Educational and Psychological Measurement

http://epm.sagepub.com/

Significance Testing of Congruence Coefficients: A Good Idea?

Ernest C. Davenport, Jr.

Educational and Psychological Measurement 1990 50: 289
DOI: 10.1177/0013164490502007

EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT
1990, 50

SIGNIFICANCE TESTING OF CONGRUENCE COEFFICIENTS: A GOOD IDEA?

ERNEST C. DAVENPORT, JR.
The University of Minnesota

Tucker's congruence coefficient is often used to compare the equality of latent structures on a given test for different subgroups. Initial use of the index was subjective; given the same congruence coefficient, one investigator could decide that a pair of factors was similar, whereas another investigator could decide that they were different. Critical values were subsequently developed for this index to provide an objective basis for decision making. In practice, these values too often suggest congruence between factors known to differ. Results of the present study suggest that significance tests of congruence coefficients are inappropriate as currently performed.

WHEN designing a test, the test developer should ensure that the latent structure is the same for all groups that must take the test. If latent structures differ for different populations of individuals, construct validity of the test is questionable. It is difficult to argue that a test measures specific attributes when it measures different attributes for different groups of examinees. The purpose of this paper was to explore properties of one index used to help determine the equality of latent structures for different subgroups on a common set of test items—namely, Tucker's congruence coefficient. The paper begins by defining this index. The definition is followed by a description of the use of congruence coefficients with special em-

Correspondence concerning this article should be addressed to Ernest C. Davenport, Jr., Department of Educational Psychology, University of Minnesota, 320 Burton Hall, Minneapolis, MN 55455.

I wish to thank Dr. James Terwilliger for assisting with this project.

TABLE 1
Illustration of Perfectly Congruent Factors

Item	Factor 1	Factor 2
1	0.03	0.018
2	0.10	0.060
3	0.08	0.048
4	0.15	0.090
5	0.60	0.360
6	0.55	0.330
7	0.85	0.510
8	0.22	0.132

phasis on significance tests for the index. The paper ends with an example that illuminates points raised in the discussion.

Definition of the Congruence Coefficient

The congruence coefficient (Equation 1) measures the qualitative similarity of two factors by ascertaining the degree to which the magnitude of factor loadings on the target factor ($F1$) has a pattern similar to the magnitude of loadings on the comparison factor ($F2$). Factors $F1$ and $F2$ are congruent when items which have high loadings on $F1$ have relatively high corresponding loadings on $F2$, while items with small loadings on F1 also have relatively small loadings on $F2$. For perfect congruence, the corresponding loadings for both factors must be equal within a constant of proportionality.

$$r_c(f_1,f_2) = \frac{f_1'f_2}{\sqrt{(f_1'f_1)\,(f_2'f_2)}} \tag{1}$$

(The equation is in vector notation. The factor loadings for the target factor are represented by $f1$ and loadings for the comparison factor are represented by $f2$.)

Both factors in Table 1 are qualitatively the same. That they are defined by the same items (5, 6, and 7) leads to similar interpretations for each factor. Factor 2 is simply a weaker version of Factor 1. In fact, the factors are perfectly congruent, as multiplying each loading in Factor 1 by 0.6 yields the corresponding loadings for Factor 2. One should note that the congruence coefficient for this pair of factors is 1 as expected (calculations not shown).

The congruence coefficient is similar to Pearson's product-moment correlation. Both indices measure the relatedness of scores on a pair of vectors. They also have similar formulas which lead to equivalent results when the mean for both vectors is zero. They differ in that Pearson's correlation measures the degree to which vectors are related under permissible transformations for an interval scale ($f2 = af1 + b$). Congruence coefficients measure a more stringent relationship; the degree to which vectors are related under permissible transformations for a ratio scale ($f2 = af1$).

Conventional Application of Congruence Coefficients

The congruence coefficient is used to determine whether an 'interpretable' factor (latent trait) defined for one group of examinees is present in the structure derived for another group of examinees. Given a common set of test items, an investigator obtains independent factor solutions for each group. Next, each structure is independently rotated (usually by the Varimax procedure) to obtain interpretable factors. Finally, congruence coefficients are computed for factors of one structure versus factors of the other structure. The magnitude of the congruence coefficient is used to determine which factors are similar for the two groups of subjects. When the congruence coefficient is high, factors are deemed congruent.

Originally, this decision was subjective. Given the same congruence coefficient, different investigators could arrive at different conclusions concerning the equality of a pair of factors. Subsequent development of critical values (Cattell, 1978; Korth and Tucker, 1975) provided an objective basis for making decisions. However, there is a problem with using these critical values. Too often the magnitude of the congruence coefficient is significant for factors that are known to differ. Cattell (1978. p. 254) states: "In using r_c it will be found that rather the more r_c's between factors that one definitely *knows to be different* reach indices of *moderate significance* than one would expect to find."

Exploration of the Problem

To explain the excessive number of Type I errors suggested by these critical values, one needs to explore characteristics of r_c relative to the significance test. The fact that higher congruence values suggest similar factors along with the convention of testing statistics against the value 0 sets up the false impression that the significance test is Ho: $\rho_c = 0$ versus Ha: $\rho_c \neq 0$. The formula for r_c shows that it is expected to be 0 when the pair of vectors is orthogonal. However, Korth and Tucker's (1975) critical values were developed for independent vectors, a less stringent condition (Rodgers, Nicewander, and Toothaker, [1984]).

Distinct independent vectors need not be orthogonal. They can have varying degrees of congruence. Intra-structure factors are by definition distinct. But, within-structure congruence for rotated vectors is often non-zero. Furthermore, the expected magnitude of r_c differs by the nature of the content areas represented in the factor structure. One expects higher congruence between distinct factors for a battery of tests within a content area than between factors created from a battery of tests over several content areas. Given that the expected magnitude of the congruence coefficient (between distinct factors) varies depending on the relationship of the items, one should formally account for these differences when developing critical values. Thus, one set of critical values will be insufficient. The magnitude of the critical values should be increased when comparing structures based on a single content area, whereas the values should be adjusted downward when comparing factors derived from a battery that covers several content areas.

To provide conceptual justification for the previous statements, one needs to remember what it means (cognitively) for factors to be orthogonal (i.e., have zero congruence). The congruence coefficient is zero, if the inner product of the vectors is zero. For this event to occur (with non-trivial vectors) at least one of the vectors must have both positive and negative loadings. However, it is not surprising for almost all factor loadings of a rotated structure to be positive (with the remaining loadings being effectively zero). The concept of a general factor and the belief that a high level of performance in one cognitive area should not lead to a diminished level of performance in another leads to an expectation of positive manifold for test data. Furthermore, the extent to which the structure exhibits positive manifold is a function of the "relatedness" of the content areas represented by the factors. As it is mathematically impossible for structures with all positive loadings to have inner products of zero between distinct factors within a structure, these factors will have nonzero congruence between their vectors.

This paper begins by stating that the congruence coefficient was developed to ascertain inter-structure similarities. Yet, the preceding discussion focuses on expected congruence for intra-structure comparisons. The reason for that emphasis is that intrastructure relationships influence the expected congruence of interstructure factors as illustrated with the following example.

Example

Structures A and B, each consisting of three factors (A1, A2, A3, and B1, B2, B3), are assumed to be equivalent. Given principal axis structures, one expects congruence only between corresponding factors. Thus, one anticipates congruency between factors A1 and B1, A2 and B2, and A3 and B3. Also, as factors within a structure are orthogonal one would predict B1 (because it is significantly related to A1) to be unrelated to both A2 and A3. Likewise, little congruence is expected between the other noncorresponding factors (B2 with A1 or A3 and B3 with A1 or A2). If these equivalent principal axis structures are rotated such that factor loadings within structures become related, Factor A1 is still expected to be related to Factor B1. However, if factors A1, A2, and A3 become related, Factor B1 is likewise anticipated to be related to factors A2 and A3. Factor B2 will still be related to Factor A2. However as Factor A2 is related to factors A1 and A3, Factor B2, is expected to be similarly related to factors A1 and A3. Finally, factors B3 and A3 are expected to be related, but Factor B3 is also predicted to be related to factors A1 and A2.

The preceding example is illustrated in the following analysis using data from National Assessment of Educational Progress (NAEP). NAEP is a national organization created to track the educational progress of the nation: a task accomplished by periodically testing representative samples of the nation's youth in several learning areas. The current study uses a subset of data from the special 1976 mathematics assessment. This subset of data is from package 2, 13-year-old, while males, in the eighth (modal) grade.

The 1370 eligible subjects are randomly assigned to two groups of size 685. For each subject scores of 0 or 1 were assigned to the 46 questions in the mathematics assessment depending on whether the subject's response to the question was incorrect or correct, respectively. An inter-item matrix of tetrachoric correlations was computed separately for the two groups of subjects on 36 questions. Ten of the 46 items were deleted, because of extreme marginals. (At least 94% of the students answered each of these items correctly.) The resulting correlations were then factor analyzed separately for each group. As subjects were assigned randomly to the two groups, it is reasonable to predict that the resulting factor structures should be equivalent. The resulting structures are shown in Table 2.

Next, intra- and inter-structure congruence coefficients were calculated for the factors belonging to both groups. Table 3 shows results for the principal axis solutions. As predicted, the loadings on Factor 1 for both groups are highly related. Also, both Factor 2's are congruent and both Factor 3's are similar. The r_c values for each of these comparisons are significant, 0.99, 0.74, and 0.57 respectively (the critical values being indicated in Cattell [1978], Korth and Tucker [1975]). One should note, too, the lack of congruence for

TABLE 2
Factor Structures

	Group 1						Group 2					
	Principal Axis			Varimax			Principal Axis			Varimax		
ITEM	F1	F2	F3	V1	V2	V3	F1	F2	F3	V1	V2	V3
Q02	34	31	10	30	−6	35	34	18	43	−1	18	55
Q03	64	12	−13	55	27	27	48	11	8	32	19	34
Q04	35	−6	27	3	24	37	37	9	15	19	17	33
Q05	56	20	11	38	15	44	53	24	6	39	11	42
Q06	48	−14	−10	30	39	12	44	0	−8	37	20	15
Q07	44	12	−3	36	16	25	41	9	−10	38	11	17
Q08	66	4	10	38	34	43	62	8	−5	50	24	30
Q09	62	9	−35	65	28	7	54	5	−10	47	21	21
Q10	59	−9	−42	59	42	−6	57	−5	−27	57	25	5
Q11	47	9	22	21	19	44	25	24	29	4	3	45
Q13	65	1	−21	55	36	17	58	14	−12	52	15	26
Q15	62	17	10	41	21	46	54	43	8	42	−4	55
Q17	57	3	−7	42	30	24	53	16	12	33	19	42
Q19	31	14	−10	33	6	13	39	13	−10	37	7	19
Q20	61	9	−22	56	28	17	61	8	−22	60	18	18
Q21	29	16	−6	30	3	16	33	−19	15	10	36	14
Q22	41	8	2	29	17	25	39	8	5	26	15	26
Q23	31	−13	6	10	28	16	36	30	48	0	11	66
Q24	67	10	−5	51	30	34	59	−1	−12	50	27	18
Q25	64	−16	−2	35	50	26	45	−33	19	13	56	15
Q26	28	14	0	25	4	19	17	−8	−20	24	8	−10
Q29	50	−1	32	12	29	50	42	37	4	35	−6	44
Q30	41	−9	2	21	31	19	38	2	−15	38	13	8
Q31	39	8	39	4	15	53	44	17	20	22	16	44
Q32	65	−2	3	39	38	35	54	4	3	38	26	29
Q33	60	21	−11	55	17	29	68	8	−46	81	14	5
Q34	36	23	18	23	2	40	34	6	0	25	13	19
Q35	76	−44	15	19	79	36	64	−53	20	22	82	13
Q36	62	−50	2	15	76	16	66	−64	2	34	85	−5
Q37	63	−47	−3	21	74	14	67	−48	9	33	76	9
Q38	48	5	26	17	23	46	41	2	3	28	21	22
Q39	50	13	5	35	17	34	47	14	−7	41	11	25
Q40	49	16	−19	50	15	16	47	9	−28	54	8	8
Q41	60	−11	14	25	43	37	60	−30	25	20	63	27
Q42	39	−7	−20	34	29	2	46	−4	−13	41	23	10
Q46	56	13	7	37	21	38	44	3	−1	33	20	21

Note.—Loadings are multiplied by 100 with 2 significant digits kept.

noncorresponding factors; −0.17 for Factor 2 of Group 1 with Factor 3 of Group 2 being the strongest. These results suggest a one-to-one correspondence between the factor structures (as expected, because the examinees were randomly assigned to the two groups).

The results for the principal axis structures are as expected (i.e., congruence between corresponding factors and no congruence between noncorresponding factors). Table 4 shows the inter- and

TABLE 3
Principal Axis Congruence Coefficients

		Group 1		Group 2		
		M12	M13	M21	M22	M23
Group 1	M11	0	0	99	1	2
	M12		0	−1	74	−17
	M13			−2	14	57
Group 2	M21				0	−0
	M22					0

Note.— All values are multiplied by 100. With respect to M_{ij} the first subscript represents the group number (1 or 2) and the second subscript represents the factor number (1, 2, or 3). Thus, M23 identifies the third factor for the second group of subjects. Values inside the box are inter-structure congruence coefficients. Values outside the box are intra- structure congruence coefficients.

intra-structure congruence coefficients for the Varimax rotated structures. One may observe, again, the high congruence between the ith factor of the first structure with the *i*th factor of the second structure. The congruence coefficients were 0.94, 0.91, and 0.87, respectively. Unfortunately congruence coefficients for the *i*th factor of Group 1 with the jth factor of Group 2 ($i \neq j$) are also high. In fact, each factor in the first structure is significantly related to each factor in the second. Instead of facilitating one's decision, such results increase confusion.

The Varimax results can be explained as follows. The structures are expected to be equivalent. Thus, intra-structure relationships of the factors will be mirrored in inter-structure comparisons between the structures. Given that the intra-structure congruence between Factors 1 and 2 for the first group is 0.73 and that the structures should be equivalent, one would anticipate a similar degree of congruence between Factor 1 for the first group and Factor 2 for the

TABLE 4
Varimax Congruence Coefficients

		Group 1		Group 2		
		M12	M13	M21	M22	M23
Group 1	M11	73	72	94	62	69
	M12		68	76	91	58
	M13			74	60	87
Group 2	M21				61	63
	M22					44

second group (i.e., M11 and M22). Similarly one would expect a congruence coefficient close to 0.72 between Factor 1 for the first structure and Factor 3 of the second. (The observed values are 0.62 for the comparison of M11 and M22 and 0.69 for the comparison of M11 and M23, respectively). As foretold, the relationship of factors within structures influences the expected relationship of factors between structures.

Conclusion

Results of the preceding example are extreme. They are caused (in part) by the fact that all items were from a single content domain, mathematics. Expected congruence between noncorresponding factors should be smaller from a battery covering a wider range of topics. Diversifying the content would decrease the extent of positive manifold exhibited by the data.

However, the major point has been illustrated. In addition to the other factors taken into account to develop critical values for r_c (e.g., number of items, number of factors), one should consider the interrelationship of intra-structure factors. The amount of positive manifold within a structure influences the expected relationship of factors between structures. As this component was not considered in the previous attempts to develop critical values for the congruence coefficient, these critical values are inappropriate as currently applied to rotated factors.

REFERENCES

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

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

Rodgers, J. L., Nicewander, W. A., and Toothaker, L. (1984). Linearly independent, orthogonal, and uncorrelated variables. *The American Statistician*, 38, 133–134.