PSYCHOMETRIKA—VOL. 29, NO. 3
SEPTEMBER, 1964

MEASURES OF INVARIANCE AND COMPARABILITY IN FACTOR ANALYSIS FOR FIXED VARIABLES

SAMUEL R. PINNEAU

SAN FERNANDO VALLEY STATE COLLEGE

AND

ALBERT NEWHOUSE

UNIVERSITY OF HOUSTON

New procedures are presented for measuring invariance and matching factors for fixed variables and for fixed or different subjects. Two of these, the coefficient of invariance for factor loadings and the coefficient of factor similarity, utilize factor scores computed from the different sets of factor loadings and one of the original standard score matrices. Another, the coefficient of subject invariance, is obtained by using one of the sets of factor loadings in conjunction with the different standard score matrices. These coefficients are correlations between factor scores of the appropriate matrices. When the best match of factors is desired, rather than degree of resemblance, the method of assignment is proposed.

Determining the extent to which factors obtained in different studies are the same has been a problem which has plagued factor analysts since Thurstone [11] proposed that the generality (or invariance) of factors is a major goal in an adequate factor analytic program. The difficulty of the task becomes apparent when one considers that a different problem is involved if there is a change in tests, subjects, or both, or in the method of analysis. Methods of measuring invariance have been critically evaluated by Henrysson [6], Wrigley [13], and Harman [5]. A number of these are considered in the present paper in relation to new methods of measuring invariance.

For N individuals and m tests, the complete system of factor equations is expressed by

$$Z = FA.$$

In this equation Z is an N by m matrix of standard scores for N subjects on m tests, F is the matrix of factor scores for the N subjects on q factors, and A is the matrix of loadings of the q factors on the m tests, where q is the rank of the m by m intercorrelation matrix $Z'Z$. The factor score matrix is given by the equation

$$F = ZA^{-1},$$

where A^{-1} is the general inverse [9] or pseudoinverse [3] of matrix A. If the factor axes are rotated by a given amount, i.e., if the matrix of factor loadings A is postmultiplied by a rotation matrix T of order m, new matrices of factor loadings and factor scores, B and G, respectively, will be obtained.

$$AT = B,$$
$$Z = GB,$$
$$G = ZB^{-1}.$$

The relationship between the scores for a given factor in the rotated matrix G and the scores for the same factor of the unrotated matrix F is the same as that which exists between the factors of the two matrices A and B, since Z is a constant; that is, the degree of correlation among the factors and among the factor scores is a function of the angle of rotation. The correlation between the two factors, if they are unit vectors, is equal to their scalar product, which in turn is equal to the cosine of the angle of separation [5].

Measures of Invariance for Fixed Samples

While the model thus far presented is of limited value for practical application (where a large number of variables is involved), it illustrates at a fairly simple level one possible solution to the problem of invariance, namely, determining the angle of separation between the factors in question.

For the case in which a fixed sample but different variables are involved, Wrigley and Neuhaus [14] have suggested using coefficients of congruence. In the terms used here, and as set forth by Harman ([5], pp. 259–260), this procedure entails computing for the two studies matrices of orthogonal factor scores, F, an N by q matrix, and G, an N by q^* matrix. The factor scores of the one matrix are then related to those of the second matrix in order to obtain the correlation matrix R.

$$R = \frac{1}{N}\begin{bmatrix} F' \\ G' \end{bmatrix}[F \quad G]$$
$$= \frac{1}{N}\begin{bmatrix} F'F & F'G \\ G'F & G'G \end{bmatrix}$$
$$= \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix},$$

in which R_{12} $(= R'_{21})$ is a set of coefficients of congruence.

If the two sets of factor scores are transformed so as to be maximally congruent and orthogonal within each set, the coefficients of congruence are equivalent to canonical correlations. Horst [8] has recently presented a solution for generalizing canonical correlations to m sets of data.

The coefficient of congruence appears to be the most natural method of matching factors ([5], p. 260), and would seem to be a fairly direct measure of the similarity of factors obtained from different sets of data. Before accepting such a conclusion, consider a hypothetical study in which one obtains the standard score matrix Z. Suppose that a second standard score matrix Y is formed by rearranging the rows of the matrix Z by a random assignment without, however, changing the row headings. Taking the two matrices as if they contained scores obtained on two different occasions, one would find, if one took from the two matrices scores on the tests with the same names (that is, columns of the same location in the respective matrices), that the test-retest scores would be unrelated, except by chance. However, if matrices of correlations are obtained for the two sets of data and these are factor analyzed, the factor loading matrices will be identical, as will be the correlation matrices, assuming the same rounding and machine errors. On the other hand, the factor scores of the first study will be unrelated to the factor scores of the second study, except by chance. Thus the coefficient of congruence confounds changes in the subjects with changes in the factors.

Put in a slightly different way, in a factor analytic study in which one obtains the same measures on two different occasions on the same sample, one seeks to determine the comparability of the two different methods of weighting the *same* measures for obtaining scores on composite variables (factors). In using the coefficient of congruence, however, one is actually comparing the two different methods of weighting when applied to the different sets of measures. The means of avoiding this problem are evident. The two factor score matrices F and G are obtained by using different sets of factor loadings A and B but the same standard score matrix Z; i.e.,

$$F = ZA^{-1},$$
$$G = ZB^{-1}.$$

The Coefficient of Invariance—Different Samples

The correlation of the scores on one factor from the first matrix F with scores on a factor from the second matrix G is termed the *coefficient of invariance* r_{IV}, where the subscript I stands for invariance and the subscript V stands for fixed variables. This coefficient reflects both the angle through which one would need to rotate the second factor in order for the factors to coincide and the differences between them at that location in factor composition, i.e., configuration of the loadings. (The angle of separation by itself provides an overestimate of the invariance of the factors. In computing such an angle one would need to assume that the factor spaces for the two factor matrices were the same, an issue to which attention is directed in a later section.)

If, prior to the computation of the coefficients of invariance, one of the four methods set forth by Horst [8] is used to obtain a transformation of the two factor score matrices F and G to maximum congruence while at the same time insuring mutual orthogonality of the factors within each set, the coefficient will reflect only the difference in factor composition. A different designation, say the coefficient of factor similarity, would seem to be appropriate in this instance, since one has discarded the original criteria for positioning the vectors. In essence, the question has been changed from how invariant are the factors from one set of data to another to how similar they can be made to be by rotation. A comparison of the coefficient of similarity for the two factors with the cosine of the angle of rotation required to maximize their congruence indicates the extent to which their dissimilarity is a function of the positions of the vectors and the extent to which it is a function of differences in composition of the factors. Changes in subjects and changes in the methods of weighting the variables are, of course, confounded in the coefficient of similarity if the factor score matrices are based on different standard score matrices rather than on the same matrix.

While the logic of the coefficient of invariance was developed using for illustrative purposes matrices of data on the *same variables* and *subjects* on two different occasions, its application is not restricted either to longitudinal data or to data on only two samples. It is directly applicable to data collected on the same variables on different samples. On the other hand, as developed thus far, the coefficient of invariance is not readily applicable to data for which Wrigley and Neuhaus [14] developed their coefficient—different variables, fixed samples.

With respect to the coefficient of invariance, the question arises as to whether values obtained using A and B with Y would be different from those obtained using A and B with Z. Slight differences in the results are certainly to be expected as a function of chance fluctuations in the shapes of the distributions for the variables and of the differences in the homogeneity of the samples (or of the sample in question on two occasions). Greater than chance differences would certainly raise a question as to the comparability of the samples.

Configurational Invariance

The coefficient of invariance which has been presented reflects what Henrysson [6] has called configurational invariance, since it is not sensitive to differences in the numerical size of the loadings. A measure of relationship reflecting numerical differences, as well as configurational changes, is the intraclass correlation, which may well be a useful measure when a coefficient of numerical invariance is desired [4].

Tucker's Coefficient of Congruence

It may be desirable to contrast the coefficient of invariance with the most frequently suggested measure of invariance for fixed variables and different samples, that is, Tucker's coefficient of congruence ([12], p. 43). This measure is the sum of the cross products of the loadings for the two factors under consideration divided by the square root of the product of the sums of the squared loadings. Let

$$A^{\text{normed}} = \begin{bmatrix} \dfrac{1}{\sqrt{\sum_{i=1}^{m} a_{i1}^2}} & & \\ & \ddots & \\ & & \dfrac{1}{\sqrt{\sum_{i=1}^{m} a_{iq}^2}} \end{bmatrix} A,$$

and similarly for B^{normed}; then $\phi = A^{\text{normed}} (B^{\text{normed}})'$, and its elements are given by

$$\phi_{ij} = \frac{\sum_{k=1}^{m} a_{ki} b_{kj}}{\sqrt{\sum_{k=1}^{m} a_{ki}^2 \sum_{k=1}^{m} b_{kj}^2}}.$$

(Note: If A and B consist of q and q^* rows of normal orthogonal eigenvectors, then the denominator of ϕ_{ij} is equal to one.) This is also the procedure which has been recommended by Burt ([1], p. 185) and by Wrigley and Neuhaus [14]. As Harman points out ([5], p. 258), this measure is not a correlation, since the raw loadings used in the formula are not deviates from their respective means and the summations are over the number of variables rather than the number of individuals.

Tests of significance are not available for this index, and its interpretation is ambiguous. Consider as an example the sets of loadings obtained on four variables in two separate studies—.01, .03, .02, .04, and .91, .93, .92, .94, respectively. The patterns are identical, although the points are separated by a constant of .90. In this case the coefficient of congruence is only .92. Consider as a second example two sets of factor loadings whose values are as follows: .2, .4, .1, .3, and .3, .2, .4, .1. In this case the coefficient is .70. Indeed, the lowest value the coefficient can take for the "poorest" configuration of the second set of loadings (i.e., that pattern which results in the smallest sum of cross products) is .67. Thus, factors whose loadings are of the same sign will have high coefficients of congruence. The consequence of this for centroid factor analysis is that the first factors from two different matrices (for fixed variables and different subjects) will almost always have high coefficients of congruence because of the high proportion of large positive loadings.

Correlation of Loadings

Correlating factor loadings also gives ambiguous coefficients. Consider, for example, a factor from matrix A which has loadings varying between .00 and .85, and a factor from a second matrix B whose distribution of loadings has the same shape but whose loadings vary between $-.85$ and $+.85$. Since in the process of computing correlations one converts the raw scores into standard scores, one would be giving equivalent standard score value to loadings of .00 on the first factor and to the very high negative loadings of $-.85$ on the other. One thus equates a variable which contains none of the common variance of the factor with one which shares a great deal of the common variance of the factor on which it loads. To avoid equating loadings which have quite different meanings, the correlation between the squared loadings could be obtained. This procedure presents an equally difficult problem. Loadings which previously had opposite meanings are equated, e.g., $+.85$ and $-.85$.

Horst's suggestion, in his development of relations among m sets of measures ([7], p. 133), that the rows of his supermatrix are variables and the columns are factors, i.e., that the supermatrix should consist of a combined matrix of two or more sets of factor loadings, appears to be similarly limited by such considerations. Means of avoiding this limitation were suggested at the beginning of this section. The problems presented by Tucker's coefficient of congruence and by the correlation of factor loadings are avoided in the coefficient of invariance.

Indices of Proportionality

To use this procedure, according to Cattell's account, "It is necessary that the two experiments have the same variables and yield the same factors *but that the variance of each factor in one experiment shall be different,* through accidents of sampling or through deliberate manipulation of experimental conditions, *from that of the corresponding factor in the second*" ([2], p. 246). When this is the case, the two factors are rotated to the one position in which the loading pattern of one factor corresponds to the loading pattern of the similar one of the other study. (Numerically, of course, the loadings of the two factors may differ at this position.) It is fairly evident that when this rotation is effected, one discards whatever criteria one established for positioning the vectors in the first place, a criticism which also applies to canonical correlations and, as earlier noted, to the coefficient of factor similarity. In addition, some degree of subjective judgment may on occasion be required in deciding which factors to rotate to proportionality.

In contrast, the coefficient of invariance can be meaningfully applied whether or not the variance of each factor in one study is different from that of the corresponding factor in the other. In the case in which the factors are exactly the same, the matrix of coefficients of invariance is a permutation matrix. The use of the coefficient of invariance does not require modifying the position of the vectors or dropping the earlier criteria of rotation, nor does it require that the factors be the "same" either in number or in apparent similarity.

If factors of greater generality are desired, as appears to be the goal for which Cattell strives, this "collapsing" of the factors of the two studies can be achieved by a factor analysis of the two sets of factor score matrices combined; for example, of the $N \times q$ matrix F and the $N \times q^*$ matrix G:

$$[F \mid G] = Z = F^*A,$$

where the combined matrix Z is an $N \times q + q^*$ standard score matrix, A is the matrix of loadings of the $q + q^*$ "tests" on the p new factors, and F^* is the matrix of scores of N subjects on these factors. Since F and G are based on the same standard score matrix rather than on different standard score matrices, the major factors obtained will be composites of the congruent factors.

Matching Factors and the Assignment Problem

Using the matrix of invariance coefficients, one can match the factors from the two studies on the basis of the high relationships. While using the coefficient in this manner would be as defensible as other methods used in this way, such an interpretation would seem to be somewhat naive since the coefficient reflects the *degree* of invariance. Thus the correlation of scores on a given factor from one matrix with scores on the factors in the other matrix shows the extent to which the factor in question is similar to each of the factors of the other matrix.

As an alternative to using the coefficient of invariance as a procedure for obtaining the best match between factors, when the same number are available in both sets, the writers propose the *method of assignment.* (Silver has presented an algorithm for the assignment problem [10].) For factors which should be paired, the differences between loadings should theoretically be constant. In practice, of course, this will not be the case; however, the standard deviation of the differences should be small. To establish the basis for matching factors, let σ_{ij} be the standard deviation of the differences between the loadings of the ith factor of the first matrix and the jth factor of the second set. Since only one factor in the first may match one factor in the second set, one takes a set of σ_{ij}'s ($i = 1, \cdots, q$, and j is a permutation of the values $1, \cdots, q$) such that the sum of the squares is a minimum; i.e., so that the sum of the variances, Σv_{ij}, is a minimum. Computationally, this is the assignment problem for the $q \times q$ matrix of $V = (v_{ij})$. (Whether or not a unique solution has been obtained could be determined by increasing the magnitude of the selected σ_{ij}'s one at a time, determining whether another set of σ_{ij}'s is obtained whose Σv_{ij} is as small as that originally obtained. In actual practice the original solution for a given set of data will, except in a very rare instance, be unique.) This procedure as currently developed is only applicable in instances in which the same number of factors are obtained from the two matrices. When this is not the case, the procedure would need to be modified.

Where a match for every factor is desired, but it is not required that one factor in one set match only one factor in the other set, the assignment could be made so that those showing the smallest standard deviation of differences are paired (or grouped).

Comparability of Factor Space

In this section a procedure is presented for measuring the extent to which the same factor space is occupied by factors obtained on data from different samples of subjects, the factors being treated as vectors of unit length. The same subjects and variables are utilized in computing the factor scores for the two orthogonal loading matrices. Hence, one can combine the two factor score matrices so that for each subject the number of scores Q is equal to q, the number of factors of the first matrix, plus q^*, the number of factors of the second matrix. (The number of factors from the two matrices need not be the same.) If the complete matrix of intercorrelations between factor scores, R_{IV}, with order $q + q^*$, is determined, it will be as follows.

$$R_{IV} = \begin{bmatrix} R_{qq} & R_{qq^*} \\ R_{q^*q} & R_{q^*q^*} \end{bmatrix}.$$

Matrix R_{qq} and matrix $R_{q^*q^*}$ consist of the intercorrelations among the factor scores for the first and second study, respectively. Assuming that the intercorrelations for each of the two matrices are zero,

$$R_{IV} = \begin{bmatrix} I_{qq} & R_{qq^*} \\ R_{q^*q} & I_{q^*q^*} \end{bmatrix},$$

and concern focuses on the matrix R_{q^*q}, which is equal to R'_{qq^*}. If the coefficients of invariance in R_{q^*q} and R_{qq^*} are squared and their columnar sums are obtained, one obtains for each factor a measure of the extent to which its variance can be accounted for by factors of the other study. Thus one can use the coefficient of invariance to show the extent to which two factors from different studies share common variance and, when the factors are orthogonal, to show the extent to which the variance of a given factor is common to the total variance of the second study. In addition, assuming orthogonal factors in both matrices, the sum of all the squared coefficients in either of the two matrices provides a measure of the extent to which the factor space occupied by the factors in one study is comparable to that occupied by those in the second study.

In the case of the coefficient of similarity, all nondiagonal values of matrix R_{q^*q} are zero and the coefficients in the diagonals directly reflect the extent to which the congruent factors share common variance. The sum of the squared values in the diagonal in this matrix provides a measure of the extent to which the two sets of factors occupy the same factor space.

Different Factor Analytic Solutions

The logic basic to the invariance coefficient R_{IV} has been set forth for the special case in which the number of factors obtained is equal to the rank of the matrix. In practice, this is rare in the behavioral sciences regardless of which method of extracting factors is used. Further complicating the problem is the frequent assumption that in terms of factors the total variance of statistical variables consists of three components—common variance, specific variance, and error variance. If this assumption is made, the total number of factors exceeds the number of variables, and an inverse does not exist for the factor loading matrix. One may, however, by using the general inverse or pseudoinverse matrix obtain a unique solution in the least squares sense [3].

Alternatives include the complete estimation method of linear regression [5] and procedures utilizing the reproduced correlation matrix, i.e., the total common variance of the factors extracted. The order of matrix used is equal to the number of factors derived and hence is typically much smaller than that of the matrix whose order is equal to the number of variables involved. The assumption is made that the observed correlational matrix would be equal to the reproduced correlational matrix provided no specific or error variance were involved. The reproduced correlational matrix is then factor analyzed with ones in the diagonals, factor scores being derived as earlier set forth. Factor loadings of the variables on the new factors must then be obtained. Since the new factor loadings represent rotations of the vectors formerly obtained, the task becomes one of defining the rotation matrix required to transform the old factor loadings into loadings on the new factors.

Subject Invariance—Reliability and Stability

In the development of the coefficient of invariance, individual differences were kept constant while variations were permitted both in the number of factors and in the values of the factor loadings. When the two original data matrices contain scores obtained not only on the same variables but on the same subjects, an additional question is raised; how invariant (reliable or stable) from one occasion to the other are the subjects' scores on the obtained factors? The coefficient of congruence provides one estimate of the invariance of the subjects' scores. However, the same problem is present in using it as a measure of reliability or stability as was present in using it as a measure of factor invariance; changes in the factor loadings are confounded with subject change. As an alternative, the writers propose using a measure which employs factor score matrices derived using the same set of factor loadings but the two original standard score matrices. In this way, the "item weights" are kept constant for the same variables in the two matrices. Hence,

$$F = ZA^{-1},$$

$$F^* = YA^{-1}.$$

The correlation of scores on one factor from matrix F with scores on the same factor from the second matrix F^* is termed factor score reliability or, when longitudinal data are involved, the coefficient of subject invariance, r_{SI}, where the subscript S stands for fixed subject, and the subscript I stands for invariance in the factor loadings, i.e., fixed factor loadings.

Some variations would be expected in the values obtained if the coefficients of subject invariance were based on the matrix of factor loadings B rather than on A. Greater than chance differences are to be expected, of course, when the factors differ in composition or in number. The procedure is also applicable when there are n sets of measures rather than just two.

Summary

The coefficient of invariance is a measure of the degree of invariance of factors obtained from data collected on the same variables on the same or different subject samples. Basic to the computation of the coefficient are the two sets of factor scores obtained when one utilizes the different sets of factor loadings with one of the original standard score matrices. The use of the coefficient of invariance is further extended in measuring the extent to which the variance accounted for by a factor in one set is accounted for by factors in any one of the other sets. In addition, it provides the basis for an over-all measure of the comparability of the space spanned by the different sets of factors. The proposed coefficient of factor similarity is recommended when concern focuses on the extent to which sets of factors can be made congruent rather than on the invariance of factors across studies.

The method of assignment is suggested as an alternative procedure to matching factors on the basis of the highest invariance coefficients. This method is applicable when one wishes to obtain the best match between factors and when the matrices are of the same rank. A modification is suggested when the numbers of factors for the two matrices differ.

When the concern of the investigator is focused on the invariance of the factor scores, i.e., on their stability or reliability, the coefficient of subject invariance is suggested. In the computation of this measure, different standard score matrices are utilized but the factor loading matrix is constant.

REFERENCES

[1] Burt C. Factor analysis and canonical correlations. *Brit. J. statist. Psychol.*, 1949, **1**, 95–106.
[2] Cattell, R. B. *Factor analysis*. New York: Harper, 1952.
[3] Greville, T. N. E. Some applications of the pseudoinverse of a matrix. *SIAM Review*, 1960, **2**, 15–22.
[4] Haggard, E. A. *Intraclass correlation and the analysis of variance*. New York: Dryden, 1958.
[5] Harman, H. H. *Modern factor analysis*. Chicago: Univ. Chicago Press, 1960.
[6] Henrysson, S. *Applicability of factor analysis in the behavioral sciences: A methodological study*. Stockholm: Almqvist and Wiksell, 1957.
[7] Horst, P. Relations among *m* sets of measures. *Psychometrika*, 1961, **26**, 129–149.
[8] Horst, P. Generalized canonical correlations and their applications to experimental data. *J. clin. Psychol.*, 1961, Monograph Supplement No. 14, 331–347.
[9] Horst, P. *Matrix algebra for social scientists*. New York: Holt, Rinehart and Winston, 1963.
[10] Silver, R. An algorithm for the assignment problem. *Comm. ACM*, 1960, **3**, 605–606.
[11] Thurstone, L. L. *Multiple-factor analysis*. Chicago: Univ. Chicago Press, 1947.
[12] Tucker, L. R. A method for synthesis of factor analysis studies. Personnel Research Section Report No. 984. Washington, D. C.: Department of the Army, 1951.
[13] Wrigley, C. S. Objectivity in factor analysis. *Educ. psychol. Measmt*, 1958, **18**, 463–476.
[14] Wrigley, C. S. and Neuhaus, J. O. The matching of two sets of factors. *Amer. Psychologist*, 1955, **10**, 418–419.

Manuscript received 11/15/62
Revised manuscript received 10/12/63