	.249	.001	.410	.001	.671	.001	.532	.001

Zpsi = .693 (psi = .60)		Zpsi = .867 (psi = .70)		Zpsi = 1.099 (psi = .80)		Zpsi = 1.472 (psi = .90)		Zpsi = 2.647 (psi = .99)	
Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
.653	.018	.825	.020	1.090	.027	1.325	.040	1.900	.088
.666	.012	.847	.014	1.131	.018	1.390	.030	2.171	.082
.660	.005	.847	.005	1.144	.007	1.401	.014	2.348	.068
.645	.005	.824	.006	1.078	.010	1.318	.013	1.877	.029
.671	.003	.860	.003	1.138	.005	1.389	.009	2.156	.029
.679	.002	.875	.002	1.167	.004	1.426	.005	2.363	.023
.667	.002	.855	.002	1.129	.004	1.387	.005	2.152	.015
.680	.001	.876	.001	1.165	.002	1.427	.003	2.360	.014

Note. For each distribution 200 replications were performed.
[a] NV = Number of variables.
[b] NS = Number of subjects.

same or decreased; (f) regardless of the value of psi, as the number of variables increased, the variance of each distribution either stayed the same or decreased; (g) the sampling distributions of the congruence coefficient behaved similarly to those of the correlation coefficient only when psi was equal to zero (on the basis of the results of the *t* tests, goodness-of-fit to normality tests, and chi-square tests performed in the present study).

Table 3 presents a summary of the three statistical tests that were performed on each of the sampling distributions. If the congruence coefficient behaved similarly to the correlation coefficient then there would be no "*X*'s" on the table and all the statistical tests would have led to a retention of the null hypothesis.

The findings of this study have implications for researchers who wish to use the congruence coefficient in order to determine whether factors have been replicated. When psi was equal to zero, the

TABLE 3
Summary of Statistical Tests Performed on Each Distribution

		.00/.00			.10/.100			.20/.203			.30/.310			psi/^{Z}psi .40/.424			.50/.549			.60/.693			.70/.867			.80/1.099			.90/1.472			.99/2.647		
NV[a]	NS[b]	D[c]	t[d]	χ^{2}_{v}[e]	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v	D	t	χ^2_v
10	50					X			X			X			X	X		X			X			X				X	X	X	X	X	X	X
10	100					X			X			X			X						X				X		X	X	X	X	X		X	X
10	200					X			X			X			X			X			X			X			X	X	X	X	X			
30	50			X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X
30	100			X		X	X		X	X		X	X		X	X		X	X		X	X			X		X	X		X		X	X	X
30	200			X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X			X			X	X
50	100			X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X
50	200			X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X		X	X	X	X	X

Note. For each distribution 200 replications were performed. An "X" in this table means that the null hypothesis for the test was rejected, $p < .01$.
[a] NV = Number of variables.
[b] NS = Number of subjects.
[c] D = Goodness of fit to normality test.
[d] t = Test for equal means.
[e] χ^2_v = Chi square test of equal variances.

sampling distribution of the congruence coefficient behaved similarly to the sampling distribution of the correlation coefficient. When 10 variables were used, the obtained variances were equal to the expected variances. However, when the number of variables used was larger than 10, the chi-square tests for equal variances were rejected. Therefore, a researcher could use the hypothesis-testing procedure available to test the null hypothesis that rho is equal to zero to test whether psi is equal to zero. As the obtained variance for the sampling distributions (when psi is equal to zero) were equal to or less than the expected variances, this statistical test would be a conservative one of the null hypothesis. The magnitude of the sample congruence coefficient is an indication of the strength of the relationship between the two factors.

Whenever psi is some value other than zero and whenever a sample congruence coefficient is used as an estimate of a population congruence coefficient, researchers should be aware of the amount and direction of bias of the sample congruence coefficient. For example, when psi is greater than .50, the congruence coefficient can be expected to be negatively biased. Thus, a researcher could conclude that an obtained sample congruence coefficient of greater than .50 will usually be an underestimate of the actual population value.

Further investigation of the sampling distribution under more combinations of sample size and number of variables would provide information as to whether the results reported in this paper are representative and generalizable. The results found in this study have indicated that the stability of the congruence coefficient increases as number of variables and subjects increase. The results of this study also have revealed that when the population congruence coefficient is zero, the sampling distribution of the congruence coefficient is relatively stable and is similar to the sampling distribution of the correlation coefficient (especially when 10 variables are used).

REFERENCES

Barr, A. J., Goodnight, J. H., Sall, J. P., Blair, W. H., and Chilko, D. M. (1979). *SAS user's guide*. Raleigh, NC: SAS Institute,

Blommers, P. and Lindquist, E. F. (1960). *Elementary statistical methods in psychology and education*. Boston: Houghton Mifflin.

Burt, C. (1948). Factor analysis and canonical correlations. *British Journal of Psychology, Statistical Section*, 1, 95–106.

Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron*, *1* (Part 4), 3–32.

Glass, G. V. and Stanley, J. C. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice Hall.

Gorsuch, R. L. (1974). *Factor analysis*. Philadelphia, PA: W. B. Saunders Company.

Korth, B. A. (1978). A significance test for congruence coefficients for Cattell's factors matched by scanning. *Multivariate Behavioral Research*, 13, 419–430.

Korth, B. A. and Tucker, L. R. (1975). The distribution of chance congruence coefficients from simulated data. *Psychometrika*, 40, 361–372.

Korth, B. A. and Tucker, L. R. (1979). Erratum for the distribution of chance congruence coefficients from simulated data. *Psychometrika*, 44, 365.

Nesselroade, J. R. and Baltes, P. B. (1970). On a dilemma of comparative factor analysis: A study of factor matching based on random data. EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, 30, 935–948.

Nesselroade, J. R., Baltes, P. B., and Labouvie, E. W. (1971). Evaluating factor invariance in oblique space: Baseline data generated from random numbers. *Multivariate Behavioral Research*, 6, 233–241.

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

Tucker, L. R. (1951). *A method of synthesis of factor analysis studies* (Personnel Research Section Report No. 984). Washington, DC: Department of the Army.

Wrigley, C. and Neuhaus, J. O. (1955). The matching of two sets of factors. *American Psychologist*, 10, 418.